

【改革与发展】

“双一流”建设成效评价中案例指标的特征、问题与反思

朱军文 邵玲芝

【摘要】案例指标已成我国高等教育评价的重要依据。以首轮“双一流”建设周期总结中137所高等学校提交的1353个典型案例为样本,以《统筹推进世界一流大学和一流学科建设总体方案》确定的建设任务和改革任务为聚焦点,分析案例指标总体特征和使用中存在的问题。典型案例在任务覆盖面、结构完整性、成效呈现、证据形式及其来源等方面呈现四个特征,即案例指标指向的建设任务和改革任务不均衡,案例指标构成要素缺失的案例数量少但涉及学校范围广,案例呈现的建设成效与“双一流”建设战略目标有差距,建设成效的证据形式未能完全摆脱已有定量评价的弊端。“双一流”建设战略目标和建设任务决定案例指标在其成效评价中具有适切性,案例指标的使用价值有待进一步挖掘。建设任务的实际成效作为案例的核心部分,如何呈现和评价仍然缺乏标准,亟待研究。案例指标中佐证建设成效的证据形式及其来源亟待拓展,作为评价指标的典型案例需有更为翔实的撰写要求与评分细则。

【关键词】高等教育;“双一流”建设;成效评价;案例指标

【作者简介】朱军文,华东师范大学教育学部教授;邵玲芝,华东师范大学教育学部博士生(上海 200062)。

【原文出处】《教育研究》(京),2022.12.91~100

【基金项目】本文系教育部哲学社会科学实验室上海师范大学、上海市教育科学研究院教育大数据与教育决策实验室2022年度重点资助项目“高等教育评价中的案例指标研究”(项目编号:2022zdsys0001)的研究成果。

近年来,在高等教育评估实践中,案例指标日益受到重视。教育部办公厅发布的《关于开展2016—2020年“双一流”建设周期总结工作的通知》指出,每所高等学校在总结中提交不超过10个典型案例,旨在总结首轮“双一流”建设的典型成效、创新举措和成功经验。2021年,教育部、财政部、国家发展和改革委员会联合印发的《“双一流”建设成效评价办法(试行)》指出,成效评价实行定量评价与定性评议相结合,各高等学校提交的典型案例及其他写实性材料,由专家进行定性评议。此前,教育部学位与研究生教育发展中心(以下简称学位中心)组织开展的第四轮学科评估,已单独设置“社会服务贡献案例”观测点,通过“代表性案例”观测学科的社会服务贡献。2020年启动的第五轮学科评估,将强化典型案例评价作为基本原则,且社会服务案例所占权重进

一步提升。^[1]案例指标已成我国高等教育评价的重要依据。

一、问题提出

从国外高等教育评价实践看,英国在2014年高等教育卓越研究框架(Research Excellent Framework, REF)中首次引入案例指标,以评价科学研究影响力。^[2]随后,荷兰标准化评估指南(Standard Evaluation Protocol, SEP)^[3]、澳大利亚社会互动与影响力(Engagement and Impact, EI)^[4]评价等相继引入案例指标。英格兰高等教育拨款委员会(Higher Education Funding Council for England, HEFCE)等研究发现,案例指标是评价高等学校科研影响力的最佳方式,^[5]能够有效缓解归因问题复杂带来的评价难题^[6]。

采用案例指标评价影响力有效可行,但亦存在不足。^[7]分析英国高等教育卓越研究框架案例,哈兹

拉吉(Khazragui, H.)和哈德逊(Hudson, J.)发现,案例主要是叙述,缺少硬性信息,难以准确反映科研影响力。^[8]案例指标呈现内容千差万别,案例之间无法进行比较。通过对案例指标进行内容分析,格兰特(Grant, J.)和辛里奇斯(Hinrichs, S.)发现,支撑影响力的定量证据多种多样,其使用和表述存在不一致现象。^[9]同时,作为评价指标的案例需要遵循特定格式和规则,这可能导致许多类型影响力无法提供确凿的证据材料。克尼尔(Kneale, P.)等学者发现,参评机构选择案例似乎特别受科学研究与其影响力之间证据的可用性影响,某些无法提供证据的研究领域可能不会被呈现。^[10]曼维尔(Manville, C.)等学者调查访谈了一些参与英国高等教育卓越研究框架案例撰写的工作人员,发现学校提交的案例并未涵盖科学研究的所有影响。^[11]此外,案例指标需借助同行进行评价,而同行评价的固有缺陷可能会影响案例指标的效果。^[12]通过分析英国高等教育卓越研究框架,有学者发现,主专家评估小组成员审查子专家评估小组决定的方式,意味着科研影响力案例的评分可能较少取决于其内在价值,更多取决于两个评估小组决定的一致性。^[13]

为了充分发挥案例作为评价指标的作用,一些学者也提出了改进建议,主张采用更加标准化的信息,使案例指标具有横向可比性。鲍尔(Power, M.)提出,案例指标需要有一个固定的规范核心。^[14]斯塔尔(Star, S.L.)和格里瑟默(Griesemer, J.R.)指出,为实现评价目的,必须对案例指标边界进行限制。^[15]班多拉-吉尔(Bandola-Gill, J.)和史密斯(Smith, K. E.)认为,让案例指标评价行之有效的关键正是它们标准化表现的能力。^[16]通过对2014年英国高等教育卓越研究框架的案例文本进行内容分析,帕克斯(Parks, S.)等人对常见的科研影响力证据提出具体的标准化操作意见。^[17]基于标准化的内涵,笔者认为,案例指标的标准化是对案例叙述的内容信息、影响力类型及其表述、证据形式、评分等级及评价程序等,提出统一要点、统一规格、统一流程的过程。^[18]

二、研究方法

以清晰界定案例指标概念及其效度观测维度为基础,本研究以“双一流”建设战略任务为指向,采用

文本内容分析方法进行分析。

(一)案例指标及其效度观测维度

作为高等教育评价指标的案例,是高等学校在人才培养、科学研究、社会服务或其他类型办学活动等方面,开展的代表性工作及其实际成效或影响力的完整呈现,是按照明确的要求系统阐述代表性工作的目标指向、实施过程、实际成效或影响力的文字表述与证据的集合。高等教育评价案例指标的构成要素,包括代表性工作的执行主体、拟解决的问题和举措等核心内容、成效及影响、支撑证据及其来源等四个部分。课题组此前的研究表明,高等教育评价的案例指标与《深化新时代教育评价改革总体方案》提出的改进结果评价、强化过程评价、探索增值评价、健全综合评价,具有很好的适切性。^[19]“双一流”建设成效评价的案例指标,是各高等学校和建设学科落实总体方案确定的战略任务过程中,完成的典型性工作及其实际成效的完整呈现。作为“双一流”建设成效评价指标的案例,不是一段文采飞扬的文字,而是一个结构完整的标准化例证,应具有包括典型性工作聚焦的关键任务、实施的主要举措、取得的实际成效、可靠的证据等支撑材料在内的基本框架,应具有完备翔实的内容信息、具体可观测的成效类型及其表述、客观可信且易于横向比较的证据数据资料。

“双一流”建设成效评价的案例指标能否准确、有效地呈现各高等学校建设任务和改革任务的完成程度,即案例指标的效度,可从三个维度观测。一是内容维度,即案例指标对建设任务和改革任务的覆盖程度和内容结构完整性。二是实际成效呈现,包括案例受益群体指向是否明确及对受益群体的贡献程度。三是证据形式及其来源,包括案例指标提供的证明建设成效的证据类型及其来源的适切性。

(二)研究样本与清理

本研究以首轮“双一流”建设周期总结工作中137所高等学校提交的1353个典型案例为样本,以《统筹推进世界一流大学和一流学科建设总体方案》确定的建设任务和改革任务为聚焦点,分析案例指标总体特征和使用中存在的问题。首轮建设周期总结工作,有126所高等学校按要求提交了10个典型

案例,1所高等学校提交了11个案例,5所高等学校提交了9个案例,有3所高等学校提交8个案例,有2所高等学校分别提交6个和7个案例。为便于对文本进行内容分析,首先对所有案例的文本格式、内容结构等进行整理,统一案例标题、正文等格式。案例样本字数总计1158815字,平均每个案例约856字。所有案例学校名称均匿名。

(三)数据资料分析

本研究采用NVivo11质性研究分析软件对案例文本进行标记、编码和描述分析,主要包括文本类型建构分析和文本主题分析两阶段编码。

文本类型建构分析阶段的编码工作,是基于总体方案确定的战略任务,提炼归纳案例文本属性和特征,进行推论式建构类目。将整个案例文本作为基本分析单元逐一编码,以形成属性空间,再根据属性相似性,对每个自由节点进行聚类组合,形成建构类型模式。文本类型建构分析阶段的编码结果,形成人才培养、科学研究、社会服务、教师队伍建设、文化传承创新、国际交流合作、制度建设等7个类型。

将1353个案例划分至7个类型,再进行文本主题分析。采用推论式类目建构方法,基于案例指标效度内涵推论得出案例4个主题类目,包括聚焦的关键任务、实施的主要举措、取得的实际成效及提供的支撑证据,并以整个案例作为基本分析单元进行初步编码,将相应的文本段划分至各个主题类目。初

步编码结束,编辑、整合归属于同一主类目的所有文本段,采用开放式编码法分别为4个主类目,归纳式地界定基本维度和创建子类目。

三、研究结果

基于文本内容分析,1353个典型案例在任务覆盖面、结构完整性、成效呈现、证据形式及其来源等方面呈现四个特征。

(一)案例指标指向的建设任务和改革任务不均衡

案例指标以呈现科学研究和人才培养贡献为主,对建设任务和改革任务覆盖不够完整。案例指标指向的建设任务和改革任务分布表明,1353个案例在7个类型建设任务上的数量分布相对不均,各类型案例数量差异较大。能够提供7个类型案例的仅有30所学校。有14所学校提供了4个类型案例,有7所学校只提供了3个类型案例。在典型案例中,属于“科学研究”类型的有429个,所占比例为31.7%。以“科学研究”类型案例为主的高等学校有94所,占比68.6%;其中的1所学校所提供的10个案例中,有7个属于“科学研究”类型。“人才培养”类型有278个,所占比例为20.5%。“社会服务”类型有199个,所占比例为14.7%。“制度建设”类型有122个,所占比例为9.0%。“国际交流合作”类型有119个,所占比例为8.8%。“教师队伍建设”类型有115个,所占比例为8.5%。“文化传承创新”类型有91个,所占比例为6.7%。(见表1)

表1 案例指标指向的建设任务和改革任务分布

类型		案例数或学校数(个)	占总体比例(%)
案例类型	科学研究	429	31.7
	人才培养	278	20.5
	社会服务	199	14.7
	制度建设	122	9.0
	国际交流合作	119	8.8
	教师队伍建设	115	8.5
	文化传承创新	91	6.7
学校类别	7个类型案例	30	21.9
	6个类型案例	50	36.5
	5个类型案例	36	26.3
	4个类型案例	14	10.2
	3个类型案例	7	5.1

注:学校类别根据学校提供的案例指标指向的建设任务和改革任务分布数量划分。

结果显示,有78.1%的高等学校并未充分利用案例指标,更好地呈现学校和学科总体建设方案确定的各项战略任务取得的效果。各类型案例数量分布不均衡,部分类型建设任务和改革任务缺少案例,可能会削弱高等学校间的横向可比性,也可能导致无法对高等学校的部分建设任务和改革任务完成情况进行完整的综合评价。“双一流”建设学校提供的案例数量,建设任务和改革任务分布差异明显,有些高等学校没有用满给定的10个案例额度。

一方面,是因为管理部门未对案例指向的建设任务和改革任务类型分布及其基本规范做出明确要求;另一方面,与建设学校对成效的理解仍然停留在传统的数量化指标相关。一些建设学校和学科还没有充分意识到案例指标在“双一流”建设成效评价中的作用和价值。特别是,在呈现标志性工作成效上的价值。案例指标在建设任务和改革任务分布上的不均衡,包括个别学校对科学研究案例的偏好,一定程度上折射高等学校“重科研轻教学”的观念根深蒂固。

(二)案例指标构成要素缺失的案例数量少但涉及学校范围广

作为评价指标的案例,需要具备典型性工作的关键任务、主要举措、实际成效及支撑证据等四个要素的标准化结构。由于“双一流”建设周期总结工作提供了案例撰写的基本格式要求,案例指标的内容结构总体上较为完整。但部分建设任务的案例指标尚存成效“支撑证据”缺失,与该建设任务的成效难以表述不无关系。

各高等学校提供的案例中有1258个,即93.0%的案例具备基本完整的内容结构,有59所学校提供的95个案例缺乏能够证明建设成效所必不可少的证据。结构不完整案例数量少,但涉及43.1%的建设

学校。没有客观的证据支撑的案例,其中,“制度建设”类型案例有22个,占比18.0%;“人才培养”案例有33个,占比11.9%;“教师队伍建设”案例有10个,占比8.7%;“文化传承创新”案例有6个,占比6.6%;“科学研究”案例有20个,占比4.7%;“国际交流合作”类型有4个,占比3.4%。比较而言,“社会服务”案例的内容结构及要素基本完整。(见表2)

高等教育评价案例是系统阐述典型性工作实施过程及实际成效的文字表述与证据的集合。^[20]部分案例缺乏证据支撑,则难以证明典型性工作与实际成效之间的直接或间接关系,以及实际成效的受益群体范围和贡献程度,可能影响专家判断的客观性,制约了案例指标的有效性。案例指标内容结构不完整的原因,首先,管理部门对如何利用案例指标尚处于摸索阶段,制定的成效评价管理办法缺少操作性指引或操作细则,尚未对案例指标的构成要素、内容要点及评估标准等作详细规定。其次,建设学校和学科对案例指标的理解不够全面,对建设任务及应实现的预期成效缺少前瞻性计划和长期跟踪的数据积累。另外,案例指标主要以文字形式呈现以及表述偏“软”等特点,很容易造成无统一框架、无成效观测点、无优劣比较基准等困惑。

(三)案例呈现的建设成效与“双一流”建设战略目标有差距

案例指标呈现的建设成效,一方面应清晰表明实际成效的受益群体,另一方面应呈现对受益群体产生影响的范围和程度。案例指标基本能够清晰提及受益群体以及对受益群体的深远影响或实际贡献,一些案例同时指向多个受益群体以及对受益群体的多个方面影响或贡献。(见表3)

有78.4%的“人才培养”案例受益群体指向人才培养模式探索,实际贡献则主要涉及学生创新能力

表2

案例指标的内容结构完整性

案例指标 内容结构	占案例总 数比例(%)	占高等学校 总数比例(%)	分建设任务类型的占比(%)						
			人才培养	科学研究	社会服务	教师队伍 建设	文化传承 创新	国际交流 合作	制度建设
完整	93.0	56.9	88.1	95.3	100	91.3	93.4	96.6	92.0
不完整	7.0	43.1	11.9	4.7	0	8.7	6.6	3.4	18.0

注:案例指标的内容结构完整性数据来源于案例4个主题类目的编码结果。

表3 部分居前的案例指标成效受益群体及实际贡献的类型

建设任务	受益群体	实际贡献	比例(%)
人才培养	探索人才培养模式	为思想政治教育改革实践提供借鉴	78.4
		为高等学校教育教学改革提供参考	
	服务学生发展	学生创新能力提升	47.1
		学生学术水平提高	
科学研究	推动科学技术前沿	突破重大科学问题和工程技术难题	80.2
		在国内外学术界产生重要影响	
	服务经济社会发展	为生产提供新产品和新技术	35.0
	探索科研和学科建设模式	引领优势学科和领域建设	29.6
社会服务	服务经济发展和产业转型升级	面向行业企业转移转化科技成果	55.8
		为区域振兴提供科技和人才支撑	
	服务政府决策咨询	为政府部门提供资政专家和报告	37.2
教师队伍建设	教师专业素质与能力发展	科研创新能力提升	87.8
		思想政治素质提高	
		教学业务能力提升	
	探索教师队伍建设路径	在人才引进上取得新突破	68.7
文化传承创新	服务文化繁荣发展	为文化传播提供平台和活动载体	96.7
		为文化传承培育学术作品	
		为文化建设提供咨政建议	
国际交流合作	服务教育国际影响力提升	为人文交流提供平台和活动载体	92.4
		为国家争取学术和科技话语权	
		为全球教育治理贡献中国方案	
		为受援国发展提供科技和人才支撑	
制度建设	服务高等教育建设路径探索	在现代大学制度改革上取得新突破	100
		为高等学校党建工作提供示范引领	

注:依据受益群体和实际贡献比例由高到低,选取部分内容予以呈现。

得到提升、为思想政治教育改革实践提供了借鉴。在“科学研究”类型案例中,有80.2%指向科学与技术前沿发展,有35.0%指向经济社会发展,实际贡献涉及突破重大科学问题和工程技术难题、在国内外学术界产生重要影响、为生产提供新产品和新技术。“社会服务”类型的案例受益群体以服务经济发展和产业转型升级、政府决策咨询为主,实际贡献主要涉及面向行业企业转移转化科技成果、为政府部门提供资政专家和报告、为区域振兴提供科技和人才支撑。“教师队伍建设”类型案例指向教师专业素质与能力发展占比87.8%,实际贡献主要涉及科研创新能力提升、在人才引进上取得新突破等。“文化传承创新”类型案例指向文化繁荣发展占比96.7%,实际贡献主要涉及为文化传播提供平台和活动载

体、为文化传承培育学术作品等。“国际交流合作”类型案例指向教育国际影响力提升占比92.4%,实际贡献主要涉及为人文交流提供平台和活动载体、为国家争取学术和科技话语权、为全球教育治理贡献中国方案。“制度建设”类型案例全部指向高等教育建设路径探索,实际贡献主要涉及在现代大学制度改革上取得新突破等。

典型性工作成效及实际贡献是案例评价的核心。从上述描述性统计分析可见,案例指标呈现的建设任务的受益群体相对明确。但从“双一流”建设服务国家和社会发展战略目标看,案例指标呈现的建设成效仍不准确,建设任务的目标群体指向存有偏离,与一流目标的建设要求还有差距。比如,“人才培养”类型接近八成案例以服务人才培养模式

探索为主。“十年树木，百年树人”，在较短的建设周期，如何对人才培养成效进行更为合理的界定，是案例指标使用的难题。但人才培养类型案例过于集中在培养模式探索，显然也是其局限所在。案例提供的建设任务典型性工作受益群体指向不准确，取得的实际成效表述不聚焦，可能直接影响案例指标在“双一流”建设成效评价中的实际价值。当然，各类典型案例呈现的受益群体及实际贡献表述偏离建设目标任务和战略方向，可能意味着建设学校对案例指标的理解或对重点建设任务理解有偏差，误以条件类、过程类等改善作为建设成效或实际贡献表征。

(四)建设成效的证据形式未能完全摆脱已有定量评价的弊端

客观证据是案例指标不可或缺的组成部分。典型案例用以佐证“双一流”建设成效的证据形式包括奖项类、称号类、排名类等数量型证据，也包括政策咨询、媒体报道等写实性证据，多数案例同时采用多种证据形式。高等学校提供的用以佐证建设成效的证据仍以统计型数据为主，写实性证据表述相对欠缺，未能完全摆脱定量评价中“五唯”惯性思维的桎梏。有1296个案例提供了数量型证据，占比95.8%；其中604个案例仅提供数量型证据。数量型证据中，以教育教学奖励、科研奖励、竞赛奖项等各类获奖证据排名第一；其次是各类人才称号、先进单位等；位居第三的是教学中心、示范基地、重点实验室等基地

类证据；随后是获批重大项目、重点研发计划等项目类证据；部分案例提供了论文、著作、大学或学科排名、授权专利等数量型证据。有749个案例采用了写实性证据形式，占比55.4%，其中有57个案例仅提供写实性证据。写实性证据中以成果获政府批示采纳居首，其次是新闻媒体正面报道，第三是同行高度评价的证据，实现重大成果突破的证据占比16.2%。(见表4)

大学和学科排名是典型案例用以佐证建设成效的重要证据来源，选取标准均以有利性为原则，缺少可比性和连贯性。比如，针对大学或学科排名类证据，不同案例参考的排行榜各不相同，均选取排名表现最好的为证据。有52.9%的案例用到了基本科学指标数据库(Essential Science Indicators，以下简称ESI排行榜)的数据，45.1%的案例使用了软科世界大学学术排名数据，有15.7%引用了学位中心发布的学科评估结果。可见，典型案例对佐证建设成效的证据来源的引用有着明显的各取所需倾向。(见表5)

证据类型难以摆脱“五唯”的窠臼，对大学和学科排名的依赖，从一个侧面体现了“双一流”建设学校在教育评价上，长期以来形成的思维惯性一时难以根本扭转。教育评价改革“破”与“立”并举不容易。办学者和学校管理者需要对建设成效有更为全面深刻的理解，也亟待相关部门和专家研究提出具有质量和贡献属性的高等教育建设成效评价新指

表4 案例指标的证据类型及分布

证据形式		案例数	占案例总数比例(%)
数量型证据	奖项类	803	62.0
	称号类	767	59.2
	基地类	633	48.8
	项目类	595	45.9
	论文类	200	15.4
	著作类	156	12.0
	排名类	141	10.9
	专利类	119	9.2
写实性证据	政府批示采纳	411	54.9
	新闻媒体报道	199	26.6
	同行高度评价	151	20.2
	成果重大突破	121	16.2

注：案例总数分别按照数量型证据案例数和写实性证据案例数统计。

表 5

案例指标中以排名为证据的情况统计

证据来源	证据类型	占案例总数比例(%)
第三方排名机构	ESI 排行榜	52.9
	软科排名	45.1
	QS 世界大学排名(QS World University Rankings)	17.6
	U.S.News 全球大学排名(U.S.News & World Report Best Global Universities Rankings)	15.7
	自然指数排名	13.7
	泰晤士高等教育世界大学排名(Times Higher Education World University Ranking, THE)	7.8
	金平果排名	5.9
	校友会排名	5.9
	中国大陆最具创新力大学百强榜	2.0
学位中心	学科评估	15.7

注:案例总数根据引用大学和学科排名证据的案例数统计。

标。依赖排名特别是对本校或本学科有利的排名为依据已成普遍现象,昭告案例指标使用中证据选取亟待转换观念。

四、结论与讨论

案例是“双一流”首轮建设成效评价的重要指标,且规定了每所学校提交数量的限额。在强化发展质量与实际贡献的评价改革导向下,案例指标通过对代表性工作的完整呈现,具有传统指标所不可替代的作用,须发挥应有作用。通过对首轮“双一流”建设周期总结中建设学校提供的案例进行文本内容分析,从案例指标的建设任务和改革任务覆盖面、内容结构要素的完整性、建设成效的受益群体指向和实际贡献呈现以及证据形式与来源等方面,可见,案例指标在使用中已经开展有益尝试,也有值得进一步完善的思考。

(一)案例指标在“双一流”建设成效评价中的使用效果有待进一步提升

“双一流”建设战略目标和建设任务决定案例指标在其成效评价中具有适切性,但案例指标的使用价值还有待进一步挖掘。案例指标以写实性方式呈现建设学校在重点建设过程实施的典型性工作的问题导向、原创价值和服务社会实际贡献的特点,切实体现了“双一流”建设以深化内涵发展、彰显优势特色、服务国家急需的导向,与深入推进世界一流大学和一流学科建设的战略目标和建设任务相契合。可

以认为,案例是“双一流”建设成效评价中可依靠的重要指标。

目前,对案例指标的理解还不够全面系统深入,对其在成效评价中如何更好地发挥作用还不够清晰。案例使用可能仍然停留在作为证明材料,而不是独立的评价指标阶段。故此,一些建设学校和建设学科尚不够重视,按照规定每所学校可以提交 10 个案例,但有一批学校没有用满额度。一些学校提供的案例类型集中在科学研究任务部分,未能提供完整涵盖人才培养、科学研究、社会服务、教师队伍建设和、文化传承创新、国际交流合作、制度建设等 7 个类型的典型案例。以一定数量相互补充的典型案例组合为指标,可以较好地呈现总体办学情况,实现充分描述建设方案战略任务和综合评价一所高等学校建设成效的目标。^[21]为了提升案例指标在“双一流”建设成效评价中的使用效率,不仅需要对典型案例总体数量提出要求,也要对案例指向的建设任务和改革任务类型分布有基本规定,从而保障典型案例作为评价指标使用上的完整性。

(二)实际成效作为案例指标构成要素的核心部分仍然缺乏评价标准

“双一流”建设成效评价案例指标总体上具备相对统一且较为完整的内容结构,但建设任务的实际成效作为案例的核心部分,如何呈现和评价仍然缺乏标准,亟待研究。案例指标的天然优势在于,可以

文字形式全面系统阐述典型性工作与实际成效的形成过程,完整呈现典型性工作聚焦的目标或问题、实施思路、主要举措、多种类型成效及其支撑证据等。^[22]要充分发挥案例指标过程本位的显著优势,离不开若干典型性工作构成要素的支撑。

总体上,“双一流”建设成效评价的案例指标基本具备相对统一且较为完整的内容结构,但部分类型案例的实际成效呈现存在受益群体定位偏离、支撑证据要素缺失或不够准确等问题。实际成效是案例指标的核心,有证据支撑的受益群体范围和贡献程度是专家定性评议的主要依据。为提升案例指标在“双一流”建设成效评价中的使用效度,需要进一步明确规定:完整的案例指标结构须由典型性工作的关键任务、主要举措、实际成效及支撑证据四个部分组成。同时,考虑到文字属性偏软容易造成案例呈现的内容和格式参差不齐,也要求对案例指标的每一构成部分内容要点作统一规定。^[23]在典型性工作成效的类型和程度划分表述方面,需要根据战略目标和建设任务,结合高等学校的基本功能深化研究,精准定位受益群体及实际贡献,对不同类型成效及其表述形式进行标准化。

(三)案例指标中佐证建设成效的证据形式及其来源亟待拓展

“双一流”建设成效评价案例指标没有摆脱定量评价中“五唯”的惯性思维桎梏,佐证建设成效的证据形式及其来源较为单一。基于案例指标的评价是同行专家判断。同行专家赖以做出评价决策的依据是高等学校提供的事实性陈述及各类证据。通过分析发现,总体上“双一流”建设成效评价案例指标的支撑证据未能摆脱“五唯”的窠臼和各类排名的束缚,佐证建设成效的证据类型更偏向形式化的定量数据,写实性证据表述相对欠缺。

由于高等教育办学活动本身的复杂性,案例指标中建设成效仍然停留在将学生升学情况、学校和学科排名表现、发表论文数量和刊发期刊等级、获得各类奖项或称号等外在的数量表现作为人才培养、科研贡献、机制创新等建设成效的证明,部分丧失了以典型性工作案例为指标的意义,未能充分发挥其在评价改革中的价值。定量方法不仅偏离了自身所

具有的质量属性,更会加剧高等学校建设“短平快”的肤浅与“急功近利”的浮躁,^[24]导致难以实现高质量内涵式发展要求。案例指标引用的证据来源类型参差不齐,影响了案例所代表的典型性工作之间的横向可比性和专家判断的客观性。为提升案例指标在“双一流”建设成效评价中的使用效度,需要进一步拓展佐证建设成效的证据形式及其来源,对高等教育重点建设任务的效果有更为深刻的理解,对其证据呈现形式、来源类型等作统一规定,包括统一数量型证据的呈现形式以及写实性证据的表达方式,针对同一类型的证据数据资料,统一证据来源和渠道。

(四)作为评价指标的典型案列需有更为翔实的撰写要求与评分细则

“双一流”建设成效评价须完善典型案例撰写要求,并制定案例指标评分细则。对案例提供标准化的撰写指引,是降低案例撰写随意性与专家评价主观性的关键。案例指标需要标准化结构,需要对案例叙述的内容信息、成效类型及其表述、证据形式及其来源、评分等级及评价程序等统一要点、统一规格、统一流程。^[25]英国高等教育卓越研究框架对案例主要构成部分的内容要点、影响力的内涵及其表述形式、证据表述格式等案例撰写要求、评估标准、相关权重等专家评价细则均进行了系统规范。我国首轮“双一流”建设周期总结中对案例撰写已经提供了原则性要求,对案例选取范围有了明确的规定,对案例撰写的内容结构、篇幅和编辑排版等提出了注意要点和统一规定。但未对案例内容各构成要素的核心要点部分做出具体要求,也未对专家基于案例开展同行评价的标准和分值权重给出参考细则。案例指标主要以文字形式呈现,在撰写过程中若无明确具体的内容要求和实施细则,文字表述偏软的特征将使得案例的撰写千差万别,无法进行横向比较。为了推动案例指标在“双一流”建设成效评价中的广泛应用,需要进一步完善典型案例撰写要求,对案例内容的各个构成部分的具体信息、受益群体范围和实际贡献程度、证据形式及其来源等进行规范,做适当的统一和标准化处理。为此,制定一套客观、透明的案例评分指南势在必行。一方面,可为案例

撰写提供参照,也为专家的同行评价提供了指引;另一方面,也可引导建设学校以质量和贡献为导向,推动建设任务高质量发展。

案例指标作为以多种证据支持的对典型性工作及其实际成效的实例叙述,其呈现的建设任务对战略目标的实现程度,仍然依赖同行专家进行判断。“双一流”建设成效评价中案例指标指向的建设任务实际成效大小、任务目标的达成度、与其他建设学校同类型建设任务实际成效高低的比较等,均需要依赖同行专家的评议判断。提升案例指标的标准化程度,既方便同行专家进行评价,也可最大限度降低同行专家评价的主观偏差。

参考文献:

- [1]教育部学位与研究生教育发展中心负责人就《第五轮学科评估工作方案》答记者问[EB/OL].http://www.moe.gov.cn/jyb_xwfb/s271/202011/t20201102_497816.html.
- [2]Higher Education Funding Council for England. Assessment Framework and Guidance on Submissions[EB/OL]. <https://www.ref.ac.uk/2014/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf>.
- [3]Association of Universities in the Netherlands., et al. Standard Evaluation Protocol 2015–2021: Protocol for Research Assessments in the Netherlands[EB/OL]. <https://universiteiten.vannederland.nl/files/documenten/Domeinen/Onderzoek/SEP2015-2021.pdf>.
- [4]Australian Research Council. EI 2018 Framework[EB/OL].<https://www.arc.gov.au/engagement-and-impact-assessment/ei-key-documents>.
- [5]Technopolis Group. REF Research Impact Pilot Exercise Lessons-Learned Project Feedback on Pilot Submissions[EB/OL]. https://www.ref.ac.uk/2014/media/ref/content/pub/refresearchimpactpilotexerciselessons-learnedprojectfeedbackonpilotsubmissions/re02_10.pdf.
- [6]Grant, J., et al. Capturing Research Impacts: A Review of International Practice[R]. Santa Monica: RAND Corporation, 2010.
- [7]Manville, C., et al. Assessing Impact Submissions for REF2014: An Evaluation[R]. Santa Monica: RAND Corporation, 2015;刘兴凯,左小娟.英国大学科研影响力评估机制及其启示[J].中国高教研究,2015,(8).
- [8]Khazragui, H. & Hudson, J. Measuring the Benefits of University Research: Impact and the REF in the UK[J]. Research Evaluation, 2015,(1).
- [9]Grant, J. & Hinrichs, S. The Nature, Scale and Beneficiaries of Research Impact: An Initial Analysis of the Research Excellence Framework (REF)2014 Impact Case Studies[R]. London: King's College London and Digital Science, 2015.
- [10]Kneale, P., et al. REF 2014: Higher Education Pedagogic Research and Impact[J]. Higher Education Academy, 2016.
- [11]Manville, C., et al. Preparing Impact Submissions for REF 2014: An Evaluation: Findings and Observations[R]. Santa Monica: RAND Corporation, 2015.
- [12][18][19][20][25]朱军文,邵玲芝.高等教育评价中的案例指标及其标准化[J].高等教育研究,2021,(1).
- [13]Watermeyer, R. & Chubb, J. Evaluating 'Impact' in the UK's Research Excellence Framework (REF): Liminality, Looseness and New Modalities of Scholarly Distinction[J]. Studies in Higher Education, 2019,(9).
- [14]Power, M. Counting, Control and Calculation: Reflections on Measuring and Management[J]. Human Relations, 2004, (6).
- [15]Star, S. L. & Griesemer, J. R. Institutional Ecology, Translations and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39[J]. Social Studies of Science, 1989,(3).
- [16]Bandola-Gill, J. & Smith, K. E. Governing by Narratives: REF Impact Case Studies and Restrictive Storytelling in Performance Measurement[J]. Studies in Higher Education, 2022,(9).
- [17]Parks, S., et al. Guidance for Standardising Quantitative Indicators of Impact within REF Case Studies[R]. Santa Monica, and Cambridge: RAND Europe, 2018.
- [21][22][23]邵玲芝.“双一流”建设成效评价案例指标结构特征及其有效性研究[D].上海:华东师范大学,2022.
- [24]朱军文,刘念才.高校科研评价定量方法与质量导向的偏离及治理[J].教育研究,2014,(8).