

国外科学研究计量评价的研究进展

朱少强

(广东工业大学 经济管理学院, 广东 广州 510520)

摘要:国际上有关科研计量评价的研究比较深入和系统,具体内容涉及到定量评价指标、文献计量与引文分析、命名现象分析、内容分析、学术研究对社会影响力的测度等多个方面,研究人员遍布美、英、澳、荷、丹麦等主流发达国家,其中澳大利亚 REPP 项目的工作尤其引人注目。

关键词:科研评价;科研计量;文献计量;科学学

中图分类号:G316 **文献标志码:**A **文章编号:**1008-5831(2008)02-0072-06

在进行科学及科学项目的评价时,国际上普遍采用同行评议的方式,这也包括对人文社会科学的评价。最早的同行评议源于1416年对专利申请的审查。20世纪30年代以后,美国率先把同行评议引进到科研项目经费申请的评审工作中,此后这一方法为欧美国家广泛采用,成为国际学术界通行的学术水准评价手段^[1]。随着科学计量学日趋成熟,在20世纪70年代末至80年代初,欧美发达国家出现的引文计量法开始向同行评议制提出了谨慎的挑战。早期的计量学诸量化指标主要是为了科学决策而开展的科学质量和科学发展趋势的研究。科学计量学指标真正应用到科学家个人研究成就的评估过程中,大约是20世纪80年代中后期和90年代初开始的。尽管科学计量学起源于西方发达国家,并率先开发出SCI、SSCI、A&HCI等引文计量工具,但首先把科学计量学指标应用于科学家职称与成就评定的则是前东欧社会主义诸国。在西方发达国家,似乎对引入科学计量学指标方法要谨慎和缓慢得多。直到20世纪90年代初,英国高等教育基金委员会才在《1992年科研评价实施条例:评价标准》中明确宣布同行评议是科研评价的基石的同时,同样公开表示他们将向包括文献计量学指标在内的可能作为同行评议补充的评价方法和建议敞开大门。大约也在同一时期,在德国、奥地利、日本等国家的大学或科研院所的职称评定过程中,也程度不同地引进了科学计量学和文献计量学的方法和指标^[2]。

一、定量评价指标

澳洲国立大学社会科学研究院科研评价与政策项目组(REPP),是澳大利亚首屈一指的系统性学术评价中心,其研究的焦点领域为科研绩效和澳大利亚学术研究结构的深层次定量分析。在ARC(澳大利亚研究理事会)、CSIRO(联邦科学与工业研究组织)、NHMRC(国家卫生医药研究中心)等资助下,还定期对科学出版物进行文献计量分析。REPP强调,为科研评价开发一些新的定性和定量指标很有价值。同时,在人文社会科学、工程技术领域,文献计量方法不怎么适用,所以需要为这类学科开发一些其他的测度方法和指标^[3]。

2005年REPP在一份研讨报告中,对英、法、澳等国家所采用的定量学术评价

收稿日期:2007-01-19

基金项目:国家社科基金重点项目“我国人文社会科学评价体系的构建与实证分析”(05AZX004)

作者简介:朱少强(1975-),男,广东工业大学经济管理学院讲师,博士,主要从事信息管理与科学评价等研究。

欢迎访问重庆大学期刊网 <http://qks.cqu.edu.cn>

指标进行了详细的综述研究和理论剖析。提出采用定量指标进行科学研究评价的基础步骤一般有 7 个方面^[4]。

其一,搞清楚基本概念,区分什么是研究质量(research quality)、效果(impact)、学术水平(scientific excellence)等。

其二,对定量指标进行“角色定位”,确定是完全取代同行评议,还是作为同行评议的辅助手段。多数学者赞同后一种观点,即定量指标作为一种信息工具,与同行评议相互补充、相互比较、相互启发,为同行评议提供辅助性的评价信息,尤其要防止科研管理层对定量指标作用的盲目拔高和误用。

其三,指标的选取。每一类指标都有自己的优缺点,所以学者们建议采用多项指标予以综合,且定量指标也不能脱离同行评议等其他评价方式单独使用,在对个人或小范围群体进行评价时尤其如此。如何选取指标和确定指标权重方面,有的是采取向专家学者发放调查问卷、征询意见的办法,有的则建议采用统计分析方法,选择那些相关度最高的指标组合。格伦采尔认为,极有必要开发一些标准化的评价指标,以增加文献计量分析结果的可靠性、有效性,使文献计量数据之间相互兼容。但他同时指出,标准化并不意味着文献计量专家必须遵循同一个标准。

其四,注意因评价而产生的负面后果。如果评价结论影响资金分配或者学术地位(prestige),那么也很可能会影响研究人员及其管理层的日常行为。无论对定性评价还是定量评价都是如此。可能的后果之一是目标的错位,即把在评价中获取高分当作最高行动纲领,而不关注正常的学术研究目标是否达成。学术机构会根据评价规则进行“博弈”(game-playing)。另一个可能的后果是引起学术活动自身的改变,学术研究人员因评价而产生一些适应性行为,比如大量往国际期刊投稿而歧视国内期刊,重复发表论文,申请低风险的、主流的研究项目等。

其五,认识和处理各学科领域之间的差别。不同的学科领域,其出版和引用方式是不同的。就理科而言,发表期刊论文很重要,文献的引文期限较短,但数学引文期限较长。工科论文较少,引文期限最短,专利更加重要,但计算机学科则以国际会议论文及其报告、发言最能反映其成就。文科更加注重专著、作品等,论文只占学术研究成果总量的近一半。同时,即使在同一学科内部,不同专业、研究方向的出版和引用行为也有差别。在一些小学科,由于研究人员总数较少,可能平均被引量就不如那些大学科。理想的办法是把评价限制在一个小的研究领域内,真正遵循“同类比较”(same with the same)的原则。

其六,选择聚合(评价单元)的级别。对于科技政策应以学科领域作为评价对象,还是以研究机构、

研究团体作为评价对象,学者们存在一些分歧。前者是基于认知的观点,后者是基于管理操作的观点。另外,根据统计学原理,要保证评价单元之间的差别不是随机现象,评价单元就不能太小。例如引文量的分布是高度偏斜的:少数论文被大量引用,大多数论文被少量引用,甚至无人引用。如果分析的对象规模太小,就可能导致个别行为决定整体结论的现象。部分学者建议,研究团体的出版物数量在 50 份以上,或每年 10~20 份出版物,3~5 年一评,才能基本满足统计条件。还有人建议如果出版物份数在 10 份以下,应用中位数代替平均值。研究团体被认为是文献计量分析得以成立的最小单元,而对于研究人员个人的评价则是不合适的。

其七,评价结果的呈现。定量分析的结果以什么方式发布,对它的用途、有效性及人们的认可度有相当大的影响。对研究质量的评判没有绝对的标尺,所以一个简单的数字不能揭示任何东西,而必须要有一个对照的标准,以此来评判绩效。这时,统计方法常被用来检验被引率是否显著地高于或低于平均水平。排名表(ranking tables)是比较流行的数据表示方式。这些排名通常用一组定量指标计算得出,标准的文献计量测度如出版物数量、引文数量等常常在这类指标中占据核心地位。许多人尝试为不同类型的指标赋予一定权重,并以此构建一个综合的指数系统。通过综合指数就可以为被评价的对象计算得分,并给出排名表。但这些指标加权体系都不可避免地陷于武断(arbitrary),目前还没有关于权重筛选的经验基础或实验支持。这类排名表通常很能吸引眼球,但往往容易被夸大和误用。

REPP 报告还根据各类研究文献对定量评价指标进行了归纳整理,给出了 40 个指标的名称、描述、现有应用、操作性定义、计算方法、对数据来源的要求等。并将所有定量指标分为出版物指标、引文指标、结构特征指标、其他非文献计量指标 4 类,研究了各类指标的应用范围、局限性及其指标间的相互关联。

澳洲国立大学 REPP 成员的 C Donovan 和 L Butler 将定量评价指标分为质量指标(quality indicators)、影响力指标(impact indicators)两类,其中质量指标又分为标准文献计量指标(规定范围内的学术出版物数量、篇均被引率、用影响因子加权的期刊论文数)、非标准文献计量指标(总被引数、用出版商或期刊声誉加权的出版物数量、论文浏览与下载次数)、文献计量以外的其他指标(荣誉和奖励、学术团体成员、学术团体职务、参加学术会议、杂志编委、访问学者)三类。选取了十几个代表性指标,对经济学领域的 2 个教学型和 3 个研究型的学术机构进行试点统计研究,观察这些指标值是否因学术机构的性质(偏重教学还是科研,以基础研究、应用研究还是政策研究为主)不同而存在明显差异。研究结论认

为,将标准文献计量方法扩展到非 ISI 来源项目,并不能提供额外的有用信息,论文浏览与下载次数不能视为对研究质量的有效测度,学术荣誉指标不能单独作为对研究质量的有效测度,但可以提供一些有用的背景信息。对研究质量、学术荣誉、社会影响 3 类不同级别的指标,应分别予以界定,而不应笼统地混为一谈^[5]。

英国 J S Katz 对指标得分与团体规模的相关性研究成果表明,传统的科研评价指标未能很好地处理科研绩效与机构规模之间的非线性关系,而这可能导致因为规模不同而高估或低估其科研绩效。实证的数据表明,学术认知、学术影响与学术团体的出版规模之间存在幂函数关系(power law relationship),幂指数可能大于 1 或小于 1。同样,在合作性研究的数量与学术机构规模之间也存在幂函数关系。根据上述认识,作者开发了一些与规模无关(scale-independent)的学术绩效测度指标^[6]。

二、文献计量与引文分析法

W Glänzel 与 U Schoepin 以 1993 年 SCI 和 SSCI 年度累积索引作为数据来源,选取“连续出版物占参考文献比例”、“平均引文年限”、“平均引文数”3 个指标,比较了文科与理科的引文差别。通过研究发现,第一个指标是区分文科、理科的敏感特征指标,但数学、技术性学科、一些文科领域等在该特征指标的表现上存在着交叉;在某些文科领域,引文不仅指向专著、预印本、报告等非连续出版物,而且还指向一些非学术性文献。因此,在将标准的文献计量方法移植应用到文科领域时,需要对原有的信息流模型进行重大改进^[7]。

丹麦 J A Wallin 指出了文献计量方法,主要是引文分析方法应用于科研评价的可能缺陷、注意事项、应对措施等。他指出,文献计量方法是天然定量的,但却被用来作出有关定性特征的判断。那么实际上,各种文献计量实践的主要目的,是将“科学质量”等不可见因素转换为一个可操作的实体(manageable entity)。与同行评议相比,文献计量法的优点是很容易得到大规模的应用。尽管许多专家在各类研究文献中提出了警告,但文献计量方法仍然被广泛地、不加选择地应用(uncritical use)于评价和资源分配目的。用引文分析指标来评价科研绩效,是基于这样一个假设,即:科学研究质量与获得的引文数量之间是简单的线性关系。但实际上,引文行为受到学科领域、出版物类型、作者等因素的影响,同时也与科学研究的类型、其长远意义等有关。例如,有关综述、方法论的文献明显具有较高的被引率;一些包含错误观点或论证的文献吸引了比中等质量文献更多的引文;而一些知名的开创性研究,其被引率并没有人们期望的那么高。文献计量方法具有高度的“工具依赖”性(methodology-dependent),但对所有这些指标都有必要进行各种数据的规范化(data normali-

zation)处理。文献计量应用于科研评价目的具有很多局限性(pitfalls),因此需要熟练的技能、批判的眼光,以及对该学科领域的确切了解,才能正确实施文献计量分析与解释分析结论^[8]。

A Vincent 和 D Ross 为评价大学教师的科研绩效介绍了引文分析技术及其指标。手工的引文计量可以计算篇均引文率,引文及时性(recency),作者自引率,专著、期刊论文及其他资料在参考文献中的比例、论文篇均长度、期刊自引率等。而用电子数据库则可以计算排除自引的期刊被引率、影响因子、即年指标和被引半衰期。这些分析的结果可以用来评价大学教师及其研究活动,也可以用作选择学术期刊的依据。必须指出,引文数据并不能取代同行评议,同时必须注意到有许多因素会影响到引文率,如语言、期刊的历史长短与格式、刊期及主题领域等^[9]。

J S Kotiaho 从研究引文行为出发,探讨引文分析方法可能带来的误差。其中之一是引文歧视。如来自第三世界国家的论文可能在引文中受到明显歧视;引文可能会指向某些学术权威、合作者、朋友,或仅仅提供与研究有关的背景资料,而不是指向与论文最相关的文献;还有一些作者姓名排序导致的无意识的歧视。其二是人们倾向于转引综述性文献,而不是原始文献,导致真正的引文链被掩盖。另外还有错引现象,如拼错作者姓名、杂志名称、卷期页码等,导致引文丢失。尽管这些问题并不显著,但由于这些偏差可能会影响对研究人员个人的评价,并涉及到竞争性的基金申请、职位聘任等利益问题,所以运用引文分析作为评价工具时应该持谨慎和批判的态度,考虑各种可能的偏差因素^[10]。

荷兰莱登大学 Anthony F J van Raan 主张文献计量分析可以客观、透明地评价科研业绩,是继同行评议之后不可缺少的评价元素,在研究团队、大学院系、研究所等对象层次上尤其如此。传统上,同行评议被认为是定性的评价方法,文献计量指标则被认为是定量的。但这其实并不绝对,同行评议中就有一些定量元素,比如在权威性期刊上的出版物数量。与之相对,对著作的引用可被视为一种评判,视为同行科学家对该著作“投赞成票”。借助深层次的文献计量分析,不只是提供一般的“附加数据”,而且能迫使专家们重新审视他们的判断,并提供一些挑战性的新见解。尽管应用在较低统计水平,如个人层次上,文献计量的表现不尽如人意;但对于较长时间内一个研究团体的评价来说,它确实是反映研究绩效、尤其研究质量的强有力指标。除“发表论文数”、“总被引次数”、“篇均被引次数”3 个常见计量指标外,Raan 还开发了“未被引证论文的百分比”、“某期刊的篇均被引率”、“世界范围内某学科的篇均被引率”、“国内某学科的篇均被引率”及其相互之间的比值等复杂指标,以 ISI 数据库为工具,分析了德国医学研究所的科研绩效情况。研究得出了一些很有

意思的结果。并认为所谓“国际标准化影响指标”,即“篇均被引次数”除以“世界范围内某学科篇均被引率”的比值,是反映科研绩效的最佳指标(‘crown indicator’)^[11]。

三、命名分析、内容分析等其他评价方法

命名现象(eponymy)是指用原创者的名字为某一现象、定理、理论、原则、发明或方法流程命名。命名可以说是科学界的最高荣誉,因为科学家的成就将进入科学史,并从此不朽。当命名现象成立时,科学家的姓名不再被单独提起,并且/或者也不出现在参考文献列表中,而是整个地成为了某一学科领域内科学交流语言的一部分。命名现象实际上是一种隐含的引用,原始文献的被引率会因命名及对二手文献的引用而被低估。命名现象可以作为一种事后的评价依据,但它的评价周期可能长达数百年,具有历史性,且只适用于那些具有崇高学术地位的杰出学者。A Fernández - Cano 和 I M Fernández - Guerrero 两名西班牙学者利用一部教育领域的人名索引、一个本地数据库及非正式专家访谈对该国教育领域内的命名现象进行了研究^[12]。

英国 R J Ormerod 指出,同行评议与引文分析都有各自的隐含前提:同行评议假定论文被一份有声誉的杂志采用,就证明论文质量高;引文分析方法假定论文被引用证明其学术影响,而学术影响是学术质量的一个次级指标。尽管同行评议与引文分析都有各自的弱点,但发展历史都比较长,应用也很普遍。Ormerod 则尝试着将内容分析法应用于科研评价。他认为,内容分析的价值不在于获得与同行评议或引文分析相一致的结果,而在于因不一致而揭示出来的额外信息。尽管内容分析法比较复杂、耗时,不如同行评议和引文分析有效、可靠和实用,所提供的信息也很有限,但仍然可以作为同行评议和引文分析法的有益补充^[13]。

四、学术研究的社会效果评价

2005 年英国经济与社会研究理事会(ESRC)举办了一个为期两天的“学术研究的学术外效果评价研讨会”(Non - academic Impact evaluation),来自圣安德鲁斯大学管理学院的 H Davies 等提交了一份与会报告,探讨评价社会科学研究的学术外效果的方法^[14]。该文给出了“学术研究的学术外效果”概念定义,即“学术研究结论对政策、管理、专业实践、社会行为或公共舆论(public discourse)的影响。这些影响可以是工具性的,如对政策、专业实践、行为变化的影响;也可以是观念性的,如改进人们关于社会事务的知识、理解、态度等。”总结了两种主要的评价途径。一是前向追踪法,即从研究本身出发,追踪其所产生的后果。学术研究通常透过以下几个方面产生社会效果:知识的生产,如同行评议过的论文;研究能力的提升,如研究生的培养与职业发展;政策或产品开发,如写进政府规划纲要或协议;使特定群体

受益;更广泛的社会效益,如因国民健康或生产率提高而带来的经济效益等。二是在用户群体中理解学术研究的用途。这种方法通常会有典型案例分析,往往还包含对政策制定者进行简单调查。这种方法提供了一定的灵活性,它不仅找出特定的知识传播路径,而且可以探索出一些未知的传播路径、研究者与用户之间的互动及其后果等。

经济学家们试图找出方法,来评价农业研究的经济效果,并为此进行了一些实证研究。但是他们却很少关注如何评价社会科学研究的社會效果。1997 年国际粮食政策研究所(IFPRI)首次举办了有关这一问题的研讨会。2001 年 11 月荷兰外交部与国际粮食政策研究所召集一群研究者,继续就这一主题进行研讨。研讨的结论归为两类,一是如何测量政策导向性社会科学研究的經濟效果,二是如何提升该类型研究对于政策制定的影响力。对于前一个问题,与会人员提出了如下值得关注的方面:评价对象单元的规模;研究绩效的分配(attribution);从供给或需求出发的分析方法;如果某项研究得出意外的结果,那么研究的价值会提升;指标的选择,如经济福利的变化,社会分配结果的变化,文献计量指标,政策变化过程时间的缩短等;时间滞后效应;事前评价与事后评价^[15]。

国际粮食政策研究所的 H E Kilpatrick 认为,政策导向性社会科学研究的市場回报,应该由那些为研究项目付钱的客户来评判。对于特定的研究项目而言,事前与事后的评价都是需要的。事前评价用以决定这项研究是否应受到资助;事后评价则用以决定与之有关的下一轮项目是否应受到资助。事后评价可能基于研究结果所产生的影响而作出。他提出了用以评价社会科学研究的 9 种方法。首先重复原始研究,然后可选择以下评价方法:(1)文献计量分析;(2)同行评议;(3)案例分析;(4)用户评估;(5)成本/收益分析或社会回报率测度;(6)回归分析;(7)决策运筹模型、数据包络分析(DEA);(8)运筹学与计算经济学的最新进展——模型仿真法;(9)各种方法的综合^[16]。

两名农业经济学家 D E Schimmelpenninck 和 G W. Norton 尝试用经济学方法分析农业经济研究所带来的直接效益。他们认为,农业经济研究的最终产出是提供信息,许多研究旨在设计或改进制度,贝叶斯决策理论与经济剩余分析是评估这些信息的可能途径。他们探讨了结合两种方法来对农业经济研究进行实证性评价的优缺点,并利用从关键决策者那里得来的先验和后验概率分析评估了三个案例。将这些概率导向一套“状态/行动”集,显然对于只有一个政策制定中心或具有一致决策程序的情况是最适宜的^[17]。

澳大利亚 C Donovan 总结归纳了见诸澳大利亚及国际上的各种评价指标^[18]。

其一,在评价学术研究的社会效益方面,澳大利亚国内研究机构所采用的指标有:商业化应用、与产业界的联系、为政府提供咨询等。如产业化的收入、商业化披露的次数、孵化产业公司(start-up company)数、用户满意度、国际合作项目数、人力和资本支出、政策提案、专家委员会成员资格及发言邀请、研究生数、博士后数、参与合作研究项目成员数等。

其二,澳大利亚研究理事会(ARC)所采用的评价指标有:合作机构的数量、获得横向资金的数量、作者合著和合作申请专利的次数。该理事会2003年发布一份题为《知识的富足:澳大利亚研究理事会资助研究项目的投资回报》的报告,用纯粹的经济收益观点评价所资助的研究项目。但该报告也指出了社会效益的其他方面,如基础知识体系的构建、技能的提高、决策的改善、卫生及环境与文化的提升等,尽管没有为这些方面拟定评价指标。

其三,澳大利亚国家卫生医药研究中心(NHMRC)的“绩效管理框架”则包含了多维测度。其中“知识利用”之“卫生建议的作用”指标有:咨询建议被非学术性引用的次数、利益相关者对建议与信息的知晓率和满意度。“知识转移进入卫生政策和实践”指标有:在此基础上进行更深入研究的趋势、引起公共卫生政策和实践的变化、承担新的咨询任务。“增进的商业活动”包含商业产出的增长。“交流与合作的加强”指标有:合作与伙伴关系的增长,包括从合作者获得外部资助的水平,建立国内、国际合作及伙伴关系的数量。“社区联系的增进”指标有:为公众及利益相关者提供顾问服务的趋势、与其他机构联系的增多、对该中心产品的需求、利益相关者对该中心作用与建议的了解程度及满意度水平、社区对中心所表述事项的了解及对这些事项态度看法的转变。

其四,在国际上,新西兰、荷兰与英国在开发研究效益评价指标上处于领先。新西兰的科研绩效指标有:为用户开发新的或改进的产品、流程、服务,为用户撰写的报告、出版物和发言(包括各种未经同行评审的非学术出版物、网站、电子文档等)、伙伴关系(含合同、联系)、建立这些联系的总数及简短描述。荷兰采用的“关联度”指标,是指学术研究对社会发展及解决悬问题的影响。英国正在努力构建测度社会效益的方法,提出“学术研究提供者及其用户都应该实际参与到评价过程中来,并就研究(评价)范围给出各自的观点”。英国艺术与人文研究理事会(AHRC)正在开发“一套十分新颖的效益评价方法,该方法基于知识的用户,而不是知识的生产者进行评价”。

五、小结

国际上有关科研计量评价的研究比较深入和系统,具体内容涉及到定量评价指标、文献计量与引文分析法、命名现象分析、内容分析、学术研究对社会

影响力的测度等多个方面,研究人员遍布美、英、澳、荷、丹麦等主流发达国家,其中澳大利亚 REPP 项目的工作尤其引人注目。鉴于科研计量评价在中国科学研究评价体系中占有相当重要的地位,跟踪和研究国外有关科研计量评价的进展状况,及时总结国内外各类评价经验是很有必要的。同时,西方国家尽管在理论与实证研究上都远远领先于中国及东欧诸国,但在科研评价实践中的应用却很谨慎。这种“慢决策、快执行”的策略和态度也很值得学习。

参考文献:

- [1] 刘明. 学术评价制度批判[M]. 武汉:长江文艺出版社, 2006:67.
- [2] 蒋国华. 科学学的起源[M]. 石家庄:河北教育出版社, 2001:249-251.
- [3] Research Evaluation and Policy Project[DB/OL]. [2007-01-07] <http://www.repp.anu.edu.au/>.
- [4] Research Evaluation And Policy Project. Quantitative Indicators For Research Assessment - A Literature Review. Literature Review For Arc Linkage Project; The Strategic Assessment Of Research Performance Indicators[DB/OL]. [2007-01-02] [http://www.repp.anu.edu.au/Literature %20Review3.pdf](http://www.repp.anu.edu.au/Literature%20Review3.pdf).
- [5] Claire Donovan and Linda Butler. Testing Quantitative Indicators Of The Quality And Impact Of Research In The Social Sciences: A Pilot Study In Economics[DB/OL]. [2007-01-02] Repp Discussion Paper 05/2 September 2005. http://www.repp.anu.edu.au/papers/20050912_econ_workingpaper.pdf.
- [6] KATZ J S. Scale - Independent Indicators and Research Evaluation[J]. Science and Public Policy, 2000,27(1):23-36.
- [7] GLÄNZEL W, SCHOEPIN U. A Bibliometric Study Of Reference Literature In The Sciences and Social Sciences [J]. Information Processing and Management, 1999(35):31-44.
- [8] WALLIN J A. Bibliometric Methods: Pitfalls and Possibilities[J]. Basic & Clinical Pharmacology & Toxicology 2005; 97, 261-275.
- [9] VINCENT A, ROSS D. On Evaluation of Faculty Research Impact of Citation Analysis[J]. The Journal of Applied Business Research, 2000 16(2):1-13.
- [10] KOTIAHO J S. Ethical Considerations in Citing Scientific Literature and Using Citation Analysis in Evaluation of Research Performance [J]. Journal of Information Ethics, 2002,11(2):10-16.
- [11] ANTHONY F J, Van RAAN. The Use of Bibliometric Analysis In Research Performance Assessment and Monitoring of Interdisciplinary Scientific Developments [J]. Technikfolgenabschätzung - Theorie Und Praxis/Technolo-

- gy Assessment – Theory and Practice, 2003, 12(1):20 – 29.
- [12] ANTONIO FERNANDEZ-CANO, LNES M FERNANDEZ-GUERRERO. Eponymy For Research Evaluation; Spanish Cases From The Educational Field[J]. Research Evaluation, 2003, 12(3):197 – 203.
- [13] ORMEROD R J. Is Content Analysis Either Practical Or Desirable For Research Evaluation? [J]. Omega 2000 (28): 241 – 245.
- [14] DAVIES H, NUTLEY S, WALTER I. Approaches to Assessing The Non – Academic Impact of Social Science Research[DB/OL]. Report of The Esrc Symposium on Assessing The Non – Academic Impact of Research 12Th/13Th May 2005. [2007 – 01 – 02] http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Images/non-academic_impact_symposium_report_tcm6-16593.pdf.
- [15] INTERNATIOANAL FOOD POLICY RESEARCH INSTITUTE. Impact Evaluation; Assessing The Impact Of Policy – Oriented Social Science Research[EB/OL]. [2007 – 01 – 02] <http://www.ifpri.org/pubs/ib/ib5.pdf>.
- [16] KILPATRICK H E. Some Useful Methods for Measuring The Benefits of Social Science Research[EB/OL]. Impact Assessment Discussion Paper No 5. [2007 – 01 – 02] <http://www.ifpri.org/impact/iadp05.pdf>.
- [17] SCHIMMELPFENNIG D E, NORTON G W. What is The Value of Agricultural Economics Research? [J]. Amer J Agr Econ 85(1)(February 2003):81 – 94.
- [18] DONOVAN C. A Review of Current Australian and International Practice in Measuring The Quality and Impact of Publicly Funded Research in The Humanities, Arts and Social Sciences [EB/OL]. http://www.repp.anu.edu.au/papers/200511_disc03_review.pdf, 2007(1):2.

International Advancement of Quantitative Research Evaluation

ZHU Shao-qiang

(School of Economics and Management, Guangdong University of Technology, Guangzhou 510520, China)

Abstract: This paper introduces international research advancement of quantitative research evaluation, which concerns about several methodologies, such as quantitative indicators, bibliometrics and citation analysis, eponymy and content analysis, and non – academic impact evaluation. Researchers distribute over the western world, and among which the work of Research Evaluation Project and Policy from Australia is especially outstanding.

Key words: research evaluation; scientometrics; bibliometrics; science of science

(责任编辑 彭建国)