

文章编号: 2095-1663(2013)05-0063-06

# 卓越导向的大学本科科研评价框架研究

沈秋坦 周荣庭

(中国科学技术大学公共事务学院, 安徽 合肥 230026)

**摘要:** 研究梳理了国内外主要的科研评价方法, 建立了以标准引用影响力、Nature & Science 指数、高引论文和高水平期刊论文四项指标构成的卓越科研评价框架。此框架以客观性、国际标准、质量优先等为特点, 摒弃了传统方法对科研投入及科研产出总量的衡量, 强调对高水平科研成果的评价, 同时, 卓越科研评价框架为大学机构提供了三维立体评价数据, 并且遵循用户驱动的原则, 由用户选择维度和指标。

**关键词:** 研究型大学; 科研评价; 卓越; 文献计量

中图分类号: G40-058.1

文献标识码: A

## 一、引言

研究型大学作为国家创新体系的重要组成部分, 其科技创新能力直接关系到国家整体的创新实力。迈入 21 世纪, 以日、韩、欧盟等为代表的发达国家和地区, 以及中、印等为代表的发展中国家均提出了一流研究型大学建设计划, 而其中提升研究型大学的科研创新能力为重中之重, 如香港“大学教育资助委员会”于 1997 年设立的“卓越学科领域计划”、韩国于 1999 年开始实行的“智慧韩国 21 工程”、日本于 2002 年起实施的“21 世纪卓越科研基地计划”等。与此同时, 主要发达国家也开始重视对本国研究型大学的科研活动的评价, 有些国家甚至将评价结果与经费分配挂钩, 以此来促进大学科研能力的提升, 如英国政府于 1986~2008 年间开展的高校学术研究评估 (Research Assessment Exercise) 和将于 2014 年实行的卓越研究评估框架制度 (Research Excellence Framework)<sup>[1]</sup>, 澳大利亚政府实行的科研质量框架制度 (Research Quality Framework)<sup>[2]</sup>

等。这些国家层面的科研评价均重视对科研产出的质量和社会影响的评价<sup>[3]</sup>。

相比而言, 我国针对研究型大学的科研评价体系尚不完善, 主要体现在“重数量轻质量”<sup>[4]</sup>, 而量质失衡的评价体系在一定程度上充当了指挥棒的作用, 影响了高校的科研管理制度和科研产出质量。据中国科学技术信息研究所统计<sup>[5]</sup>, 中国自 2009 年起连续三年发表 SCI 论文数量位居世界第二, 其中逾八成来自于高校, 2011 年发表 SCI 论文总量占世界比例 11%, 而 Cell、Nature 和 Science 三大名刊的发文量仅占世界的 2.4%, 10 年间 (2002~2012) 被引次数处于世界前 1% 的论文也仅占世界份额的 7.3%。由此可见, 中国虽为“论文大国”却非“论文强国”, 在高水平科研产出上的表现与世界水平仍有较大差距。随着国内高校从规模扩张逐渐转为内涵发展, 其科研战略也从以前的“唯数量”“唯论文”转变为现在的“重质量”“重高水平论文”, 而战略得以成功转变的关键在于先转变传统的科研评价机制。本研究立足于国内外科研评价的理论与实践, 去粕存精, 采用主流的文献计量指标设计了以卓越为导

收稿日期: 2013-06-29

作者简介: 沈秋坦 (1981—), 男, 黑龙江大庆人, 中国科学技术大学公共事务学院博士研究生。

周荣庭 (1970—), 男, 浙江东阳人, 中国科学技术大学科技传播与科技政策系主任, 教授, 博士生导师。

向的研究型大学排名评价框架,一方面坚持使用客观性、国际可比性指标,另一方面坚持质量优先、卓越导向的原则,将重视投入和产出数量的传统科研评价体系转变为重视产出质量的卓越科研评价体系,为科研绩效评价的实践提供理论支撑,旨在进一步促进我国研究型大学在高水平科研评价工作方面的创新。

## 二、已有科研评价方法评述

目前的科研评价方法主要有两类<sup>[6]</sup>,即基于价值理性的质化评价方法和基于工具理性的量化评价方法,常见的质化评价方法主要是同行评议,而量化评价方法主要是采用量化科研指标评价,由于后者具有客观、可比较性强、评价成本低、简单易行等特点,使用最为广泛<sup>[7]</sup>。总的来说,国外开展的国家层面的科研评价通常采用质化和量化相结合的评价方法,而学术界以及大学排名等民间机构的评价通常采用量化为主的评价方法。

### (1) 同行评议

同行评议认为不同的学科应有不同的评价标准,其核心是由同一学科领域的专家对该学科领域的科研产出的学术水平进行评价。目前国际上通行的做法是由接受评价的学术单元(或个人)自报代表性科研作品,然后由专家组对科研作品的学术水平进行打分,最终得到学术单元(或个人)的科研质量得分。同行评议不仅自身存在主观性强、公正性难以保证<sup>[8]</sup>等问题,同时,我国实行同行评议还存在以下难点:其一,我国学术共同体的规范性、自律性不高,如何选择专家组成员,如何保证专家组打分的公平公正,均是实施的难点;其二,通过同行评议评价机构科研水平的工作量异常大,一方面由机构自报数据可能存在失真现象<sup>[9]</sup>,需要人力或者相关制度进行排查,另一方面对每篇(件)代表作进行人工认定本身便是一项浩大的工程。

### (2) 量化指标评价

量化指标评价的相关研究主要有两类,一类是基于“投入-产出”的评价,同时考虑了科研投入和产出指标,主要采用生产函数<sup>[10]</sup>、数据包络分析(DEA)<sup>[11-12]</sup>等方法;另一类是对科研产出的评价,主要采用权重指标法<sup>[13]</sup>。从选取指标来看,科研投入指标主要有科研人员数量、科研经费、科研平台数量等;科研产出指标主要有国内外数据库收录论文

数量及引用情况、专著数量、科技成果奖励数量(国家级、省部级奖)、授权专利数量、科研项目立项数量等。科研的“投入-产出”评价重在解释投入与产出之间的关系,强调科研产出的效率;而科研产出评价则更聚焦于对科研产出规模的评价。不同的评价方法以及指标的选取均有各自的局限性,如DEA方法对指标个数选择的限制<sup>[14]</sup>,权重指标法的评价结果受权重确定方法影响较大<sup>[15]</sup>,等等。

### (3) 大学排名体系中的科研评价

科学研究能力(成果)是国内外大学排名普遍采用的评价维度。表1显示,7种中国大学排名和7种世界大学排名选取的科研指标权重从15%到100%不等。从科研类指标的配置规则看,中国大学排名与世界大学排名主要存在以下两点不同:其一,覆盖面不同,中国大学排名采用的科研类指标通常覆盖科研投入和科研产出多种指标,而世界大学排名则更加集中的使用科研产出指标;其二,导向性不同,中国大学排名重在评价科研投入与产出的规模、科研资源的多寡,仅有两岸四地排名、人大版排名和网大排名同时设置了科研产出总量和人均产出的指标,其余排名均以衡量规模为重,世界大学排名则几乎不考虑科研产出总量的因素,或在设置总量指标的同时,更加重视衡量科研产出的质量,通常采用高引论文、高水平期刊论文和平均被引次数等指标。

虽然不同的评价有不同的指标和方法,但是论文产出评价已经成为国内外科研评价的重要组成部分。基于国际论文引文数据库(如SCI数据库等)的论文数和引证数等文献计量指标是衡量科研产出的通用指标。虽然不少学者认为利用SCI引文数据库评价科研时存在学科之间的不均衡<sup>[16]</sup>、语言歧视<sup>[17]</sup>等问题,但在评价机构的科研产出时仍是首选的评价工具,同时,运用SCI数据库进行评价的价值在于其客观性,可重复性,可检验性以及国际可比性。然而,目前我国采用的文献计量指标雷同性较高、缺乏创新性,多数指标仍停留在科研论文、著作等的发表数量上,缺乏对高水平科研论文评价的实践与学术探讨。

## 三、卓越科研评价框架

卓越科研评价的核心思想是遴选高质量的科研成果作为评价对象,以此来评价大学机构的科研能力。与同行评议通过专家来认定科研成果的水平不

表1 国内外主要大学排名涉及的科研指标及其权重一览表

大学排名	涉及的科研指标	权重
武书连排名	国内外英文数据库论文及引用、学术著作引用、艺术作品、专利授权、科学与技术奖、人文社会科学奖、国家大学科技园	>44.7%
武大版排名	科研队伍与基地(创新团队、国家重点实验室等),科研项目与经费(国家自然科学基金项目、国家社会科学基金项目等),科研产出(专利申请与授权数、国外数据库收录论文数等),成果质量(国家科技奖、人文社科奖、Nature & Science 论文等),效率与效益(人均产出率、万元产出率)	45.31%
校友会排名	重大科研成果(国家科技奖励、Nature & Science 论文等),科学创新基地(国家级实验室、国家大学科技园),基础科研项目(承担 973 首席项目、国家自然科学基金项目、国家社会科学基金项目等)	46.7%
两岸四地排名	科研经费(总量/师均)、顶尖论文(总量/师均)、国际论文(总量/师均)、国际专利(总量/师均)	55.0%
网大排名	国家级实验室、SCI 论文(总量/人均)、EI 论文(总量/人均)、SSCI 论文(总量/人均)、CSSCI 论文(总量/人均)	36.6%
人大版排名	SCI 论文(总量/师均)、SSCI 论文(总量/师均)、CSSCI 论文(总量/师均)	20.0%
HER 排名	国家级实验室、国家科技奖励、承担 973 首席项目、国家自然科学基金项目、国家社会科学基金项目、国家发明专利、科研人力投入、近一年科研论文产出	20.0%
上海交大世界大学学术排名	Nature & Science 论文、SCIE & SSCI 论文、师均学术表现	>40%
泰晤士报高教增刊世界大学排名	研究能力(学术声誉、师均学术论文、研究收入等),篇均论文引用量,师均工业来源研究收入	65%
QS 世界大学排名	师均论文引用次数	20%
莱顿世界大学排名	论文平均引用(MCS)、考虑学科因素的论文平均引用(MNCS)、引用前 10% 的论文的比例(PPtop10)	单指标排名,无权重
台湾高等教育评鉴中心科研论文绩效排名	学术生产力(近 11 年论文数、当年论文数),学术影响力(近 11 年论文被引用次数、近两年论文被引用次数、近 11 年论文平均被引用次数),学术卓越性(近两年 h 指数、近 11 年高被引文章数、当年高影响期刊论文数)	100%
SCImago 科研机构排名	科研论文产出、国际合作论文比例、归一化影响力、前 25% 期刊论文的比例	100%
Webometrics 世界大学排名	前 10% 高引论文数量	15%

注:整理自各大学排名 2011 版指标体系

同,本研究从高水平科研成果在文献计量指标上具有的统计学意义出发,通过引用次数和期刊影响因子来鉴定高水平科研成果。

#### (1) 设计原则和思路

卓越科研评价框架的设计遵循四项原则:保障国际可比性、顾全学科多样性、聚焦产出评价、坚持质量优先。

其一,框架设计全部采用国际可比标准。评价框架不仅是对国内研究型大学科研水平的评价和比较,还应该发挥国内大学和国际一流大学的标杆比较作用。本研究采用的数据均源自汤森路透公司的数据库及数据产品,如 SCI 引文数据库、ESI 基本科学指标、JCR 期刊影响因子报告等,这些数据库拥有全球所有科研机构及大学的详尽数据,便于为中

国高校建设世界一流研究型大学设立世界标准。

其二,框架设计尽最大程度考虑学科因素,将不同学科之间的差异性导致的误差降到最低。本研究拟根据汤森路透的学科分类体系,按学科确定每个学科的高水平期刊、高引论文,实现学科-领域-机构评价的全覆盖;同时根据各大学拥有的进入 ESI 数据库的世界前 1% 的学科(简称“ESI 学科”)数量来划分大学的科研规模层次,ESI 学科数量超过 10 个的为密集型研究型大学,数量介于 5~10 个的为一般研究型大学,数量介于 1~5 个的为教学研究型大学,在进行具体指标数据比较时,将科研规模作为大学分类的依据。

其三,评价指标全部采用科研产出指标。根据大学排名的柏林原则<sup>[18]</sup>,科研产出在一定程度上比

科研投入更能反映大学机构的科研实力。首先,科研产出的数据较投入更容易获得,而且数据质量较易控制;其次,科研产出数据的国际可比性比投入更强,一些如科研人力的投入很难进行国际比较;最后,由于国内高等教育资源配置以及相关科研制度的特殊性,科研投入的多寡往往与大学的经济社会地位相关,并且受外围环境因素影响较大,并不能直接反映大学机构的科研实力。

其四,优先评价高水平科研成果。此原则是对第三条原则的补充,单纯评价科研产出很容易使评价体系陷入“规模论”的泥沼,因此,本研究在设计产出指标时,摒弃传统的科研产出总量或总引用率指标,紧紧围绕高水平科研成果设计高端产出指标,使得一些注重科研质量的高校能在评价中脱颖而出。

#### (2) 指标体系和维度

通过总结国内外大学排名以及国外主要的科研评价体系,本研究根据以上设计原则选择了四项指标来评价我国研究型大学的高水平科研产出,分别为,标准引用影响力(Normalized Citation Impact,简称 NCI)、Nature & Science 出版指数(Nature & Science Publishing Index,简称 NSI)、高引论文(Highly Cited Papers,简称 HCP)和高水平期刊论文(Papers Published on Top Journal,简称 TJP)。

NCI:是目前使用最为广泛的衡量机构论文引用影响力的指标之一<sup>[19]</sup>,它的核心思想是降低不同学科的引用差异在科研评价中的消极影响,将每个学科的引用情况的世界平均值作为基准引用影响力,具体计算方法如下:

假设某机构拥有  $n$  个学科,其在学科  $i$  上发表的论文总数为  $P_i$ ,总计被引次数为  $C_i$ ,学科  $i$  的基准引用影响力为  $R_i$ ,则该机构的 NCI 值为  $\sum C_i / \sum P_i R_i$ ,其中  $i=1,2,3,\dots,n$ 。

NSI:《Nature》和《Science》作为目前最具影响力的综合类世界级期刊,大学在这两种期刊上的发文量通常作为考量其科研实力的重要指标之一。自然出版集团也于 2011 年开始每年一次发布中国地区的自然出版指数,用来衡量中国科研机构(含大学)的科研实力,NSI 借鉴了自然出版指数的计算方法,计算了各大学在《Nature》和《Science》上的发文折合数。

HCP 和 TJP:分别为各学科领域引用次数较高的论文和各学科领域高水平期刊上发表的论文。论文的被引次数和期刊的影响因子分别是评价论文和

期刊质量的常用指标,结合使用 HCP 和 TJP 可以最大限度的搜集高水平论文,使在高水平期刊上发表的引用次数不高的论文和在普通期刊上发表的引用次数很高的论文均纳入对机构高水平科研成果的评价中。

为了获取高水平论文的细节数据,便于评价机构分析高水平科研成果的结构,本研究针对 NSI、HCP 和 TJP 三项指标分别在机构贡献和影响力档次上进行了扩展评价。

机构贡献:由于科研论文作者的排序不同代表其不同的贡献<sup>[20]</sup>,因此,不同于传统排名指标将论文数量简单相加,本研究将机构在 NSI、HCP 和 TJP 三项指标上的得分区分为主创贡献率和合作贡献率,前者即机构发表的第一作者和通讯作者文章数,后者则是将其他合作论文通过 L. Egghe 提供的贡献率折算方法<sup>[21]</sup>折算成的文章折合数,大学机构的总贡献率(即指标得分)为主创贡献率和合作贡献率两项加总。细分贡献率有利于机构分析高水平论文的合作构成,长期应以提升高水平论文的主创贡献率为目标。

影响力档次:为了将论文的“卓越”这主观概念加以客观的标准描述,分析大学机构在多个质量层次上的高引论文数量,本研究将 HCP 和 TJP 指标分别按照论文引用次数和期刊影响因子基准值均分三档,其中 HCP 为 TOP0.1%、TOP1% 和 TOP10% 三档,分别统计了中国大学在 ESI 数据库各学科领域的引用次数分别为前 0.1%、前 1%、前 10% 的高引论文数量;TJP 为 TOP5%、TOP10% 和 TOP25% 三档,分别统计了中国大学在 JCR 数据库中各学科领域影响因子分别排名前 5%、前 10% 和前 25% 的期刊上发表的论文数量。

#### (3) 数据来源及计算流程

卓越科研评价框架的指标中,NCI、HCP 和 TJP 三项指标均遵循从学科-领域-机构逐步归总计算的流程,NSI 指标则直接通过 SCI 数据库采集数据后计算得出。明确学科分类是计算流程中的一个重要步骤,本研究将 ESI 学科和 JCR 学科合并为自然科学、工程技术、生命医药、农业资环、社会科学和交叉领域六大领域。每项指标的数据来源与计算流程如图 1 所示。

#### (4) 结果展示

卓越科研评价框架下的评价结果展示以多样性、系统性为特点,为每个大学机构提供了三维立体

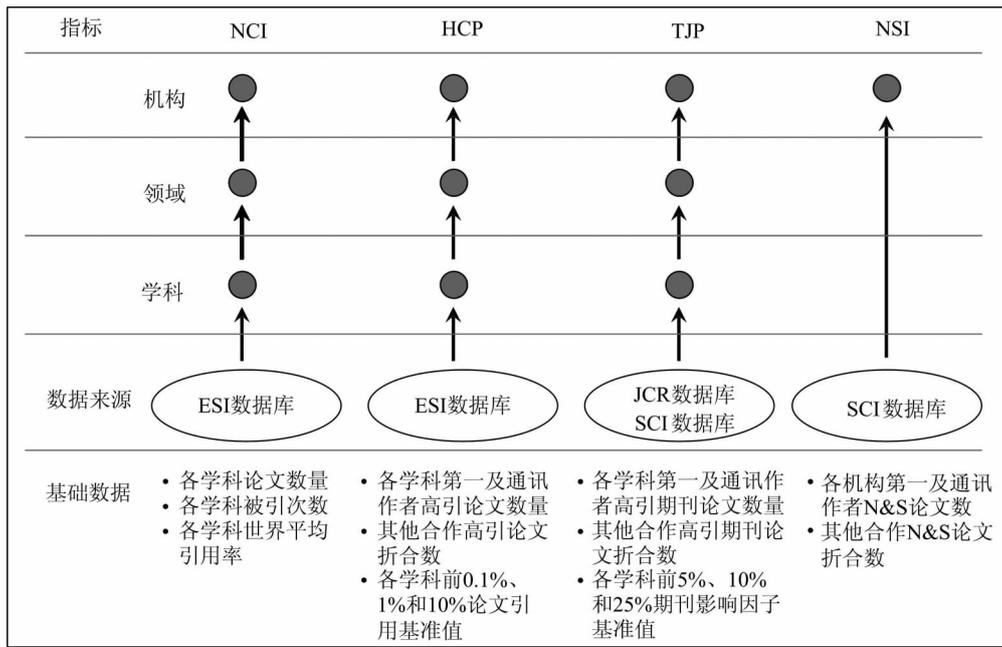


图1 各项指标的数据来源与计算流程示意图

评价数据(见图2),并且遵循用户驱动的原则,由用户选择维度和指标。

#### 四、结论与展望

我国目前还没有成形的受到认可的科研产出评价体系,常见的大学排名在评价科研产出时重视科研产出总量的评价,不足以展示大学在科研产出质量上的优劣。本研究中提出的卓越科研评价框架从引用次数和高水平期刊上全面挖掘高水平科研成果数据,为学科、领域和机构评价提供了全面系统的评价数据。相比传统的量化评价工具,卓越科研评价框架具有以下创新性:

其一,将评价对象聚焦于高水平科研论文上,从被引次数和高水平期刊两个角度挖掘高水平科研论文,基本将水平较低的论文排除在外,有利于对研究型大学高水平科研成果的全面评价;

其二,将贡献率和影响力档次引入评价体系,有效避免了单纯的对高水平科研论文总量的评价,有利于分析研究型大学高水平科研成果的原创性以及产出结构,使基于评价结果的科研战略规划工作更具有针对性;

其三,评价结果的展现形式更加科学和人性化,以指标集和多维度立体为展现形式,有限避免了传统的“指标\*权重”的主观性,有利于用户根据需求获取评价数据或对数据进行二次加工,大大提升了原始数据的利用率和利用价值。

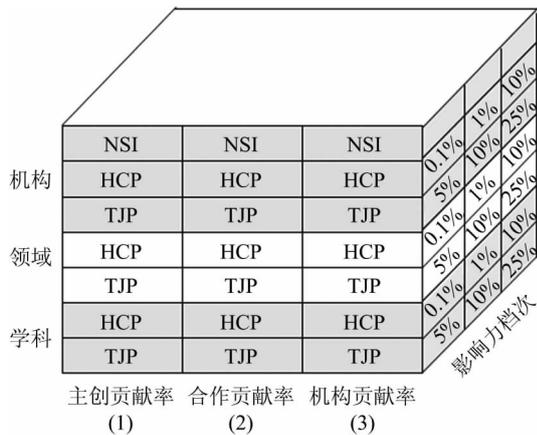


图2 卓越科研评价框架的多维度立体评价效果

其一,三维立体评价。评价框架不仅展示了各细分学科、领域和整个机构的高水平论文产出情况,还展示了不同影响力档次和不同贡献率属性的高水平论文产出情况。

其二,指标集独立展示。由于各项指标之间无法确定孰轻孰重,所以本研究并不采用设定权重再相加的方法展示每个大学机构最终的“卓越科研分数”,而是建立指标集,让用户各取所需,选择自己需要的指标、评价维度、影响力档次以及贡献率属性,最终提取目标数据。指标集的形式有利于进行机构间不同指标的交叉比较分析。

卓越科研评价框架虽然为卓越科研评价提供了一个全新的思路,但仅处于概念设计状态,由于涉及数据量庞大,其实现还需计算机技术、数据挖掘技术等一系列多学科知识的支持与配合。同时,卓越科

研评价框架主要数据来源于汤森路透数据库,对于工程学科、人文社科类科研产出不能很好的反映;评价对象以高水平论文为主,对于论著、专利等其他研究产出没有全面反映。

#### 参考文献:

- [1] 武学超. 英国学术研究卓越的生成逻辑与启示[J]. 中国高教研究, 2012, (6): 22-26.
- [2] 丁宇, 黄艳霞. 澳大利亚 RQF 科研评价制度述评[J]. 科学学与科学技术管理, 2008, (5): 29-33.
- [3] 刘莉. 欧洲各国大学排名评价及其启示[J]. 科学学与科学技术管理, 2005, (9): 86-90.
- [4] 叶继元. 有益遏制学术评价形式化数量化[N]. 中国教育报, 2012-03-28.
- [5] 中国科学技术信息研究所. 2012 年中国科技论文统计结果[EB/OL].
- [6] 隋萌萌. 高校科研评价体系之问题及其实现策略浅析[J]. 教育教学论坛, 2013, 86(5): 263-265.
- [7] 刘贵华, 柳劲松. 教育科研评价的中国难题[J]. 高等教育研究, 2012, 33(10): 25-29.
- [8] 龚旭. 同行评议公正性的影响因素分析[J]. 科学学研究, 2004, 22(6): 613-618.
- [9] 梁立明, 蔡全胜, 李国亭, 等. 再谈高校科研绩效评估的自报指标与源生指标[J]. 科学学与科学技术管理, 1999, 20(3): 13-15.
- [10] 肖晋芬, 仲实. 科研投入产出模型在高校科研绩效评价中的应用[J]. 科技管理研究, 2010, (2): 81-82.
- [11] 王晓红, 王雪峰, 翟爱梅, 等. 一种基于 DEA 和多指标综合评价的大学排名绩效评价方法[J]. 中国软科学, 2004, (8): 156-160.
- [12] 段永瑞, 霍佳震. 基于数据包络分析的高校科研绩效评价[J]. 上海交通大学学报, 2007, 41(7): 1074-1077.
- [13] 李锋, 葛世伦, 尹洁. 高校科研绩效评价模型研究[J]. 科技管理研究, 2009, (7): 271-272.
- [14] 孟激, John Mingers, 刘文斌. 基于 SSM 分析的科研评价系统探讨[J]. 科研管理, 2007, 28(2): 1-8.
- [15] 俞立平, 潘云涛, 武夷山. 比较不同评价方法评价结果的两个新指标——以《泰晤士报高等教育副刊》大学排名为例[J]. 南京师大学报(自然科学版), 2008, 31(3): 135-140.
- [16] 刘艳阳, 吴丹青, 吴光豪, 等. SCI 用作科研评价指标的思考——学科分布对指标公正性的影响[J]. 2003, 24(5): 59-64.
- [17] 卢秉福, 霍丽华. 刍议 SCI 在科研评价体系中的作用[J]. 科研管理(增刊), 2007, 28: 180-183.
- [18] IREG. 高等教育机构排名的柏林原则[EB/OL]. [http://www.ireg-observatory.org/index.php?option=com\\_content&task=view&id=60&Itemid=61](http://www.ireg-observatory.org/index.php?option=com_content&task=view&id=60&Itemid=61).
- [19] Moed, HF. New developments in the use of citation analysis in research evaluation[J]. Archivum immunologiae et therapeuticae experimentalis, 2009, 57: 13-18.
- [20] 王孝宁, 何钦成, 郭继军, 等. 基于文献计量学研究方法的科技论文定量评价[J]. 评价与管理, 2004, 2(2): 22-29.
- [21] Egghe L. A Review of Ranking Problems in Scientometrics and Informetrics[EB/OL]. <http://hdl.handle.net/1942/795>.

## Study on the Research Excellence Framework of Chinese Research Universities

SHEN Qiu-tan, ZHOU Rong-ting

(School of Public Affairs, University of Science and Technology of China, Hefei, Anhui 230026)

**Abstract:** This paper summarizes several research assessment methods, and presents a new research excellence framework comprised of four objective and internationally comparative indicators—the Normalized Citation Impact, Nature and Science Publishing Index, Highly Cited Papers and Papers Published in Top Journals. Unlike traditional methods focusing on the total inputs and outputs of research, this framework emphasizes high level research outcomes. It also provides universities with a three-dimensional assessment model and gives consumers freedom to choose indicators and data based on their needs.

**Keywords:** research university; research assessment; excellence; bibliometrics