

不同自然科学门类间论文学术影响力多指标综合评价的合理性研究*

郭红梅^{1,3} 何苗² 邢星¹ 金晶¹ 何钦成¹

¹中国医科大学人文社科学院 沈阳 110001 ²中国医科大学第一附属医院信息科 沈阳 110001

³中国科学院国家科学图书馆 北京 100190

(摘要) 以 2006-2010 年物理学、化学、医学或生物学三个学科门类诺贝尔奖获得者得突出贡献文献(参比论文)和与其研究领域密切相关的论文(非参比论文)为研究对象,在原有研究基础上结合统计学分析结果,从 12 个与论文学术影响力密切相关或一般相关的计量指标中最终筛选出 7 项组成综合评价体系,利用主成分分析法计算每篇论文的综合评价值(以下简称“F 值”)。结果表明,综合评价指标体系及其所获 F 值在总体上相对他引量可以更好地表征高学术影响力论文的品质,并且在一定程度上“纠正”仅以他引量对不同科学门类论文评价时出现的系统偏倚,具有在跨科学门类论文评价、比较上的可应用性。

(关键词) 综合评价体系 跨学科 学术影响力 论文评价

(分类号) G301

The Construction of the Comprehensive Evaluation System and Its Feasible Study on the Academic Influence of Papers Across Natural Scientific Disciplines

Guo Hongmei^{1,3} He Miao² Xing Xing¹ Jin Jing¹ He Qincheng¹

¹School of Marxism, China Medical University, Shenyang 110001

²Information Office of the First Affiliated Hospital of China Medical University, Shenyang 110001

³National Science Library, Chinese Academy of Sciences, Beijing 100190

(Abstract) In this paper, we studied the outstanding contribution literature (reference paper) of Nobel winners in the fields of medicine, chemistry and physics from 2006 to 2010 and the related articles (non-reference paper) of each literature. On the basis of original study and the statistical analysis, we selected 7 indicators from 12 indicators which were related to the academic influence of papers. We calculated the comprehensive value (F value) with the principal component analysis. The result showed the comprehensive evaluation system was better to characterize the quality of high academic papers and could correct the biases of system in different fields. So the comprehensive system was feasible when was used to evaluate papers across disciplines.

(Keywords) comprehensive evaluation system across disciplines academic influence paper evaluation

1 目的和意义

学术论文评价是目前科研绩效评价研究的热点之一,基本方法可分为以同行评议为主的直接评价和应用科学计量学方法及指标进行的间接评价。后者对于论文的评价研究主要集中于论文发表的期刊品质(常以期刊的 IF 值等指标表征)和论文发表后的引文分析

(常以他引量等指标表征)两个方面。论文在某一期刊的发表是作者投稿时对论文研究水平的预判和期刊编辑、论文评阅(审稿)人所做同行评议判定双向交流的结果;论文发表后以他引为代表的引文分析则反映了论文相关的学术界对其学术发现、贡献、价值和影响的接受、认可和利用的范围或程度。在不同的科学门类或学科(以下简称跨门类或跨学科)之间,单独以论文发表期刊的 IF 值或论文发表后的他引量来表征论

* 本文系国家自然科学基金资助项目“自然科学不同学科和研究领域学术论文学术影响力评价比较的可行性研究”(项目编号:70873133)研究成果之一。

收稿日期:2012-06-11

修回日期:2012-09-14

本文起止页码:62-68

本文责任编辑:高丹

文学学术影响力的不可比性和不确定性已为国内外诸多学者所共识^[1-7],而同时对大量跨门类或跨学科的学术论文开展规范、公正和真正意义上的同行评议亦难以实施。

近年来,以实现跨门类或跨学科论文相比较为目的,国内外学者对论文学术影响力的多指标综合评价进行了一些探索,如 Yan Erjia 等的被引频次加权算法^[8]、邱均平等的论文质量指数(Paper Quality Index, PQL)^[9]和吴勤的基于引证强度的论文质量评价方法^[10]等。金晶、徐佳等构建了基于他引量、刊物 IF 值及其派生指标的综合评价体系^[11-12],应用 Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) 法和因子分析法获得论文学术影响力的综合评价值。在同一学科论文中,该综合评价值与他引量等指标的排序趋势相近,还使部分具有高学术影响力特征的论文排序提前;在跨门类或跨学科间论文比较时,该综合评价值与他引量等传统指标的评价结果的基本相同,又较好地避免了因他引量差异性较大而产生的评价偏倚,提高了部分高学术影响力论文的显示度。但是,该综合评价体系中有些指标不易从 Web of Science 等数据库获得。为此,本文以更易于获取数据的指标构建了多指标综合评价体系,并用此体系对 2006-2010 年诺贝尔物理、化学、医学或生理学奖获得者突出科学贡献关联文献的评价结果,验证其在不同科学门类间论文评价与比较中的科学性和可行性。

2 资料来源与研究方法

2.1 参比论文的定义及选取

由于一般认为诺贝尔奖获得者论文具有较高的学术影响力,其质量和学术影响力为世人所公认,所以本文从诺贝尔基金会网站(www.nobelprize.org)的 advanced information 文件中获取 2006-2010 年诺贝尔物理、化学、医学或生理学奖获得者突出科学贡献的关联文献,选择文献类型为 article,作为具有高学术影响力特征的参比论文。

2.2 确定参比论文的相关论文集合

通过每篇参比论文,检索、获得与其研究领域密切相关的全部论文(非参比论文)。对诺贝尔医学或生物学奖参比论文,通过 PUBMED 数据库的 Related Article 获得其相关文献(PUBMED 生物医学数据库的相关文献按主题词相关性构建),保留文献类型为 article;对物理学和化学奖参比论文,通过 Web of Science 数据

库的 Related Articles 获取相关文献后,按文献共被引次数降序排列,从中选择文献类型为 article 的文献。为使各集合的论文在评价体系涉及的各指标上都取得完整数据,并排除因参比论文作者获奖后在该论文引用时可能发生“马太效应”的影响^[13],以诺贝尔奖获得者获奖前两年为时间截点,确定每篇参比论文的相关论文集合。

2.3 建立各相关论文集合引文数据库

在 Web of Science 数据库中检索、下载各论文集合中每篇论文的引文数据并保存为文本形式,施引文献发表时间截取至论文所在集合中参比论文获奖年的下一年。由于相关论文集合确定时以获奖前两年为时间截点,可保证每篇文献都能获得三年他引量的数据。通过自编程序,根据作者姓名和机构排除每篇论文是自引,获得每篇论文各年度的他引数据。根据自行汇总的 1992-2009 年 JCR 期刊 IF 值和期刊更名表,按刊名和 ISSN 号查找论文发表期刊的 IF 值以及各论文施引期刊的 IF 值,1992 年之前论文发表期刊的 IF 值按 1992 年的 IF 值计算。在此基础上,建立各论文的相关论文集合及其引文信息数据库。

2.4 综合评价体系指标的选择

在原有研究的基础上^[11-12],在与论文学术影响力密切相关或一般相关的 12 个计量指标(即总他引量、年均他引量、三年他引量、最大年度他引量、发表期刊 IF 值、发表期刊 IF 值/集合内同年论文 IF 均值、发表期刊 IF 值/集合内同年论文最大 IF 值、施引刊物种类、施引刊物平均 IF 值、施引刊物最大 IF 值、年均他引量/集合内同年论文年均他引量均值和三年他引量/集合内同年论文三年他引量均值)中,综合各指标的相关系数、变异系数,并综合以是否为参比论文为因变量、以 12 个指标为自变量(分别按各科学门类分计和三门类合计)所做的 Logistic 回归分析和判别分析的分析结果,兼顾评价指标体系的系统性和指标数据可获取的简捷程度,最终选定 7 个指标组成综合评价体系,即:他引量(Ct)、年均他引量(Ca)、三年他引量(C3y)、最大年度他引量(Ch)、论文发表刊物 IF 值(IF)、施引期刊种类(Ck)和施引期刊最大影响因子(Cfm)。

2.5 论文学术影响力综合评价值计算方法

利用主成分分析方法计算每篇论文的综合评价值(以下简称“F 值”)。将 F 值与他引量的评价结果进行比较,以验证该综合评价体系用于跨门类论文评价的可行性。

3 结果

3.1 纳入研究的论文有关数据

表1 各门类诺贝尔获奖者参比论文及相关论文集合的基本数据

年度 (年)	医学或生理学				化学				物理			
	获奖 人数(人)	参比 论文数(篇)	建立集合 数(个)	集合中 论文数(篇)	获奖 人数(人)	参比 论文数(篇)	建立 集合数(个)	集合中 论文数(篇)	获奖 人数(人)	参比 论文数(篇)	建立 集合数(个)	集合中 论文数(篇)
2006	2	5	2	185	1	17	9	1 531	2	4	4	1 401
2007	3	25	25	3 354	1	25	7	1 395	2	5	5	1 253
2008	3	41	41	7 298	3	15	6	1 252	3	8	8	2 352
2009	3	13	13	1 823	3	21	5	667	3	4	4	632
2010	1	15	15	2 162	3	12	12	2 345	2	11	11	1 544
合计	12	99	96	14 822	11	90	39	7 190	12	32	32	7 182
施引文献				1 027 619				406 362				431 695

综合评价体系涉及的7个指标在全部论文中,有的与发表时长正相关,有的与发表时长负相关,表明这7项指标组成的综合评价体系受论文发表时长影响较小,可削弱因被评价论文发表时间较长由总他引量、施引期刊种类等指标所导致的累积效应,如表2所示:

表2 各门类论文去重后他引量等7项指标原始值与论文发表年限的 Spearman 相关系数

指标	物理	化学	医学或生理学	三门类
他引量	-0.023	0.197 **	0.026 **	0.028 **
年均他引量	-0.388 **	-0.233 **	-0.211 **	-0.294 **
最大年度他引量	-0.159 **	-0.008	-0.048 **	-0.095 **
三年他引量	-0.199 **	-0.067 **	-0.084 **	-0.135 **
论文发表刊物 IF 值	-0.348 **	-0.194 **	-0.127 **	-0.209 **
施引期刊种类	-0.024	0.164 **	0.009	0.018 **
施引期刊最大影响因子	-0.209 **	-0.029 *	-0.024 *	-0.098 **
论文数	5 738	5 662	10 854	22 254

注: **表示 $P < 0.01$, * 表示 $P < 0.05$; 列入不同集合的同一文献按一篇文献计算论文数。

鉴于各论文集合中论文的7项指标间均有相关关系且方差不齐,KMO 统计量和 Bartlett 球形检验结果表明此数据适宜应用主成分分析,各门类或集合间的比较应以非参数检验为主(167个集合的7指标相关系数、方差齐性检验、KMO 统计量和 Bartlett 球形检验结果从略)。

3.2 跨门类或跨学科论文间他引量等传统评价指标缺乏可比性

各门类不同集合间或三门类间全部论文他引量等传统评价指标差异非常显著(见表3)。即使是获奖论文的关联论文,在不同门类间甚至在同一门类的不同获奖年度间他引量等传统评价指标的差异也非常显著(见表4、表5)。这种非常显著的差异难

以完全用学术影响力的不同解释,提示在跨门类或跨学科间的论文比较中,他引量等传统评价指标缺乏可比性。

表3 各门类不同集合间及不同门类论文间各传统评价指标差异 Kruskal Wallis 检验结果

门类	集合/ 门类数	他引量 卡方值	年均他引量 卡方值	最大年度 他引量卡方值	论文发表刊物 IF 值卡方值
物理	32	439.57 **	2 233.501 **	1 131.136 **	1 606.173 **
化学	39	508.201 **	1 102.817 **	723.745 **	2 182.739 **
医学或 生理学	96	1 269.999 **	1 868.432 **	1 323.217 **	2 837.491 **
三门类	3	424.014 **	546.717 **	259.794 **	527.057 **

注: **表示 $P < 0.01$ 。

表4 各门类获奖关联文献传统评价指标差异 Kruskal Wallis 检验结果

门类	统计量	他引量	年均他引量	最大年度 他引量	论文发表 刊物 IF 值
医学或 生理学 (N=99)	均值	444.7	19.9	47.6	16.6
	标准差	678.2	34.6	74.9	11.1
	全距	7-4481	0.13-247.08	2-503	1.07-52.59
化学 (N=90)	均值	309.7	23.2	40.9	12.8
	标准差	419.7	36.1	51.1	11.3
	全距	4-3104	0.67-194	1-267	0.83-39.19
物理 (N=32)	均值	1005.4	101.4	196.5	7.7
	标准差	1302.4	172.0	356.0	8.8
	全距	2-4338	0.05-707.5	2-1590	0.58-31.85
三门类比较 (N=221)	卡方值	7.12	8.50	9.00	17.90
	P 值	0.029	0.014	0.011	0.000

表5 各门类5个获奖年度获奖关联文献传统评价指标差异 Kruskal Wallis 检验结果

门类	统计量	他引量	年均他引量	最大年度 他引量	论文发表 刊物 IF 值
医学或生理学	卡方值	16.30 **	27.04 **	20.36 **	28.50 **
化学	卡方值	6.295	23.84 **	20.29 **	48.57 **
物理	卡方值	10.81 *	18.11 **	16.60 **	23.32 **

注: * 表示 $P < 0.05$; **表示 $P < 0.01$ 。

3.3 综合评价值未从根本上改变他引量对论文评价的总趋势

F 值与他引量不仅在各集合内论文间高度正相关(各集合数据略去),而且在各门类及三门类所有论文间均具有非常显著的正相关关系,三门类全部参比论文的 F 值与他引量也具有非常显著的正相关关系,如表 6 所示:

表 6 各门类及三门类论文他引量与 F 值的 Spearman 相关系数

门类	医学或生理学	化学	物理	三门类
适合论文(29 185 篇)	0.87 **	0.85 **	0.86 **	0.85 **
参比论文(508 篇次)	0.86 **	0.87 **	0.90 **	0.88 **

注: **表示 $P < 0.01$ 。

在同一门类内,因研究领域的相近,同一篇论文可能会被包含在不止一个论文集合中,某些参比论文亦有此现象。由于每个论文集合都是通过一篇参比论文所建立,而该篇参比论文与此集合中的其他论文涉及的研究领域最接近,彼此间也最具可比性,故将此类参比论文称为“关联集合参比论文”。分别将三门类全部论文按各自的 F 值、他引量由高至低排序后(数值最大者排序为 1,以下简称跨门类序),可见全部论文及关联集合参比论文的 F 值和他引量跨门类序间均高度正相关,如表 7 所示(因表 6、表 7 同为秩相关分析,故相应门类的 Spearman 相关系数相同):

表 7 各门类及三门类论文及关联集合参比论文 F 值和他引量跨门类序的 Spearman 相关系数

指标	医学或生理学	化学	物理	三门类
门类中论文数(篇)	14 822	7 190	7 173	29 185
相关系数	0.87 **	0.85 **	0.86 **	0.85 **
关联集合参比论文数(篇)	90	99	32	221
相关系数	0.86 **	0.87 **	0.90 **	0.88 **

注: **表示 $P < 0.01$ 。

结果说明综合评价体系的评价结果并未从根本上改变他引量评价结果的总趋势。

3.4 综合评价 F 值在不同门类论文间具有可比性

表 8 表示 F 值在门类间的差异与他引量相比已明显减小,表 9、表 10 表示 F 值在集合间、门类间的离均差平方和均为 0,无统计学意义的显著差异,与他引量相比, F 值在用于不同门类间论文的评价、比较时更具有可比性。

表 8 各门类论文 F 值、他引量部分统计量

指标	门类	平均值	标准差	最大值	最小值
F 值	医学或生理学	0.000	1.925	25.903	-2.981
F 值	化学	0.000	1.903	22.129	-3.068
F 值	物理	0.000	1.737	29.867	-1.970
他引量	医学或生理学	69.331	221.027	5 186	0
他引量	化学	56.518	130.243	4 586	0
他引量	物理	60.183	228.456	4 338	0

表 9 各集合、各门类间 F 值方差分析

变异来源	离均差平方和	自由度	均方	F 值	P 值
集合间变异	0	166	0	0	1
集合内变异	102 597.272	29 018	3.536	-	-
门类间变异	0	2	0	0	1
门类内变异	102 597.272	29 182	3.516	-	-
总变异	102 597.272	29 184	-	-	-

表 10 各门类论文间他引量、F 值差异比较的 Games Howell 检验结果

指标	统计量	医学或生物学 vs 化学	医学或生物学 vs 物理	化学 vs 物理
他引量	组间均数差	12.81	9.15	3.76
	P 值	0.00	0.01	0.47
F 值	组间均数差	0.00	0.00	0.00
	P 值	1.00	1.00	1.00

注: 医学或生理学论文 14 822 篇次; 化学论文 7 190 篇次; 物理论文 7 182 篇次。

3.5 参比论文与非参比论文各评价指标跨门类序的比较

按 F 值或他引量对三门类所有论文跨门类排序后,参比论文获得的平均秩次均优于非参比论文,且具有统计学意义的显著差异,表现出高学术影响力的特征,如表 11 所示:

表 11 参比论文与其他论文 F 值、他引量跨门类序 Mann - Whitney U 检验结果

指标	参比论文	样本量	平均秩	Z 值	P 值
他引量	否	28 677	14 792.74	-30.46	0.00
	是	508	3 317.58	-	-
F 值	否	28 677	14 797.75	-31.19	0.00
	是	508	3 034.68	-	-

3.6 关联集合参比论文 F 值与他引量跨门类序的比较

各门类及三门类关联集合参比论文 F 值的跨门类序平均秩次优于他引量的跨门类序平均秩次,如表 12 所示:

表 12 各门类关联集合参比论文综合评价 F 值与他引量跨门类序平均秩次

门类	论文数(篇)	他引量平均秩次	F 值平均秩次
医学或生理学	99	4 066.94	3 368.14
化学	90	3 491.88	3 377.11
物理	32	3 262.05	2 494.72
三门类	221	3 728.61	3 245.33

各门类关联集合参比论文他引量和 F 值跨门类序 Wilcoxon Signed Ranks 检验结果 如表 13 所示:

表 13 各门类关联集合参比论文他引量和 F 值跨门类序 Wilcoxon Signed Ranks 检验结果

门类	跨门类序的排序指标	指标秩次差异	样本量	Z 值	P 值
医学或生理学	他引量(A)	A < B	41	-2.12	0.03
	F 值(B)	F 值(B)	A > B	58	
			A = B	0	
化学	他引量(A)	A < C	29	-2.61	0.01
	F 值(C)	F 值(C)	A > C	60	
			A = C	1	
物理	他引量(A)	A < D	18	-0.26	0.79
	F 值(D)	F 值(D)	A > D	14	
			A = D	0	
三门类	他引量(A)	A < E	99	-2.02	0.04
	F 值(E)	F 值(E)	A > E	122	
			A = E	0	

通过表 13 进一步看出,在医学或生理学和化学门类中,F 值较他引量能使更多的关联集合参比论文的跨门类排序靠前,且二者间的差异有统计学意义;在物理学门类中,他引量只比 F 值使略多的关联集合参比论文跨门类序靠前,但其间的差异并无统计学意义;在三门类中,F 值较他引量能使更多的关联集合参比论文的排序靠前,且二者间的差异也具有统计学意义。

从三门类全部关联集合参比论文 F 值和他引量跨门类序在 5 段百分位数内的论文数可知,在同一百分位数内,F 值的排序比他引量排序可使更多的参比论文进入其中,如表 14 所示:

表 14 三门类全部关联集合参比论文各指标跨门类序各百分位数内论文数

全部论文排序的百分位范围	按 F 值序	按他引量序
5%	109	101
10%	151	133
15%	164	158
25%	187	181
50%	207	201

从各门类 F 值和他引量跨门类序前 1% 论文的门类构成及其中参比论文的门类构成的比较中可见,按 F 值跨门类排序可使他引量均数最低和次低的化学(56.5 ± 130.2)和物理学(60.2 ± 228.5)论文更多地进入前 1%,他引量均数最高的医学或生物学(69.3 ± 221.0)论文数有所减少。可使他引量均数最低和次低的化学(309.7 ± 419.7)和医学或生理学(444.7 ± 678.2)参比论文更多地进入前 1%,他引量均数最高的物理学(1005.4 ± 1302.4)参比论文数有所减少,但进入前 1% 的三门类参比论文数增加。综合评价 F 值可以在一定程度上“纠正”仅以他引量排序时出现的偏倚,如表 15 所示:

表 15 F 值、他引量跨门类序前 1% 论文的构成

论文门类	在 F 值序中论文数(篇)	在 F 值序中百分比	在他引量序中论文数(篇)	在他引量序中百分比
医学或生理学	145	49.5	166	56.7
化学	85	29.0	65	22.2
物理	63	21.5	62	21.2
论文合计	293	100.0	293	100.0
医学参比论文	29	45.3	24	42.9
化学参比论文	23	35.9	17	30.4
物理参比论文	12	18.8	15	26.8
参比论文合计	64	100.0	56	100.0

2009 年诺贝尔物理学奖获得者之一因“有关光在纤维中的传输以用于光学通讯方面”突破性的研究而获奖的 4 篇关联论文的有关数据,如表 16 所示:

表 16 2009 年诺贝尔物理学奖获得者之一获奖关联论文有关数据

参比论文来源文献	发表年	最大他引量年度	Ct	C3y	Cave (年均他引量)	Cmax (最大化他引量)	IF	Ck	Cfm	F 值	Ct 跨门类序	F 值跨门类序
<i>Proceedings of the IEEE</i>	1966	1976	198	3	4.50	12	1.992	105	25.814	5.14	1 825	651
<i>Journal of Scien Scientific Instraments</i>	1968	1973	33	3	0.79	5	0.581	23	20.967	1.26	10 224	3 916
<i>Journal of Scien Scientific Instraments</i>	1969	1973	43	5	1.05	7	0.581	28	4.064	0.78	8 447	5 386
<i>Journal of Scien Scientific Instraments</i>	1968	1971	2	0	0.05	2	0.581	2	1.688	-0.63	24 427	16 516

1966 年在 *Proceedings of the IEEE* 发表的首篇论文(诺贝尔奖励委员会在介绍该奖获奖原因时以大篇幅述及)开创性他引 199 次,而后于 1968 年、1969 年相继发表在 *Journal of Scien Scientific Instraments*(IF 值仅为 0.581)的 3 篇关于不同波长的光在玻璃或其他介质传

输衰减特性测试试验的论文(被诺贝尔奖励委员会提及)他引分别为 2、33、43 次。4 篇论文按他引量跨门类排序在三门类全部论文中分列第 1 825、24 427、10 224 和 8 447 位,而按 F 值排序则分别提前至第 651、16 519、3 916 和 5 386 位,F 值提高了这 4 篇论文学术影

响力的显示度。

4 讨论

4.1 本研究构建的论文学术影响力综合评价指标体系的特色

他引量、发表期刊 IF 值在不同科学门类或同一科学门类的不同学科、研究领域(论文集合)间存在的显著差异不仅可能来自论文学术影响力自身的差异,也有可能源于科学门类、学科或研究领域本身引文行为特征的固有差异。鉴于他引量、发表期刊 IF 值等单一指标在论文评价中的局限性,特别是在跨门类或跨学科间论文评价中缺乏可比性(见表3-表5),国内外研究者开始尝试应用复合指标获得论文学术影响力的综合评价。Yan Erjia 等以论文每次被引的施引刊物论文影响分值(article influence score)和该次引用与论文发表的时间差权重相乘再求和,获得论文的被引频次加权值,因 JCR 对论文影响分值的发布始于 2007 年,故此算法仅适合对引用发生在 2007 年以后的论文的评价^[8]。邱均平等以年度论文发表期刊 IF 值与该年期刊所属分类(category)平均影响因子的比值(旨在消除不同学科期刊影响因子的差异)和论文同年被引次数与同年论文发表期刊所属分类全部期刊论文的平均被引次数(即 JCR 各分类 Aggregate Impact Factor)的比值(旨在消除不同学科论文被引的差异)相乘获得论文年度质量指数,将论文发表后各被引年度的质量指数相加即可获得论文总的质量指数,因 JCR 对各分类 Aggregate Impact Factor 的发布始于 2003 年,故此算法仅适合对引用发生在 2003 年以后的论文的评价^[9]。吴勤以论文所在期刊级别量化值及他引文献的总引证强度之和与自引文献的总引证强度的差值表征论文的质量,其中,每篇他引或自引论文须在建立引证拓扑结构树后,对树中的所有论文节点递归计算其引证强度,且作者仅给出中文期刊级别量化值,故应用的普适性被限制^[10]。本研究选取了具有易获得性、内涵互补性、较好的区分度和高学术影响力论文(诺贝尔奖获得者获奖关联论文)指向性(依据 Logistic 回归分析和判别分析的分析结果)的 7 个指标,组成综合评价体系。7 个指标与论文发表时长的相关分析证明,仅他引量和施引期刊种类与发表时长有一定程度的正相关关系,而另 5 个指标呈程度不等的负相关关系,可较好地避免各集合中论文他引数据获取时间不同对评价结果可能产生的评价偏倚。

4.2 论文综合评价 F 值用于跨门类或跨集合论文间的评价比较具有公正性、客观性和可比性

本研究中每篇论文的综合评价 F 值是在以诺贝尔奖获得者获奖关联论文检索产生的 167 个论文集合各自的范围内,即在研究领域密切相关的论文群体中,经主成分分析法所获得的,符合同类相比的公正性。评价体系的 7 项指标彼此间有程度不同的相关关系,恰是主成分分析法所必需的。指标的重要程度通过各自的因子负荷被体现在主成分的计算式中,是由各指标数据的变异程度及其间的相互关系所决定的,避免了对其主观赋权可能形成的“偏倚”。针对本综合评价体系的指标和研究对象的数据特征,主成分分析法的应用较之以往多指标评价体系综合评价值的计算通常采用的加权求和法更具客观性。表 9、表 10 显示,各集合、各门类 F 值均值皆为 0,集合间、门类间 F 值的离均差平方和亦为 0,各集合间、各门类间的 F 值均值间无统计学差异。F 值这样的数据特征使其被用于对跨门类或跨集合论文比较时具有统计学上的可比性。

4.3 综合评价指标体系及其所获 F 值在不同科学门类间论文评价、比较中的效用

本研究以诺贝尔奖获得者获奖关联论文作为参比论文,其是具有高学术影响力特征的“金标准”论文,从参比论文综合评价 F 值与他引量在全部论文跨门类排序结果的比较中,判断综合评价指标体系及其 F 值的判别效能,具有可信性和可靠性。

除特例外,在同一研究领域甚至同一学科中,以论文他引量的多少表征其学术影响力的大小是被普遍认同的。表 6、表 7 的数据表明,F 值可以保持与他引量评价结果相近的大趋势,使该指标具有可接受性。值得注意的是,在医学或生物学、化学、物理学各门类和三门类合并中,关联集合参比论文 F 值跨门类序平均秩次优于他引量跨门类序平均秩次,且在医学或生理学、化学各门类和三门类合并中有显著差异;在跨门类序的各百分位次内,按 F 值排序比按他引量排序可使更多的关联集合参比论文进入其中;F 值排序不仅使进入全部论文前 1% 的参比论文较他引量排序增加,且可改变仅按他引量排序进入其中的参比论文的门类结构;对于少数仅按他引量排序非常靠后的诺贝尔奖获奖关联论文,F 值排序可显著提高其学术影响力的显示度。因此,本文实证研究的结果表明,综合评价指标体系及其所获 F 值在总体上较他引量可以更好地表征高学术影响力论文的品质,并且在一定程度上“纠正”仅以他引量对不同科学门类论文评价时出现的学科偏倚,具有在

跨科学门类论文评价、比较上的可应用性。

参考文献:

- [1] Seglen P O. Why the impact factor of journals should not be used for evaluating research [J]. *British Medical Journal*, 1997, 314: 498 - 502.
- [2] Cowhig J. Measuring the "impact" of a scientific paper [J]. *中国科技期刊研究*, 2003, 14(2): 117 - 120.
- [3] Zitt M, Ramanana-Rahary S, Bassecoulard E. Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation [J]. *Scientometrics*, 2005, 63(2): 373 - 401.
- [4] Garfield E. The history and meaning of the journal impact factor [J]. *The Journal of the American Medical Association*, 2006, 295: 90 - 93.
- [5] 金碧辉, 汪寿阳, 任胜利, 等. 论期刊影响因子与论文学术质量的关系 [J]. *中国科技期刊研究*, 2000, 11(4): 202 - 205.
- [6] 任胜利, 王宝庆, 郭志明, 等. 应慎重使用期刊的影响因子评价科研成果 [J]. *科学通报*, 2000, 15(2): 218 - 222.

(作者简介) 郭红梅, 女, 1985年生, 博士研究生, 发表论文7篇。

何 苗, 女, 1976年生, 副研究员, 硕士, 发表论文24篇。

邢 星, 男, 1982年生, 硕士研究生。

金 晶, 男, 1983年生, 助理研究员, 硕士, 发表论文9篇。

何钦成, 男, 1947年生, 博士生导师, 硕士, 发表论文150篇。

(上接第47页)

影响力机构拥有了绝大多数的高影响力作者。这也从一个方面证明, 一个机构所拥有的高影响力作者的数量往往可以被作为评价该机构的科研能力的重要因素或指标之一。在农业科学领域的高影响力作者中, 虽然发达国家如美国、英国、法国、比利时、荷兰等国家在H指数、论文总被引次数等方面仍然占有主要地位, 但是中国(包括台湾地区)、巴西等国家在发表论文数和论文总被引次数等方面正逐步缩小与发达国家的差距。中国农业科学领域的科研实力发展迅速。高影响力作者无论在学术卓越性、学术生产力还是学术影响力方面都进入世界前列。中国科学院、中国农业大学、浙江大学和南京农业大学等研究机构也已进入农业科学高影响力机构前20名的行列。

参考文献:

- [1] 王雪梅, 唐裕华, 张志强, 等. 基于文献计量学的优秀华人科学

(作者简介) 阎素兰, 女, 1972年生, 讲师, 发表论文4篇。

杨 波, 男, 1981年生, 讲师, 发表论文10余篇。

何 琳, 女, 1980年生, 副教授, 发表论文20余篇。

- [7] 武夷山, 梁立明. 采用文献计量学指标进行科研绩效量化评价应注意的几个问题 [J]. *中国科技期刊研究*, 2001, 12(2): 100 - 111.

- [8] Yan Erjia, Ding Ying. Weighted citation: An indicator of an article's prestige [J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(8): 1635 - 1643.

- [9] Qiu Junping, Ma Ruimin, Cheng Ni. New exploratory work of evaluating a researcher's output [J]. *Scientometrics*, 2008, 77(2): 335 - 344.

- [10] 吴勤. 基于引证强度的学术论文质量评价方法研究 [J]. *情报学报*, 2007, 26(4): 522 - 526.

- [11] 金晶, 何苗, 王孝宁, 等. 不同学科领域自然科学论文学术影响力评价与比较的可行性研究 [J]. *科技管理研究*, 2010, 30(14): 279 - 284.

- [12] 徐佳, 何钦成. 不同自然科学门类论文学术影响力评价方法可行性研究 [D]. 沈阳: 中国医科大学, 2011.

- [13] 郭红梅, 金晶, 何钦成. 对诺贝尔奖获得者论文施引行为的马太效应初探 [J]. *情报科学*, 2011, 29(6): 830 - 832.

家国际影响力分析 [J]. *情报杂志*, 2010(12): 5 - 9.

- [2] 邱均平, 赵蓉英, 马瑞敏, 等. 世界一流大学及学科竞争力评价的意义、理念与实践 [J]. *科技进步与决策*, 2007, 24(5): 138 - 142.

- [3] 董瑜, 张薇, 袁建霞, 等. 1998 - 2008 农业科学研究影响力分析 [J]. *科学观察*, 2009(6): 9 - 16.

- [4] 韩欣, 刘子忠. ESI 计量分析数据库的功能与科学评价——以中国地质研究机构的科学评价为例 [J]. *地质科技情报*, 2010(1): 138 - 142.

- [5] 邱均平, 杨瑞仙. 基于 ESI 数据库的材料科学领域文献计量分析研究 [J]. *情报科学*, 2010(8): 1121 - 1126.

- [6] 张子倩, 史丽文, 李茂茂, 等. 基于 ESI 的中国农业大学农业科学发展现状及趋势分析 [J]. *安徽农业科学*, 2011(23): 14441 - 14445.

- [7] 杨波, 黄水清, 白振田. 高影响力作者的机构分布模式研究 [J]. *图书情报工作*, 2012, 56(22): 37 - 41.

- [8] 黄水清, 张俊, 阎素兰. 黄金分割法在学科及机构评价中的应用 [J]. *图书情报工作*, 2012, 56(22): 33 - 36.