

“创新力-影响力-传承力”科研评价维度构建与实证*

■ 宋歌

东南大学图书馆情报科学技术研究所 南京 210096 东南大学经济管理学院 南京 210096

摘要: [目的/意义]为更加客观地进行科研评价,激发我国科研创新,遏制学术不端,本文从遵循科学发展规律的角度进行分析,认为科研评价维度应多元化发展,并尝试提炼出创新力、影响力、传承力3个维度。[方法/过程]首先,论述3个维度的产生机制、重要性、评价困境与进展、相互关系。其次,选取能够表征3个维度的指标进行实证研究,分析不同维度指数的相关性,不同维度指数关联评价的结果类型及发生概率,维度与扩散阶段融合评价的作用。[结果/结论]得出的结论包括:科研评价指标体系中应确立不同的评价维度,且每个维度下都具有丰富的指标;应明确每种科研指标所表征的维度;评价时宜综合采用多种指标;重视区分创新扩散阶段的融合评价。并对“创新力-影响力-传承力”的评价维度以及更广泛意义上的多维评价的应用前景进行展望,包括增强科研评价的灵活性、针对性、客观性和解释力;对于潜在人才和创新的发现;为评价指标从基于数量到基于网络结构的跃升提供广阔空间;避免“评价陷阱”;鼓励深层次研究和创新发现,匡正学术风气;引导学术出版和写作规范;增强同行评议的洞察力。

关键词: 科研评价 评价维度 创新力 影响力 传承力

分类号: G250

DOI: 10.13266/j.issn.0252-3116.2017.10.001

1 引言

在当前的科研评价指标体系中,无论是科研成果、科研人员、科研机构,还是学术期刊的评价指标,均是围绕着测度评价客体的“影响力”展开的,而同行评议也存在同样的问题。“它可能对影响力而不是‘真实的’科学质量更敏感”^{[1][126]},受到学者声望、以往成就、刊物级别、出版平台、获奖等一系列外在标识的影响。长期采用单一评价维度,对学术风气的负向驱动已经十分明显。例如,为提高指标值,个人、研究团队、期刊的过度自引和不当他引已经成为危害学术规范的暗流。又如,为片面追求影响力,各学科领域不断掀起概念炒作和跟风研究的热潮。再如,近年来论文造假事件和手段不断升级,论文造假产业也不断壮大。

实际上,科研成果在科学发展过程中所起到的作用是多元的,而科研评价体系应囊括其中的主要维度。如此,不仅能较为真实地反映评价客体对科学的推进,而且不同维度相互制约,形成立体、客观的评价结果,有利于将各学术主体的行为导向规范。

为此,本文尝试构建科研评价中的主要维度,并进行初步实证。

2 科研评价维度构建

2.1 创新力

科学发展的原动力是科研创新力;科学发展路径由高创新力成果所标记;科学发展过程表现为科研创新与扩散以及再创新与再扩散的循环往复和交叉重叠。因此,从维护、促进科学发展的角度来说,科研评价中最重要的评价维度就是创新力维度。然而,由于同行评议自身的局限以及影响力指标的易得,创新力评价维度始终模糊不清,难以确立。半个多世纪以来,越来越多的研究、调查揭示^[2-5]:科学突破往往需要经过很长时间才能被广泛接受。最好的例证莫如诺贝尔奖(以下简称“诺奖”)奖励的常常是几十年前的成果,而诺奖级成果常遭到拒稿和发表后的冷遇。据研究^[6-7],Web of Science数据库中,睡美人文献存在概率不足1%,而在诺奖获奖者的论文中,睡美人文献存在概率约为20%。虽然这些成果后来获得了认可,但很

* 本文系国家社会科学基金项目“学术创新扩散过程及创新力测度研究”(项目编号:15CTQ027)研究成果之一。

作者简介:宋歌(ORCID:0000-0001-8064-4583) 副研究馆员,博士,E-mail:songge.2000@163.com。

收稿日期:2017-02-17 修回日期:2017-05-03 本文起止页码:5-13 本文责任编辑:刘远颖

难估计有多少潜在的重要发现因其价值没有得到及时认可而不得不中断。

原始创新、超前创新或曰突破性成果遭遇科学共同体的排斥或抵制,是由于其跃出了现有的学术范式。每个科学领域都会对所研究的复杂内容做一些简单假设,并将其纳入指导该领域进行研究的学术范式中。这些范式为学者提供连贯的研究方向,降低研究过程中出现的大量不确定性,但是同时也形成了一系列假设、概念和偏见,并使之标准化。“科学家们通常意识不到这些假设的存在,即使它们已经影响到一些重要的因素,比如研究了什么,忽略了什么,以及哪些研究方法应该受到重视,哪些应该遭到摒弃。”^{[8]86-90}因此,当一个科学家遵循一个理论范式时,他的理性盲点会妨碍他看到真实情况。并且,这些东西一旦存在,就很难被认识到并加以克服。而创新者一直在领域的薄弱环节处工作,在“无疑处有疑”。正如诺贝尔物理学奖得主丁肇中^[9]多次指出的“科学是多数服从少数,只有少数人把多数人的观念推翻以后,科学才能向前发展。因此,专家评审并不是绝对有用的,因为专家评审是依靠现有的知识,而科学的进展是推翻现有的知识”。诺贝尔生理学或医学奖得主 S. Prusiner^{[10]62}也认为“虽然科学家对不符合公认的科学知识领域内的新理念持怀疑态度是很合理的,但是最好的科学发现通常来自并不符合公认范式的实验结果”。

近年,在测度评价客体创新力方面的研究有所进展。例如,“成功论文”(successful paper)^[11]的界定方式,以“关键词交叉率”^[12]测度学科创新潜力,基于网络结构变化的“变革性潜能指标”^[13]根据结构洞理论衍生的“创新潜力指标”^[14]以及判断成果创新力的 S 指数^[15]等。相对于已经较为成熟的影响力指标体系,创新力相关指标的研究还很少,也没有指标进入常规的科研评价体系。

2.2 影响力

当前,常用的科研评价指标均属于影响力维度。例如,评价期刊的影响因子 IF、评价成果的被引量指标、用来考察科研人员高被引论文数的 h 指数以及这些指标的改进形式,包括特征因子、引用中位数、Nature Index、Google Scholar Metric 的 h5、众多的 h 类型指数等。在同行评议中,提供给评议专家用于辅助判断的成果发表平台、论文转载和获奖情况等也反映的是影响力维度。评价客体的影响力对于科学发展有巨大推动作用。然而,如果仅仅关注影响力维度,科研评价就无法打开新局面。

影响力的形成是“信息觅食”^{[16]14-16}的结果。在科学发现中找出具有创新性的想法就像是觅食。科研人员作为知识的觅食者和创造者,要找的是一个新想法、新理论或者科学争论中的一个新证据,因此,就涉及对前沿领域、相关知识的识别。同时,还要考虑时间、精力及其他资源的投入成本,而且要承担预期收益的不确定性。觅食过程可以表述为:科研人员对多个主题、领域的成果进行评估,判断哪一项研究更值得关注,并决定探索该领域所应投入的时间。而最优觅食,即最大限度地提高所获收益与所付出的成本之比。从收益考虑,那些具有现实意义的创新成果无疑大大增加了预期收益。因此,如果一项创新有利于解决当前社会面临的紧要问题,则更容易被发现、被发展,从而被快速地传播、扩散,衍生一系列相关的创新性成果,逐渐或迅速地形成研究热点,产生热点文献和舆论领袖,形成影响力。从成本考虑,主要方面包括创新与现存观念的相容性、技术可行性、易获得性、易理解的程度、经济因素等。因此,一些目前不具备充足的社会、经济、技术条件来实现的创新可能被觅食者放弃,而一些颠覆当前学术范式、难以被科学共同体接受的创新更加不易扩散,无法在短时期内形成影响力。科研人员普遍采取最优觅食决策就会有选择地触发创新。一些创新成果,会被大量前来觅食的科学家选定,汇集成新的密集知识流,从而在宏观上涌现科学知识结构的变迁,即知识领域的扩张或新的知识领域的诞生以及核心研究领域的迁移。

影响力与创新力之间的关系颇为微妙。通过对影响力形成过程的分析,可见二者的联系体现在:创新力虽然从成果完成之时就完整地蕴含在成果之中,但是创新力的彰显是随着成果被理解,继而传播、被利用而逐渐展开的。因此,一方面,影响力不能代表创新力,但是可以用影响力指数的增长情况表达成果创新力被认可的情况;另一方面,成果创新力不能决定其影响力指数最终的大小,因为影响力不完全是由成果的创新力引起的,还受到其他诸多因素的影响,包括作者声誉、发表平台、可见度、语种、表达技巧、学科属性、创新扩散阶段等。基于以上原因,甚至一些对社会发展有着重大意义的创新成果仍然有可能被埋没,造成“迟滞承认”^[17]。例如青蒿素的发现,从其研究成果的首发到被国际学术界认可就历经了几十年的时间。

影响力维度的评价指标较为丰富,但是对于评价结果的解释需要进一步澄清。在科研评价相关研究中,存在将影响力与创新力混为一谈,用影响力指标

(如影响因子、被引量等)来说明创新力的情况。实际上,有影响力的成果,其创新力不一定高;同样,创新力高的成果,在一定时期内,其影响力也不一定高。《英国医学杂志》主编 F. Godlee 认为,一篇质量非常高的生物技术研究论文可能因为不符合杂志发表范围被直接拒稿,而且论文引用多少不能作为论文质量的标准。《自然》杂志也发现被引用次数最高的科学论文基本是一些方法类论文,而不是真正的科学突破^[18]。

2.3 传承力

没有传承,就没有创新。虽然具有创新力的成果必然具有新颖性,或曰独创性,甚至是史无前例的变革性研究,但是“创新并非真的是无中生有,而是将一些风马牛不相及的知识领域联系在一起形成的独创性研究”^[15]。科学作为一个整体,其中每个学科领域的增长都十分重要,正是因为“发现了在旧领域中的纽带和联系”导致了“新的领域不断出现”^[19]。具有传承力的成果形成各个研究领域的主要脉络,而在不同脉络节点上的链接促生了科研创新。因此,科研成果乃至学术流派的传承力是科学发展和科研创新的基石。而出版机构、科研机构甚至国家地区对于学术传承所起到的作用也值得深究。

对于成果传承力的讨论由来已久。克兰提出“倒金字塔”^[20]现象,认为主要的创新在一个领域的全部历史中不断出现,然而,这些成果是通过对这一领域早期论文的引证而联系起来的。大量论文都直接或间接地与处于塔顶的几篇论文发生联系。此外,在每一组新发表的论文的参考文献中,大约有一半使这些论文和早些时候的一小组出版物联系起来,这些出版物中的大多数和引证它们的论文在时间上是很接近的;另一半参考文献,显然是任意地把这些论文与科学文献中范围很广的一部分联系起来。加菲尔德的实证^[21]未将样本限定在单一领域,而是对有关遗传密码的多个研究领域的引证作分析,同样也验证了把论文联系起来的传承路径。

时至今日,科学家们越来越倾向于引用新近的研究成果,针对老成果是否在加速淘汰的话题,谷歌学术搜索研究团队^[22]在分析了其自有的大量论文索引后发现:引文比原文至少晚 10 年的论文比例呈稳定上升趋势,从 1990 年的 28% 上升到 2013 年的 36%。这一趋势在九大研究领域占到 7 个,并占 261 项主题中的 231 项。因此,尽管每年学术期刊以及发表的科研论文数量均呈现出巨大增长,但科学家的集体记忆仍在向历史深处蔓延。而其余淘汰率稳步增加的 30 项

主题通常是那些在过去 20 年中新近成为主流研究领域的学科,如纳米技术和艾滋病病毒。这些研究领域还没有足够的历史积淀可以被引用。

2.4 科研评价维度构建模型

科研评价维度的构建应基于对科学发展过程本质的认识,并对有利于科学发展的主要方面进行提炼,使之成为科研评价体系中相互促进,且互为制约的多元评价维度。科研发展过程是由科研创新和创新扩散相继发生,共同推动的。根据上文对科研创新力、影响力、传承力的解读可知:科学的演进依靠高创新力成果的产生,借助高影响力成果加速,通过高传承力成果形成脉络,由脉络间的链接激发新一轮创新。因此,本文基于科学发展过程,提出“创新力-影响力-传承力”科研评价维度。维度构建的基础和 3 个维度之间的关系如图 1 所示:

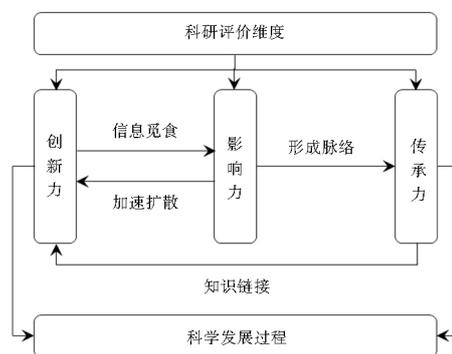


图1 科研评价维度构建模型

其中,创新力维度的确立,将提高创新成果的预期收益,降低信息觅食成本,有利于科研人员在信息觅食中触发更多、更前沿的科研创新,从而加速创新成果的产生及扩散,缩短科研创新周期,激发科研创新领域,助力我国科研创新进入快车道。传承力维度的确立,不但对那些在科学发展史上做出贡献的科研主体的明确认可,而且可以标记各个学科领域的增长与变革、萎缩与转移、融合与分离,为创新主体在不同层次和领域的知识版图中跨越边界、激发创新提供标志物。创新力和传承力维度的确立将使得影响力维度的内涵更加清晰,指向更加明确,便于澄清影响力指标在科研质量和学术水平评价中遭受的质疑。

3 实证

科研成果评价是一切科研评价的基础。因此,以成果评价为例,分别选取能够表征科研成果创新力、影响力、传承力 3 个维度的常用或适当的指标进行实证。

3.1 数据来源

结构洞理论是一项重要的科研创新,发端于社会学(sociology),在商业与经济学(business & economics)领域得到深入研究与应用,并对社会科学领域有着广泛渗透。本文借助该项创新 20 多年中发展累积的成果进行实证。除了 R. Burt 提出结构洞概念的著作 *Structural Holes: The Social Structure of Competition* 以外,在 Web of Science 平台的 SSCI 数据库以“structur* hole*”为检索词进行主题检索,限定文献类型为 ARTICLE 或 REVIEW,时间跨度为 1992-2015 年,结果命中文献 581 篇。这 582 篇论著构成的引文网络共有 3 309 条弧、10 个孤立点。单独考察孤立点文献,发现结构洞理论在其中均为次要主题。因此,样本确定为构成连通网络的 572 个成果。

一项成功的扩散,其完整过程包括 4 个阶段:起步、起飞、成熟和衰退^[23]。结构洞理论的扩散曲线见图 2。其中 X 轴表示创新扩散年代, Y 轴表示创新的累积成果数。该 S 形曲线在 2007 年的扩散加速度(diffusion acceleration, da)增量最大,而于 2014 年达到扩散速度(diffusion rate, dr)的最大值 74,形成了 S 形曲线的一阶拐点。因此,可以认为 1992-2006 年为结构洞理论扩散的起步阶段,2007-2014 年为起飞阶段。如果 2015 年以后的扩散速度不再高于之前年份,则结构洞理论扩散的成熟阶段可自 2015 年始。

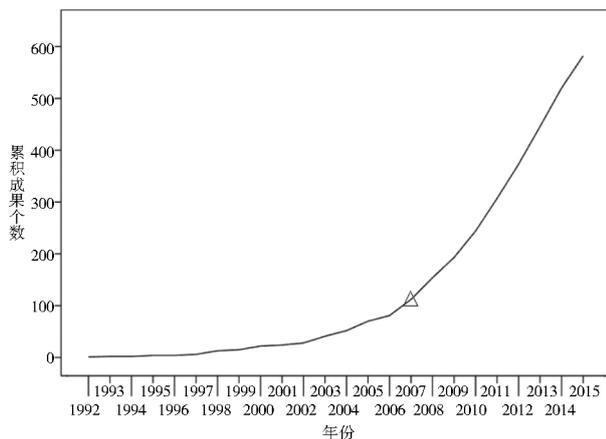


图 2 结构洞理论扩散曲线

3.2 指数选择及计算

影响力维度采用被引量作为测度指标。被引量通常被视为测度学术出版物影响力的合理指标,至少可以用来体现同行科学家对其关注的程度。创新力维度采用专门测度成果创新力的 S 指数。传承力则利用“遍历权值”(traversal weight)得出。

其中, S 指数基于引文时序网络的结构属性,从相

对已有研究的变异程度和对科学发展的影响两方面,测度一项科研成果在多大程度上推动了科学发展。其计算公式如下^[15]:

$$S_y = \frac{D_{in}}{D_{in} + D_{out}} \quad \text{公式(1)}$$

S_y 为某成果在 y 年的 S 指数, D_{in} 和 D_{out} 是成果节点在参照客体构成的引文网络中的入度和出度。其中, D_{in} 为与该成果主题直接相关的参考文献数,代表成果创新的变异程度, D_{out} 为该成果在该主题领域中从发表年至 y 年的总被引次数,代表新知识产生以后对科学发展的影响。S 指数的取值范围是 $[0, 1]$, S_y 为 0 时,成果创新力最高,为原始创新成果; S_y 为 1 时,创新力最低。

遍历权值^{[24]244-246} 是进行“主路径分析”(main path analysis)的指数。主路径分析把引文网络看作是一个科学知识或信息的传输系统,勘测有大量知识流流经的引证关系和论著,其目的在于提取传输网络中的主要脉络。而遍历权值计算的是:如果要把其他论著连接到一起,要在多大程度上需要某条引证关系或某篇论著。因此,成果节点的遍历权值反映的就是成果在科学发展过程中表现出来的传承力。计算过程为:首先,在引文网络中找出从每个“源点”(source vertex)指向每个“宿点”(sink vertex)的所有路径,计算出含有某项成果的路径数;接着,用包含指定成果的路径数,除以网络中源点与宿点之间路径的总数,由此得到的比值就是某篇成果的遍历权值。

3.3 不同维度指数相关性分析

572 篇成果在 2015 年的 S 指数、被引量 and 遍历权值的排序对比见表 1,散点图见图 3,相关分析结果见表 2。Pearson 相关分析显示 3 种指标在 p 为 0.01 的水平上两两显著相关。其中, S 指数与被引量的相关系数为 -0.468。相关系数不高的原因在图 3 中有所揭示,即创新力高的成果,其影响力不一定高,甚至有可能很低;反之,影响力低的成果,其创新力有可能很高。被引量和遍历权值的相关系数较高,为 0.731。由散点图可知,大量成果的影响力和传承力都处于低位;而高影响力成果,其传承力大多也高;还有部分影响力中等的成果同时拥有中等水平的传承力。以上情况导致了两种指标的高度相关。而不相关的部分在于:一些影响力或传承力处于中等水平的成果,其传承力或影响力较低。相关系数最低的是 S 指数与遍历权值,为 -0.372。散点图显示,在高创新力成果中,其传承力从高到低均有分布。

表1 创新力、影响力、传承力维度指数

排名	文献编号	S 指数	文献编号	被引量	文献编号	遍历权值
1	0	0.000 000	0	432	0	0.999 987
2	4	0.006 289	20	266	16	0.710 428
3	20	0.011 152	48	214	48	0.466 205
4	5	0.016 260	4	158	5	0.380 830
5	18	0.024 194	5	121	41	0.334 132
6	48	0.031 674	18	121	2	0.296 386
7	62	0.040 650	62	118	17	0.295 051
8	6	0.045 455	55	99	55	0.254 614
9	23	0.051 724	16	82	193	0.227 538
10	12	0.055 556	23	55	10	0.209 430
11	8	0.056 604	8	50	165	0.160 959
12	55	0.066 038	42	45	8	0.160 839
13	34	0.069 767	165	43	89	0.144 039
14	70	0.076 923	34	40	60	0.141 457
15	42	0.081 633	89	37	242	0.137 528
16	21	0.083 333	70	36	74	0.120 023
17	54	0.083 333	27	32	345	0.107 090
18	27	0.085 714	64	29	3	0.105 808
19	64	0.093 750	146	29	329	0.101 905
20	16	0.098 901	74	28	15	0.101 704

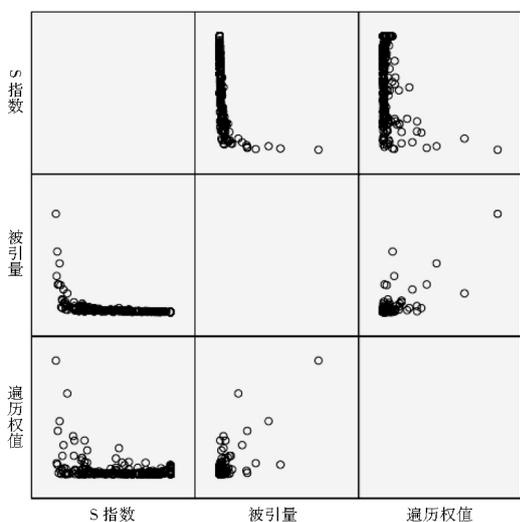


图3 创新力、影响力、传承力维度指数散点图矩阵

表2 创新力、影响力、传承力维度指数相关分析

		S 指数	被引量	遍历权值
S 指数	Pearson 相关性	1	-.468 **	-.372 **
	显著性(双侧)		.000	.000
	N	572	572	572
被引量	Pearson 相关性	-.468 **	1	.731 **
	显著性(双侧)	.000		.000
	N	572	572	572
遍历权值	Pearson 相关性	-.372 **	.731 **	1
	显著性(双侧)	.000	.000	
	N	572	572	572

注: **. 在 0.01 水平(双侧)上显著相关

3.4 不同维度指数的关联评价

若将一项评价客体 3 个维度的指数排名做相对高、低之分,则“创新力-影响力-传承力”关联评价可能出现 $2^3 = 8$ 种结果,如表 3 所示:

表3 “创新力-影响力-传承力”关联评价结果类型

维度 \ 类型	类型							
	一	二	三	四	五	六	七	八
创新力	高	高	高	高	低	低	低	低
影响力	高	高	低	低	低	低	高	高
传承力	高	低	高	低	低	高	低	高

根据发生概率,这结果可分为 5 种情况:

(1) 在一个研究领域中,大部分成果处于低水平的第五类,在任何评价维度中没有显示度。如在本例的 572 篇论著中,有 307 篇的 S 指数为 1,被引量为 0,遍历权值接近 0。

(2) 在一个研究领域中,会有极少部分的成果属于第一类,它们在各个评价维度中均处于高水平状态。如结构洞理论的开山之作文献 0,以及文献 5、8、48 等在理论、应用创新、领域影响和学术传承方面均发挥了重大作用。这类成果在现有的各评价体系中都会被凸显。而且不仅其影响力在评价指标中被彰显,其创新力与传承力也会得到同行的普遍认可。

(3) 第二、四、六、八类成果时有出现。现分别举例说明其原因。G. Walker 于 1997 年发表的文献 4 是将结构洞理论引入商业与经济学的开创性成果,创新力排名第二,仅次于文献 0,同时,影响力排名第四,而其传承力仅排第二十二名,相对较低。部分原因是由于 R. Burt 于同年也发表了将结构洞理论运用于商业与经济学领域的文献 5。虽然文献 5 在创新力和影响力方面的排名为第四和第五,都低于文献 4,但它的传承力排名第四,在很大程度上分流了文献 4 的传承作用,见图 4。因此,可以认为文献 5 属于第一类成果,而文献 4 属于第二类成果。

与此类似,文献 20 和 55 的研究主题都是企业创新,均为在商业与经济学领域研究中起到枢纽作用的关键成果。文献 20 发表于 2000 年,承接 R. Burt 和 G. Walker 对结构洞理论在社会学和商业与经济学中应用的研究,被包括文献 55 在内的商业与经济学文献引用。但是,该文献在该领域的创新扩散中仅起到了阶段性作用,其地位随后被 2005 年发表的文献 55 所取代。随着 2007 年结构洞理论创新扩散起飞阶段的到来,文献 55 被广泛引用,并孕育了近年来在该领域产生重要影响的文献 193。因此,文献 20 属于第二类成果,文献 55 属于

第一类成果,而文献 193 为第八类成果。

R. Burt 于 2000 年发表的文献 16 也属于第八类成果。它拥有第二大遍历权值 0.710 428,仅小于文献

0 的 0.999 987。该文对于结构洞理论在商业与经济学领域的迅速扩散起到重要作用,具有极高的传承力。然而,其创新力排名第二,并非开创性成果。

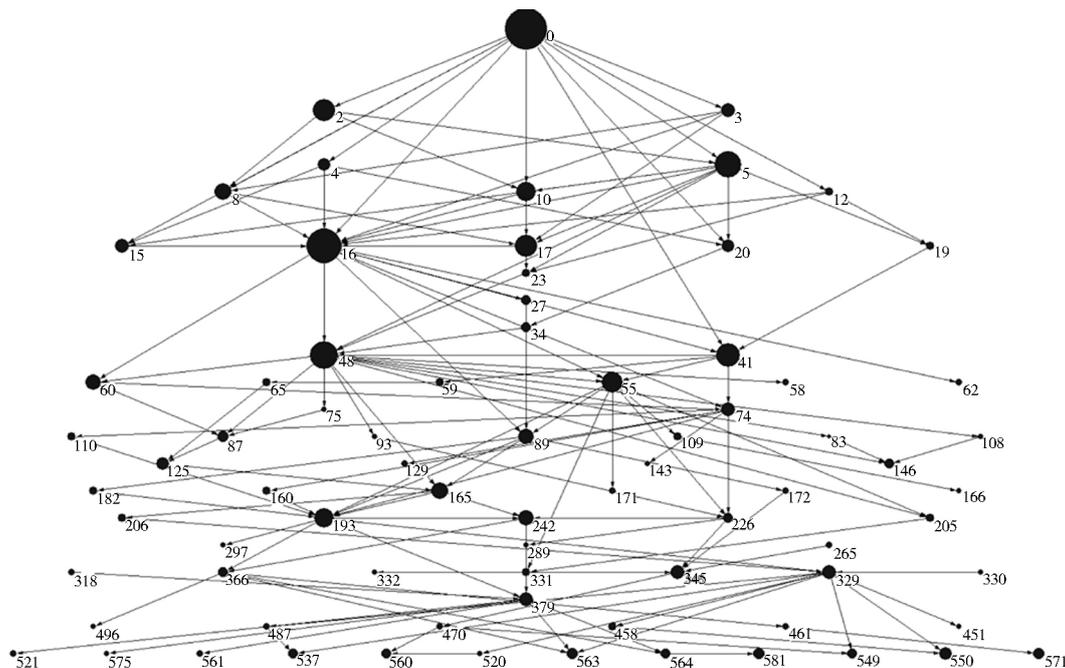


图 4 结构洞理论扩散主路径

注:节点大小根据遍历权值设定

第六类成果如 2007 年发表的文献 87。它是商业与经济学和心理学方面的跨学科研究,在该主题研究中起到承上启下的作用,但它并不像文献 27 和 50 那样是该主题领域的开创性研究。3 个维度的指数显示文献 87 是第六类成果。

第四类成果是具有潜在影响力的成果,是在当前科研评价体系中容易被忽视而具有价值的成果类型。其影响力和传承力相对于创新力而言处于低位。这类成果的极端例子就是“睡美人”文献。本例中,文献 438 的 S 指数为 0.2,创新力排名 33,属于前 5.77%,而其被引量仅为 4,传承力排名为倒数 5.42%,是典型的第四类成果。这篇发表于 2013 年的文献,研究嵌入性网络对于科研产出数量和质量的影響。虽然其并非嵌入性网络的开创性研究,但却是嵌入性网络在科研评价领域的创新应用,是潜在的科研生长点,非常值得关注。类似的还有文献 106、43 等。

(4) 第三类成果极少出现。这一情况仅发生于原始创新成果,并仅发生在创新扩散的起步阶段。原始创新往往跃出当前的学科范式,不易被科学共同体理解,在产生之初默默无闻,因而往往在发表后的较长时间内处于高创新力、低影响力的状态。同时,局部的小

范围扩散使得原始创新成果的传承力指标始终处于高位。例如文献 0,是 R. Burt 于 1992 年提出“结构洞”概念的著作。其 S 指数自始为 0,提示其为原始创新成果;遍历权值自始为 1,2009 年以后约为 0.99,一直处于高位。然而在该著作发表后的 5 年中,相关论文仅有两篇,且是 R. Burt 自己发表的。直到 1997 年才有其他研究者引用文献 0。此后历年的被引量也很小,至 2007 年才有较大突破,见图 5。因此,文献 0 在发表后的 10 多年中是典型的第三类成果,即同时具有高创新力、高传承力和很低或相对较低的影响力。

(5) 第七类成果几乎不会出现。创新力和传承力均低的成果,其影响力不可能高。

成果的关联评价结果类型可能随时间发生改变。例如,在发表 15 年后,文献 0 的创新力被越来越多的同行认可,带来影响力的显著增长,由此从第三类成果转变成为第一类成果。再如文献 438,当前属于第四类成果,具备高创新力和相对较低的影响力、传承力。随着时间推移,其影响力可能会进一步提高,转变成为第二类成果,甚至第一类成果。因此,关联评价结果类型可能会改变这一情况,这是由成果影响力和传承力的滞后性造成的。

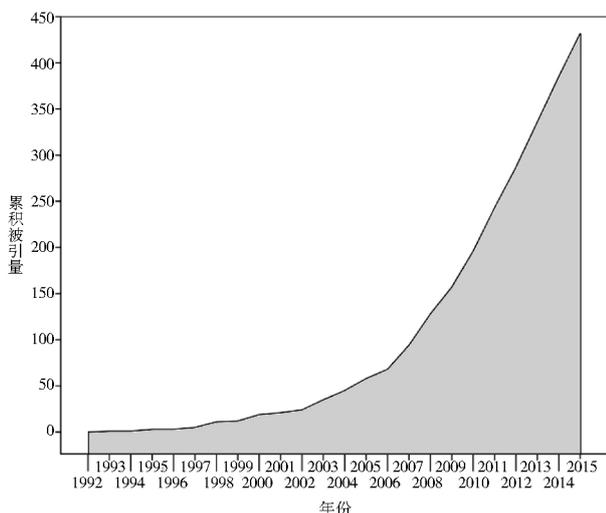


图5 文献0 影响力指数增长过程

3.5 维度与扩散阶段的融合评价

“在不同的扩散阶段会产生不同性质的创新成果。一般来说,开创性成果大部分是在起步阶段完成的;起飞阶段可能发生创新在各学科领域的大范围扩散;成熟阶段的应用创新成果较多;衰退阶段的高创新力成果往往预示新研究领域的诞生或研究范式的转变。”^[15]因此,有必要进行评价维度与扩散阶段的融合评价,以便筛选出不同扩散阶段的优秀成果。本例中,1992-2006年的78篇论著为起步阶段成果,2007-2014年的432篇论文属于起飞阶段成果。这两个阶段各自的关联评价指数排名如表4、表5所示:

表4 起步阶段关联评价指数

排名	文献编号	S 指数	文献编号	被引量	文献编号	遍历权值
1	0	0.000 000	0	432	0	0.999 987
2	4	0.006 289	20	266	16	0.710 428
3	20	0.011 152	48	214	48	0.466 205
4	5	0.016 260	4	158	5	0.380 830
5	18	0.024 194	5	121	41	0.334 132
6	48	0.031 674	18	121	2	0.296 386
7	62	0.040 650	62	118	17	0.295 051
8	6	0.045 455	55	99	55	0.254 614
9	23	0.051 724	16	82	10	0.209 430
10	12	0.055 556	23	55	8	0.160 839
11	8	0.056 604	8	50	60	0.141 457
12	55	0.066 038	42	45	74	0.120 023
13	34	0.069 767	34	40	3	0.105 808
14	70	0.076 923	70	36	15	0.101 704
15	42	0.081 633	27	32	4	0.096 650
16	21	0.083 333	64	29	20	0.083 059
17	54	0.083 333	74	28	34	0.053 538
18	27	0.085 714	17	26	27	0.048 687
19	64	0.093 750	60	24	65	0.041 396
20	16	0.098 901	22	23	23	0.041 344

表5 起飞阶段关联评价指数

排名	文献编号	S 指数	文献编号	被引量	文献编号	遍历权值
1	108	0.142 857	165	43	193	0.227 538
2	165	0.156 863	89	37	165	0.160 959
3	106	0.200 000	146	29	89	0.144 039
4	438	0.200 000	205	23	242	0.137 528
5	109	0.214 286	109	22	345	0.107 090
6	171	0.217 391	129	22	329	0.101 905
7	89	0.229 167	153	19	379	0.100 372
8	185	0.230 769	108	18	125	0.080 931
9	205	0.233 333	171	18	87	0.070 108
10	147	0.235 294	193	18	226	0.057 857
11	146	0.236 842	97	16	366	0.056 429
12	153	0.240 000	134	16	146	0.047 060
13	87	0.250 000	283	16	331	0.046 795
14	111	0.250 000	110	15	206	0.043 597
15	140	0.250 000	127	15	458	0.039 826
16	97	0.272 727	140	15	205	0.039 389
17	187	0.272 727	166	15	110	0.038 016
18	101	0.285 714	172	15	160	0.037 961
19	110	0.285 714	227	14	182	0.037 875
20	127	0.285 714	143	13	109	0.034 302

对比表1、表4、表5可见,表1和表4的区别不大,而和表5的区别很大。因此,区分扩散阶段的关联评价至少可以避免其他阶段的重要成果被起步阶段的开创性成果所淹没。举例如下:①起飞阶段创新力最高的两篇文献是108和165(S指数均小于0.2),它们都是结构洞理论扩散到商业与经济学后,在理论方面的重要发展。尤其是文献165,它在3个维度的关联评价中居于一、二名的位置,是起飞阶段最显著的第一类成果。它论述结构洞的形成,阐明网络结构的涌现来自于“结构约束”和“网络机会”这两种互补力量的相互作用。然而,如果不进行划分扩散阶段的评价,仅采用表1作为参考,则文献165的重要性并不明显。此外,在起飞阶段同属第一类成果的还有文献89、146、205、109等。②在区分扩散阶段的融合评价中,更易发现具有潜在影响力的第四类成果,如文献438和106。③高居起飞阶段传承力榜首的文献193是一篇综述性质的研究论文,属于低创新力、高影响力、高传承力的第八类成果。同样,还很容易发现近年来具有高影响力和传承力的综述论文345。

4 结论与展望

遵循科学发展规律,深入思考科研评价复杂性,科研评价维度应多元化发展。本文尝试提炼出了3个维

度,并进行了实证,得出以下主要结论:①在科研评价指标体系中应确立不同的评价维度,每个评价维度下都具有丰富的评价指标可供选择,在进行科研评价时,就能够根据不同的评价客体与目的,选择适当的指标参与评价,提高科研评价的灵活性、针对性、客观性和解释力。②应明确每种科研指标所表征的维度,包括确定是具有单一测量维度的指标,还是能够表达多个维度,及其在不同维度上的表达程度。③没有一个指标能单独提供关于科研评价的权威评估,需要多种指标结合使用。④基于科研创新扩散过程,区分创新扩散阶段进行科研评价应受到重视,多维评价与扩散阶段的融合有助于筛选不同类型的优秀成果,把握科研周期,预测科研发展前景。

对于“创新力-影响力-传承力”的评价维度,以及更广泛意义上的多维评价,其应用前景如下:

(1) 通过多维评价形成的综合评价,能够帮助人们得到关于评价客体的较为客观、完整的图景,可对其在科学发展中起到的独特作用进行多方阐释。避免埋没学术新人和具有发展潜力的理论、技术、方法、观点,维护学术生态多样性,激发更多的“创新场”,活跃创新研究。

(2) 科研评价指标的设计正面临从基于数量到基于网络结构的跃升,而科研评价维度的确立,将为新指标的研发提供广阔空间。先进的评价指标可为在大量零被引、低被引文献中识别出“睡美人”提供工具,为重要的科学发现提供标识物,为认识科研发展规律提供更多路径。

(3) “创新力-影响力-传承力”的评价维度,可避免一些“评价陷阱”,例如方法类论文往往会获得高被引的“洛瑞现象”^[25];可对于介绍国外概念、理论的非创新成果获得高被引进行解释;而“综述”作为一种特殊的论文类型,也能够多个维度上获得更为恰当的评价。

(4) 多维评价鼓励深层次研究和创新发现,降低信息觅食成本,提高收益,可在一定程度上削弱跟风炒作、重复研究、只重数量不重质量、拆分成果、忽视理论研究和基础研究、研究表面化等不良倾向。

(5) 多维评价有利于学术出版和写作规范。不必鼓励引用“近5年论文”,无需为追求参考文献数量而列举无关成果,遏制过度自引和不当他引,鼓励客观引用。

(6) 评价维度的确立,可帮助参与同行评议的科学家站在更高的视角,更长远地考虑,更清晰地审视评

价客体所具备的真正学术价值。

科研评价体系的完善对于科学王国的健康发展和我国科研实力的进步有着深刻影响。“只有和科学组织基本民主相伴随的、承认每一种思想都有发言机会的多方面的努力,才能保证有无止境的机会使那些尚在为生存而斗争的新思想产生出来。”^{[20]26}而科研评价维度的多元构建将向此迈出积极的一步。

参考文献:

- [1] 楚宾, 哈克特. 难有同行的科学[M]. 谭文华, 曾国屏, 译. 北京: 北京大学出版社, 2011: 126.
- [2] BARBER B. Resistance by scientists to scientific discovery[J]. Science, 1961, 134(347): 596-602.
- [3] ILLINOIS INSTITUTE OF TECHNOLOGY. Technology in retrospect and critical events in science[R]. Chicago: Illinois Institute of Technology Research Institute, 1969.
- [4] COLE S. Professional standing and the reception of scientific discoveries[J]. American journal of sociology, 1970, 76(2): 286-306.
- [5] STENT G S. Prematurity and uniqueness in scientific discovery[J]. Scientific American, 1972, 227(6): 84-93.
- [6] RAAN A F J V. Sleeping beauties in science[J]. Scientometrics, 2004, 59(3): 467-472.
- [7] 李江. 零被引论文都是垃圾吗? [EB/OL]. [2016-10-15]. <http://blog.sciencenet.cn/blog-4792012-994641.html>.
- [8] 罗杰斯. 创新的扩散[M]. 辛欣, 译. 北京: 中央编译出版社, 2002: 86-90.
- [9] 苏畅. 丁肇中谈科学是少数人推翻多数人观念[EB/OL]. [2016-09-11]. <http://js.people.com.cn/html/2012/06/03/1141111.html>.
- [10] 陈超美. 转折点: 创造性的本质[M]. 陈悦, 王贤文, 胡志刚, 等译. 北京: 科学出版社, 2015: 62.
- [11] KOSMULSKIM. Successful papers: a new idea in evaluation of scientific output[J]. Journal of informetrics, 2011, 5(3): 481-485.
- [12] 钱玲飞, 杨建林, 邓三鸿. 人文社会科学学科创新力单指标评价[J]. 图书与情报, 2013(2): 93-98.
- [13] CHEN C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature[J]. Journal of the American Society for Information Science and Technology, 2006, 57(3): 359-377.
- [14] 宋歌. 网络结构视域下的创新潜力指标研究[J]. 图书情报工作, 2014, 58(3): 64-71.
- [15] 宋歌. 科研成果创新力指标S指数的设计与实证[J]. 图书情报工作, 2016, 60(5): 77-86, 124.
- [16] PIROLLO P. Information foraging theory: adaptive interaction with information[M]. Oxford: Oxford University Press, 2007: 14-16.
- [17] GLANZEL W, GARFIELD E. The myth of delayed recognition[J]. Scientist, 2004, 18(11): 8-9.
- [18] 孙学军. 好论文同样会被拒稿[EB/OL]. [2016-10-11]. <http://www.cnki.net>

- tp: //blog. sciencenet. cn/blog-41174-853652. html.
- [19] HOLTON G. Scientific research and scholarship: notes toward the design of proper scales [J]. Daedalus , 1962 , 91(2) : 362 - 399.
- [20] 克兰. 无形学院——知识在科学共同体的扩散 [M]. 刘珺珺, 顾昕, 王德禄, 译. 北京: 华夏出版社, 1988: 26 63.
- [21] GARFIELD E , SHER I H , TORPIE R J. The use of citation data in writing the history of science [R]. Philadelphia: Institute for Scientific Information , 1964: 3.
- [22] 红枫. 论文淘汰率再引关注研究论文越老越吃香 [EB/OL]. [2016 - 11 - 11]. http: //news. sciencenet. cn/htmlnews/2014/11/306886. shtm.
- [23] 宋歌. 学术创新的扩散过程研究 [J]. 中国图书馆学报 , 2015 (1) : 62 - 75.
- [24] NOOY W , MRVAR A , BATAGELJ V. Exploratory social network analysis with pajek: revised and expanded second edition [M]. New York: Cambridge University Press 2011: 244 - 246.
- [25] GARFIELD E , WELLS JAMSDOROF A. Of Nobel class - a citation perspective on high-impact research authors [J]. Theoretical medicine , 1992 , 13(2) : 117 - 135.

Building and Demonstrating “Creativity-Influence-Transmission” Research Evaluation Dimensions

Song Ge

Library and The Institution of Information Science & Technology , Southeast University , Nanjing 210096
School of Economics & Management , Southeast University , Nanjing 210096

Abstract: [Purpose/significance] To form a more objective evaluation of scientific research , stimulate China’s research innovation , and curb academic misconduct , from the perspective of following the law of scientific development , this paper thinks that the dimensions of research evaluation should be diversified and try to extract three dimensions of innovation , influence and transmission. [Method/process] First of all , this paper discussed the three dimensions of the generation mechanism , importance , evaluation predicament and progress , and mutual relations. Secondly , the paper chose indices to represent the three dimensions to analyze their correlation , index of different dimensions associated evaluation type and probability , the role of fusion evaluation of dimension and diffusion stage. [Result/conclusion] The conclusions are as follows: different evaluation dimensions should be established in the index system of scientific research evaluation , and each dimension has rich indexes; the dimension of each index should be clear; multiple indicators should be used in the evaluation , value fusion evaluation. Finally , the evaluation dimensions of “Creativity-Influence-Transmission” , as well as the prospect of multidimensional evaluation in a broader sense , were discussed , including enhancing the flexibility , pertinence , objectivity and explanatory power of scientific evaluation; the discovery of potential talents and innovation , providing broad space for evaluation indexes jumping based on network structure from based on the number ; avoiding the “evaluation trap”; encouraging deep research and innovation and eradicating academic atmosphere , guiding academic publishing and writing norms; enhancing peer insight.

Keywords: research evaluation evaluation dimensions creativity influence transmission power