

北京市科技成果评价指标体系参考元素设计与测评

——以基础研究类为例

陶晓丽, 王海芸, 王 新

(北京决策咨询中心, 北京 100089)

摘要:在对北京市近年科技成果基础研究类申报材料进行系统分析的基础上,运用文献研究法和专家访谈法,建立了针对原有成果评价一级指标的主要参考元素和评价依据。应用该评价指标体系,对北京市2014年的申报材料开展了专家测试评价,并对项目测试结果排序和初审结果排序进行了比对,一致性较好。表明建立的具有参考元素和评价依据的评价指标体系具有一定的理论和实践意义,既对基础研究领域研究具有较强的针对性,又可以为评审专家提供更加公平公正合理的评价体系,更加有效地开展科技成果的评审工作。

关键词:北京市;评价指标体系;参考元素;测试分析;基础研究

中图分类号:G311 **文献标志码:**A **文章编号:**1671-1807(2015)07-0080-06

1 研究背景

科技成果评价及奖励是我国科技工作的一个重要组成部分,它对于发挥国家科技政策的导向功能,促进科技工作与经济发展相结合,以及促进科技人员的个人成长等各方面都具有重要作用。为适应科技创新形势的发展,推动科技创新工作的可持续发展,研究完善科学的科技成果评估机制,提出更为理性的评估方法,已成为新时期科技工作的重要任务^[1]。

科技成果评价指标体系是科技奖励评审的重要依据,在科技领域内起到风向标的作用。现行北京市科技成果评价基础研究类的指标在过去几年发挥了积极作用,但在实践过程中也发现了存在一些问题,需要进一步改进和完善。

2 文献综述

关于科技成果评价指标体系已经积累了一定的研究基础并取得了一些阶段性的成果,可以用来参考借鉴。学者们的研究主要集中在以下方面:①有关成果评价工具的研究。蒋悟生^[2]指出《科学引文索引》、《中国科学引文索引》和《中文社会科学引文索引》是目前对自然科学基础研究成果和社会科学研究成果较为完善、科学、客观、公正的评价工具。任胜利^[3]等认为在运用SCI评价中国的科研成果时,一定要结

合我国的客观实际具体分析SCI所统计的引证数据。宋小燕、汪克强^[4]就我国现行的科技成果评价体系在知识产权保护方面存在的缺陷和不足进行了分析。②有关成果评价方法的研究。蔡长林^[5]提出了非线性加权模型方法。林强、周勇^[6]提出了一种利用最大隶属原则与加权平均原则相结合来对科技成果进行综合定量评价的方法。王璞、赵谦^[7]根据简单多数原则引入专家动态权数,构建E-BP科技奖励综合评价智能模型。③建立科技成果评价指标体系的研究。杨远芬、郭元锋^[8]分别从科研奖励评审指标体系构建原则、类型、如何构建科研奖励评审指标体系等方面进行了论述。马鲁豫^[9]分析了不同类型成果的性质和特点,增加定量评价指标,构建多元化的科技成果评价指标体系。曹霞、喻登科^[10]从科技成果转化与知识管理耦合的思路出发,构建科技成果转化知识管理绩效评价的三维体系。④针对不同地市科技成果评价的研究。吴丹^[11]根据广东省科学技术奖励管理工作实践经验和实际需要,采用德尔菲法和层次分析法对科学技术奖评价指标体系进行研究。陈晓芳^[12]针对宁夏科技成果评价存在的问题提出了改进宁夏科技成果评价体系的建议。⑤针对不同领域、不同类型成果评价的研究。石中和^[13]以应用技术类科

收稿日期:2015-04-13

作者简介:陶晓丽(1984—),女,安徽芜湖人,北京决策咨询中心,硕士,研究方向:区域发展与区域规划;王海芸(1975—),女,吉林人,副研究员,博士,研究方向:创新政策、技术创新;王新(1984—),男,山东人,助理研究员,硕士,研究方向:科技政策与评估、中国现代科学史。

技成果评价为研究内容,分析了应用技术类科技成果评价目前存在的主要问题,构建了成果评价指标体系。赵蕾、方平^[14]运用层次分析法对某水产科研院所应用技术成果、基础研究成果和软科学成果三类成果转化的综合评价指标体系进行深入分析。

可以看出学者们的研究集中在评价工具、评价方法、指标体系,有的结合一些具体地市、有的结合一些具体领域。除了以上研究,也有学者针对国外科技成果评价^[15-17]、科技成果评价指标权重^[18-20]等进行了研究,这些研究成果和研究方法具有一定的借鉴意义。

3 基础研究类指标体系参考元素和主要评价依据设计

3.1 研究方法

本研究采取了三种主要方法:一是文献研究法、二是专家访谈法、三是实证研究法。

文献研究法是对文献进行查阅、分析、整理,并通过对文献的研究形成对事实的科学认识的方法。通过文献研究找出课题研究所需要的有用信息,同时扩展视野,启发思路。专家访谈法是通过有目的地与调查对象直接交谈来获取有用信息的方法。为了充分利用科技成果评审专家库内大量的专家资源,本研究采取个人访谈和集中研讨的方式听取了大量领域专家的意见。由于被访谈对象直接影响着最终获得的资料数据的可靠性、全面性及有效性,因此在选择访谈对象时综合考虑了多方面的因素:一是被访谈对象要具有高级以上职称,具有5年以上的相关领域研究经历,同时具有项目评审、成果评价等方面的相关经验。二是考虑专家熟悉学科和从事的研究类型,均衡领域内的不同学科,本研究分别访谈了生物医药、农林、材料领域基础研究类的专家。最后利用实际申报项目进行验证性测试是通过实际申报的推荐材料进行预审,根据定量评价参考指标体系进行打分投票来验证指标体系参考元素和评价依据的科学性、合理性。

3.2 基础研究类指标参考元素设计

基础研究类成果是指对自然界有新的发现,对自然规律有新的认识、解释,在国内外首次阐明;在科学理论、学说上有创见,在研究方法、手段上有创新,对推动学科发展有重大意义,或者对经济建设和社会发展有重要影响,其核心论文应在国内外公开发行的学术期刊上发表1年以上,并得到国内外学术界的认同。

北京市科技成果评价基础研究类的评价指标从

科学价值与意义、科学发现程度、主要学术思想和观点被学术界认可的情况、主要论文发表刊物或专业著作的影响、对经济社会发展的作用五个一级指标对成果进行评价。一级指标多为定性描述,但实际评审时要求定量打分,为了能为专家在打分时提供一定的参考意见,在专家评审思路给予引导,本研究对五个一级指标分别进行了解读和分析,提出了针对一级指标的参考元素和评价依据。

3.2.1 科学价值与意义

“科学价值与意义”是指项目所述的科学发现是否为国内外首次提出或其科学理论是否在国内外首次阐明,以及所具有的价值与意义。对这一指标进行解读,可以归结为两点,即:研究内容是否具有前沿性、研究成果是否具有的一定的价值和意义,同时有学者指出科学价值需要考虑“项目对学科的发展所起到的作用和影响”^[6]。最后将参考元素确定为:前沿性、研究意义、对学科发展的作用和影响三个方面。

在具体的评价依据指标选取上,同样采取了专家访谈法和文献研究法。①针对前沿性,在访谈农林领域基础研究类专家时,有专家认为项目的前沿性要看项目是否涉及本领域的重大理论或者现实问题,以及是不是引领了学科发展方向等。②关于研究意义,在访谈生物医药领域基础研究类专家时,有专家提出可以考察项目是否国家或北京市生物医药产业亟需解决的重大问题,也就是说是否具有较大的现实需求度及实际应用价值。文献方面,有学者提出“科技成果的评价要着重预期结果有可推广性或应用前景”^[9],据此本研究提出“理论具有一定的应用范围,可在相关领域中推广应用”。同时,翻阅历年申报材料,研究意义方面大部分项目都会阐述解决领域的共性问题,解决实际问题或者是推进了更深入的相关研究等本研究总结出了其他的评价依据。③对学科发展的作用和影响,有学者提出“对学科的作用包括对本学科发展的作用和对其他学科发展的作用两个方面”^[6],包括新思想、新理论以及起到的指导和推动作用;对其他学科发展的作用主要针对交叉学科和相关学科,据此本研究也总结出了相应的评价依据。具体见表1。

3.2.2 科学发现程度

科学发现程度指项目对自然现象和客观规律发现、认识和阐明的程度,包括科学探索与发现的深度、广度、系统性和对研究领域的开拓性,是否创建新的科学理论和学说或所用研究方法与手段是否具有创新性。

表 1 科学价值与意义的参考元素和主要评价依据

一级指标	参考元素	主要评价依据
科学价值与意义	前沿性 研究内容是否涉及本领域的前沿课题	1. 项目针对性强,涉及本领域重大理论或现实问题
		2. 项目前瞻性高,准确提出学科发展方向
	研究意义 研究是否基于重大的科学问题,在解决本领域发展中存在的问题以及推动行业进一步发展的意义	1. 解决本领域共性基础问题或关键科学问题
		2. 有助于对新现象的理解及实际问题的解决
		3. 有较大的现实需求度及实际应用价值[专家访谈]
		4. 有助于推进更深入的相关研究
		5. 理论具有一定的应用范围,可在相关领域中推广应用 ^[9]
	对学科发展的作用和影响 ^[6] 成果对本学科或其它相关学科重要基础科学问题的解决情况,以及对学科发展的影响与作用	1. 填补学科空白或形成新的学科分支
		2. 对交叉学科及其它相关学科发展作出重大贡献
3. 解决本学科重要基础问题		
4. 提出新的学术思想或理论,对学科发展具有导向作用		
		5. 对相关的应用研究起到理论指导和推动作用

基础研究贵在创新,因而对基础研究项目的科学发现程度评价就显得尤为重要。具体到科学发现程度的参考元素,重点要看该项目的创新程度^[21],据此首先提出参考元素“研究内容的创新性”,具体与已有研究相比,是否具有明显的创新性,“新”可以体现在新的研究领域、新角度、新方法、新思路^[22]等。访谈材料领域专家提出创造新的科学工具和方法,据此提出“形成新的研究思路或方法”。根据文献中提到要考虑科

技成果研究协作的跨度提出参考元素“研究的深度”^[6],主要依据项目完成人对所研究的项目涉及的研究内容及研究结果的深入程度。有的专家表示还需要考虑研究的系统性,主要参考项目完成人是否长期在某一个领域里围绕重点问题进行连续研究,形成了较为系统的整体性研究体系,所发表的论文、著作等或是研究领域是否与项目紧密相关。根据专家的建议,本研究对评价依据也作了相应总结,详见表 2。

表 2 科学发现程度的参考元素和主要评价依据

一级指标	参考元素	主要评价依据
科学发现程度	研究内容的创新性 ^[21] 研究是否具有原创性、开拓性和引领性 ^[9] ,与已有研究相比是否具有超越性 ^[6]	1. 发现新现象、新物质
		2. 揭示新规律、新机理,建立新理论、新学说等
		3. 用新角度或新技术分析问题,形成新研究思路或方法 ^[22]
		4. 对已有假说或观点进行验证,补充或修正
	研究的系统完整性 是否在某一个领域里围绕重点问题形成系统的整体性研究体系	1. 发现点之间紧密相关,属于同一个体系
		2. 主要完成人(核心研究人员)从事项目研究领域的相关工作
		3. 整体研究内容完整
		4. 完整阐明科学问题,得出完整结论
		5. 提供证据充分、全面
研究的深度 ^[6] 研究内容及结果的深入程度	1. 研究内容由面及点的程度	
	2. 研究结果由浅入深的程度	

3.2.3 学术结论和观点被国内外学术界认可和引用情况

学术结论和观点被国内外学术界认可和引用情况指他人在公开发表的科学论文、著作、教材中正面引用或评价该项目提出的学术思想、观点、方法、结论,包括引用文章的质量、数量,引用文章发表刊物的权威性,或被有关实验、实践所证实的情况。有专家提出学术水平不能只看发表文章,别人对文章的评价可能更重要,文章是否解决了科学问题也很重要。

好文章、好刊物有影响力因子的界定,所以主要看别人的评价。因此,本研究在参考元素中提出“他人评价情况”、“引文的影响力”和“他引情况”。他人评价情况主要是从评论文章、正面评论、亮点论文推荐及对成果的认可情况;引文的影响力主要是看引文发表的刊物如何、引文作者的学术认可度以及对引文本身的认可度;他引情况,主要看项目完成的所有论文的 SCI 他引总次数和 SCI 单篇论文他引次数。也有专家提出,项目完成人参加国际会议的情况可以从侧面

反映项目的学术的国际化水平,因为一般只有具备较高影响力的领域专家才会受邀作高级别国际会议的报告。综上,归纳如下表 3。

表 3 主要学术思想和观点被学术界认可的情况的参考元素和主要评价依据

一级指标	参考元素	主要评价依据
主要学术思想和观点被学术界认可的情况	他人评价情况	1. 知名专家对成果专门发表的评论文章
		2. 发文期刊或国际同行专家对成果的正面评论
		3. 获本领域亮点论文推荐
		4. 代表性引文中同行对成果的认可评价情况
		5. 成果被国内外实验室、研究机构或企业验证或采用
	引文的影响力	1. 本领域同行对引文的认可度
		2. 引文作者的同行认可度
		3. 引文发表刊物的影响力
	他引情况	1. 代表性论文的 SCI 他引总次数
		2. 代表性论文的 SCI 单篇他引次数
	国际会议情况	1. 候选人在国际学术交流与项目内容紧密相关
		2. 会议本身的学术层次、知名度和水平

3.2.4 主要论文发表刊物或著作的影响

主要论文发表刊物或著作的影响指公开发表论文的学术刊物、正式出版的学术著作在国内外学术界的影响和地位。同时也要参考项目申请人编写学术专著的情况^[23],也有文献阐述了专利是创新的佐证。可以看出科研成果的主要表现形式为著作、论文、专利等。论文和著作这样的成果发表后的“质”与“量”成为衡量科研能力的一项重要指标。“量”即是指科研成果的产出力,所说的“质”即是科

研成果的影响力和学术水平,同时考虑所发表论文、著作、专利与所申报项目内容的相关性,论文要考虑影响因子。

对科研成果定量评价不是简单地将科研成果累计作加法运算,它应该根据成果的主要形式确定定量的标准与方法,专利要考虑国际或国内专利、是否核心专利、保护范围,同时考虑是否形成了行业标准、规范,证书、软件著作权也是知识产权的表现形式,要综合评价。具体见表 4。

表 4 主要论文发表刊物或专业著作的影响的参考元素和主要评价依据

一级指标	参考元素	主要评价依据
主要论文发表刊物或专业著作的影响	代表性论文、著作 ^[23]	1. 代表性论文、著作 ^[23] 应与项目内容高度相关
		2. 发表刊物的影响力 ^[9] 和权威性
		3. 著作的学术水平和创新程度
	知识产权情况 ^[24]	1. 获得国际或国内专利授权情况
		2. 形成的标准、获得的新品种证书、软件著作权等其他知识产权情况

3.2.5 对经济社会发展的作用

对经济社会发展的作用指项目成果潜在的社会效益、生态效益、经济效益的大小。评判一个项目考察其对北京产业发展的贡献度是其中之一^[24],据此本研究提出评价依据之一“为未来产业发展奠定理论基础”。同时访谈材料领域专家,指出还要考虑成果未来转化应用的可行性,有的项目可能理论很好但未来转化的可能性很小,这就涉及到潜在的经济、社会、和生态效益的大小。根据以上认识与研究,总结如表 5。

表 5 对经济社会发展的作用的参考元素和主要评价依据

一级指标	参考元素	主要评价依据
对经济社会发展的作用	对经济社会发展的作用和影响	1. 对经济社会的发展起指导或推动作用
		2. 为未来产业发展奠定理论基础 ^[24]
		3. 成果未来转化应用的可行性
		4. 潜在的经济效益、生态效益和社会效益大小

4 测评分析

4.1 测评方法

为更好测试结果,在原有的一级指标评价系统基础上,嵌入参考元素和主要评价依据。通过组织基础研究领域 13 名专家集中参会进行在线打分,测试已形成指标参考元素和评价依据的科学性、实用性和可操作性。评审测试环境采取与实际评审相同的评审环境,专家在组织评审前一周对实际申报的推荐材料进行预审,根据定量评价参考指标体系进行打分,会上确认打分结果并提交。

4.2 测评分析

4.2.1 测试对象

选取基础研究领域 2014 年实际申报的其中 20 个项目的推荐材料(表 6),为避免测试结果与初审结果部分项目不一致所带来的争议,此处隐去具体项目

的名称,以项目 A-T 来代替。

4.2.2 测试结论

根据在线打分投票的实际情况可以看出前几位项目的排序上基本是一致的,尤其是排名第一、第二的项目,评审组的测试评审结果与实际评审结果完全一致,说明质量好的项目都能准确评价,只是排在后面的项目一致性稍差。一方面的原因是受推荐指标数限制,测试评审的项目数量少,推荐指标数相应地减少,而且由于实际评审时采用的是票决制,只要推选出拟推荐项目即可,所以专家在评审时往往比较重视对前几位有望获奖项目的评价,而对不推荐项目的打分准确度稍差;另一方面是政策宣讲的影响,两次评审前对政策宣讲的程度不同会让专家对评价标准的把握程度不同。但总的来看,测试评审结果与预期结果比较一致(表 6)。

表 6 测试评审结果与实际初审结果对比

序号	专业组	项目名称	初审结果排序	测试结果排序	序号	专业组	项目名称	初审结果排序	测试结果排序
1	基础研究 专业组	项目 A	1	1	11	基础研究 专业组	项目 K	11	17
2		项目 B	2	2	12		项目 L	12	10
3		项目 C	3	4	13		项目 M	13	16
4		项目 D	4	3	14		项目 N	14	13
5		项目 E	5	7	15		项目 O	15	15
6		项目 F	6	6	16		项目 P	16	19
7		项目 G	7	5	17		项目 Q	17	12
8		项目 H	8	11	18		项目 R	18	18
9		项目 I	9	8	19		项目 S	19	14
10		项目 J	10	9	20		项目 T	20	20

5 研究结论

本研究在北京市科技成果评价基础研究类评价原有五个一级指标的基础上,进一步细化了针对五个一级指标的不同的参考元素和主要评价依据,并且将在未来基础研究领域评审工作中发挥作用。一方面,可以更加科学、全面地对科技成果进行客观公正的评价;另一方面,可以引导专家更加合理、客观、公正的评审,提高评审工作的科学性和权威性。

当然,鉴于北京市科技成果奖的评审范围涉及领域和项目类别众多,本次研究仅涉及其中的一部分,不可能在实际评审中对所有项目都直接套用该量化指标体系,需要根据具体的项目类型再分类细化。同时,本研究不涉及指标权重合理性及一级指标的全面性、独立性和易于取得性等问题的讨论,在今后的研究中有待进一步深入,尤其是对其他领域和其他项目类别的评价指标进行深入研究,补充相对应的参考元

素和主要评价依据,为更好地开展科技成果评价提供决策依据和参考。

参考文献

- [1] 张立军,杨娟.考虑专家权重的科技成果模糊综合评价[J].科技管理,2011,24(4):1-5.
- [2] 蒋悟生.研究成果评价的重要工具:科学引文索引[J].情报杂志,2001,20(5):82-84.
- [3] 任胜利,等.SCI的引文统计指标及其与研究评价的关系[J].编辑学报,2003,15(1):70-72.
- [4] 宋小燕,汪克强.科技成果评价体系中的知识产权保护问题[J].科技管理,2003(1):24-26.
- [5] 蔡长林,王绮萍,等.科技成果奖励评价指标体系中非线性数学模型及其应用[J].软科学,1991(3):17-22.
- [6] 林强,周勇.科技成果定量评价方法研究[J].山东矿业学院学报,1995,3(1):35-41.
- [7] 王瑛,赵谦,等.基于E-BP神经网络的科技奖励评价模型研究[J].科技进步与对策,2011,5(10):111-114.
- [8] 杨远芬,郭元锋.论科研奖励评审指标体系的构建[J].决策

- 咨询,2008(11):4.
- [9] 马鲁豫. 多元化科技成果评价指标体系的建立[J]. 科技成果纵横,2009(5):31-34.
- [10] 曹霞,喻登科. 科技成果转化知识管理绩效评价体系的构建[J]. 科技进步与对策,2010,9(17):128-131.
- [11] 吴丹,王雅文,胡羚. 广东省科学技术成果奖励评价指标体系研究[J]. 中国科技论坛,2010(6):9-12.
- [12] 陈晓芳. 科技成果评价体系分析[J]. 中国高校科技,2012(1):31-32.
- [13] 石中和. 应用技术类科技成果评价及指标体系研究[J]. 北京交通大学学报:社会科学版,2007,9(3):54-58.
- [14] 赵蕾,等. 基于层次分析法的渔业科技成果转化综合评价体系研究——以某水产科研院所为例[J]. 科技管理研究,2012(14):55-60.
- [15] 陈准,郑永和. 美国国家科学基金会加强内部管理的四项要点[J]. 中国基础科学,2001(4):44-47.
- [16] 刘莹,等. 美国联邦科研机构的绩效评估制度及其启示[J]. 中国科技论坛,2007(9):140-144.
- [17] 陈宁. 美国的科技评价与科研事后评价概况[J]. 全球科技经济瞭望,2007,22(12):25-31.
- [18] 孟渠成. 开发性应用技术成果评价指标的探讨[J]. 中华医学科研管理杂志,1997,10(4):223-224.
- [19] 张立军,等. 基于路径分析的科技成果评价指标赋权方法研究[J]. 统计与信息论坛,2008,23(4):67-70.
- [20] 张立军,等. 考虑专家权重的科技成果模糊综合评价[J]. 科技与经济,2011,24(4):1-5.
- [21] 曹玮. 科技奖励立体式综合评价方法研究[D]. 长沙:湖南大学,2009.
- [22] 王永红. 科研成果定量评价的难点与方法创新[J]. 中国高校科技,2012(1):39-41.
- [23] 曹阳,等. 基于多层次灰色评价模型的生物医药产业集群创新能力实证研究[J]. 企业经济,2012(5):94-97.
- [24] 葛朝阳,郑刚,陈劲. 基础研究的经济回报率测度与评价:国外研究述评[J]. 科研管理,2003(1):44-50.

Elements Design and Test of Evaluation Index System of Scientific and Technological in Beijing

——In basic research as an example

TAO Xiao-li, WANG Hai-yun, WANG Xin

(Beijing Decision-making Consultant Center, Beijing 100089, China)

Abstract: Through the deep systematical analysis of thenotification materials over recent years of Beijing, Using the method of literature study and expert interview, this paper proposed a more advanced indicator system withreference elements and primary evaluation basis to improvethethe first-level indicators of original evaluation system. The evaluating indicator system was applied to assess the actual notification materials of Beijing in 2014 by experts and showed good consistency after comparing the rankings of experts' test results and preliminary examination of all projects. The research indicated that the evaluating indicator system with reference elements and evaluation basisproposed in this article worked well in basic areas ,it can effectively guide experts to make more fair and just evaluation.

Key words: Beijing; the evaluation index system; elements; test