世界大学排名方法及其排行榜比较研究

赵蓉英，王　嵩

摘　要：世界大学排名正因逐步影响着各国的高等教育发展战略而备受关注。通过阐述四个主要的世界大学排名方法，从指标体系设计理念、内容涵盖度和排行榜稳定性三个维度对其进行详细对比，并对它们的差异性和稳定性进行比较分析。从中发现，这些主要评价方法使用了差异较大的指标体系和权值来评价世界大学，各自方法特点鲜明；每个评价方法对应的排行榜具有一定稳定性。

关键词：世界大学排名；差异性；稳定性
中图分类号：G649.1　文献标识码：A　文章编号：1671-0169(2015)03-0131-08
DOI:10.16493/j.cnki.42-1627/c.2015.03.017

大学排行通常是运用定性、定量指标来呈现大学多样化表现的手段[1](P13)。与以监督为主要目的大学行业内部所组织的“大学鉴定”和政府组织的“教学水平评估”不同，大学评价或大学排行多以提供信息服务为主要目的[2]，一定程度上为社会各界甄别优质高等教育资源提供参考依据。这些评价或排行依据了一定的标准、采取了一定的方法，对反映高校状况的基本要素进行综合或单项的评价，并按照评价结果与设定标准的接近程度对其进行评分、排序，而纵观目前大学评价方法，其指标设置方式、角度、权重差异明显[3]。

在经济全球化和高等教育国际化的浪潮中，世界和地区大学评估与综合排名的出现也顺应了现实需求。但因各国之间政治、经济、文化甚至体制的差异，以及大学制度、职能、学术表现的差异[4]，在评估的内容、评估的方式以及如何定义质量等方面存在很大差异[5]。目前，主要的世界大学评价体系包括有《泰晤士报高等教育增刊》（Times Higher Education Supplement，THES）的 World University Rankings、《美国新闻与世界报道》（U. S News & World Report，USNWR）的 Best Global Universities、英国 Quacquarelli Symonds（QS）所发布的 QS World University Ranking、上海交通大学世界一流大学研究中心的世界大学学术排名（Academic Ranking of World Universities，ARWU）和武汉大学中国科学评价研究中心（Research Center for Chinese Science Evaluation，RCCSE）的世界大学科研竞争力排行榜。本文将对四种主要的世界大学评价体系和近 5 年的排行榜，从指标体系设计理念、内容涵盖度和排行榜稳定性三个维度对世界大学排名的主要评价体系进行比较研究，从中找出一些规律与启示。

一、评价指标体系设计理念比较

（一）英国《泰晤士报高等教育增刊》

英国 THES 是由 TSL Education Ltd. 公司出版的周刊，其指标及权重体系几经变化，2004 年至

基金项目：武汉大学 2011 年度自主科研项目“科学评价与学术规范数据库系统平台建设及开发利用研究”（114059）
作者简介：赵蓉英，武汉大学中国科学评价研究中心教授、博士生导师（湖北 武汉 430072）；王嵩，武汉大学信息管理学院博士研究生

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

2009 年《泰晤士报高等教育增刊》一直与 QS 合作进行大学排行榜的发布[6]。从 2010 年起，《泰晤士报高等教育》与世界首屈一指的数据公司 Thomson Reuters 合作，由 Thomson Reuters 负责收集和分析所有的与排名相关的数据，同时调整了评价指标体系，设置有 5 个一级指标、13 个二级指标[7]。

1. 注重全面、系统地反映高校的整体实力。THES 认为："……创建世界级一流大学应成为国家政策的重要组成部分……英国的发展想要保持现在的位置就必须采取一定的措施，而排行榜正起到了给高等教育界叫醒的作用。"[8]因此，THES 指标体系涵盖有教学投入、教学环境、教学质量、学术声誉、科研实力、毕业生质量及国际化等诸多方面，从多角度综合反映高校整体实力，并体现高等教育自身发展的规律。

2. 侧重高校科研实力。THES 设置了科研竞争力和科研影响力两项一级指标，且均设置了 30％的权重值，足见其对高校科学研究的重视。此外，THES 不光只注重科研产出的评价，更强调对科研影响力的评价，对应设置了二级指标篇均论文被引次数，并分配了达 30％的最高权重值。

3. 突出学校声誉。THES 的二级指标中，年度声誉调查和与教学有关的声誉调查分别被设置了 18％和 15％的权重。2014 年的年度声誉调查采集了超过 10 000 份的同行评议，而教学有关的声誉调查是基于具有相关经验的专家，对全球高等教育状况调查是源于超过 10 000 份反馈意见的统计结果。

（二）英国 QS

英国国际高等教育研究机构 QS 自 2004 年起，先后与《泰晤士报》和《美国新闻与世界报道》推出大学评价和世界大学评价。2014 年 QS 世界大学排名方法[9]包含有 6 个一级指标。

1. 体现主客观指标的均衡性。从整体看 QS 指标体系的设置，既有定性指标，也有定量指标，且权重各占 50％。QS 设置有两项主观定性评价指标来强调社会声誉调查，分别是学术声誉和雇主声誉，特别是学术声誉由学界评价，权重为 40％。这样的定性评价，解决了一些不能量化因素[10]。其中，2014 年 QS 的学术声誉是基于约 63 700 份同行评议的评审统计结果，而雇主声誉是基于约 28 800 个雇主调查的统计结果。而客观定量指标是生师比例、师均被引和国际化程度这三个指标进行定量评价。

2. 强调办学质量。生师比例指标强调的是教育投入，一定量的生师比可以有效地保证教学质量。QS 认为小班教学有利于每位学生都受到足够的重视与关注，能有更多与教师交流与讨论的机会，因此将生师比作为衡量教育质量的评价指标，并设权重值为 20％。虽然生师比的度量方法还存在较大争议，却是最能反映实际的指标[11]。而基于 Scopus 数据库中科研论文数与教师数计算得到的师均被引数则强调的是教师质量和科研能力。教学是高校发展的基石，科研是衡量高校水平的显著指标，而这些都离不开优秀的教学、科研人才队伍。

（三）上海交通大学世界一流大学研究中心

上海交通大学世界一流大学研究中心（前身为上海高等教育研究所）是最早对世界一流大学进行系统研究的国内机构，于 2003 年首次发布"世界大学学术排名"，并于 2005 年发起并主办了"第一届世界一流大学国际研讨会"，随后每两年召开一次。其排名指标体系主要包括教育质量、教师质量、科研成果和师均表现四个方面[12]。

1. 以提升大学的国际竞争力为目标。随着我国经济的持续高速发展和高等教育强国战略的实施，社会各界对我国建设一批世界一流大学的呼声也越来越高。我国高校与欧美等高等教育发达国家差距何在，如何缩短这些差距以提升我国高校的国际竞争力，尤其是在学术研究方面，ARWU 排名的指标设置充分体现了这些诉求。ARWU 采用在 *Nature* 和 *Science* 上发表论文的折合数和被 SCIE、SSCI 收录的论文数作为科研成果评价指标，采用获诺贝尔奖和菲尔兹奖的校友或教师折合数，及被引用次数最高的科学家数量作为教育和教师质量的评价指标。从这些高质量的评价指标可

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

以看出，ARWU 对高校的国际竞争力十分看重，但在总体上却显得较为粗略[13]。

2. 注重高端教育科研成果。高端教育科研更加注重科研队伍的培养和学术团队的建设，通过国际化交流与合作拓展学术研究视野，提升科研成果水平。ARWU 通过设置教育质量和科研成果两项一级指标，及其对应的在 *Nature* 和 *Science* 上发表论文的折合数、获诺贝尔奖和菲尔兹奖的校友折合数这两项二级指标来体现对高端教育科研成果的重视。

（四）中国科学评价研究中心

武汉大学中国科学评价研究中心自 2005 年起每两年发布一次世界一流大学与科研机构竞争力评价，2011 年后改为每年一次。RCCSE 世界大学评价方法侧重点在于科研竞争力，突出体现在对收录论文数、论文被引次数等指标的权值分配上。2014 年世界一流大学与学科竞争力评价体系做了一定调整，去除了一级指标科研发展力及对应的二级指标高被引论文占有率[14]。

1. 反映高校的综合科研竞争力。RCCSE 设置科研生产力、科研影响力等一级指标强调对科研产出数量和质量的重视，设置二级指标发明专利数来考察高校的科研创新力。此外，还将科研情况与网络排名结合起来，从不同角度反映世界一流大学和一流学科的建设与发展状况，体现高校的综合实力。

2. 注重高校的历史成就和现时水平。RCCSE 的指标体系主要反映高校办学质量和科研质量，注重指标数据的时效性。2014 年选取了进入美国《基本科学指标》ESI 数据库中近 11 年来论文总被引次数排列的 1 830 所大学，评价过程的论文各类指标数据来源于美国 ESI 数据库 2014 年 4 月 1 日—2014 年 5 月 28 日时段的数据，专利指标数据来源于美国德温特（DII）数据库 2009—2013 年的数据。这些指标都具有累积性指标的特点，在反映高校历史成就的同时，也体现出高校现有的科研竞争力。

二、具体指标体系内容涵盖程度比较

世界大学评价指标体系的构建内容是排行榜的决定因素之一。将高等教育的整个环节看作投入—产出这两个环节过程，则四个大学评价指标体系内容存在较大差异（如表 1 所示）。

表 1　世界大学评价具体指标体系内容比较

环节	一级指标	二级指标	THES	RCCSE	QS	ARWU
投入	国际化	国际国内师生比	√		√	
		国际国内师资比	√		√	
	教学环境	教师学生比	√		√	
		博硕学位授予比	√			
		博士学位授予点数量	√			
		师均收入	√			
		师资质量				√
产出	声誉调查	年度学术声誉调查/同行评议	√		√	
		年度教学声誉调查/雇主调查	√		√	
	科研成果	科研收入	√			
		师均论文产出量/卓越度	√	√	√	
		师均科研成果效益/高级别奖励	√			√
		高级别校友/教师数				√
		高级别奖励/论文数		√		√
		国际合作论文数	√			
		发明专利数		√		
	网络影响力	网络排名		√		

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

投入环节反映的是大学在维持生产和获得产出时的所需输入，通常包括了人、财、物这三方面。THES 和 QS 的指标体系中均设置了输入环节的指标，其中均设置了国际国内师生比、国际国内师资比和教师学生比，凸显出这两个指标体系对大学国际化程度的重视。此外，THES 设置了博士学位与硕士学位授予比、博士学位授予点数量和师均收入 3 个二级指标，占其总权重的 11%，体现了 THES 指标对大学软硬件条件的考察。而 ARWU 设置有诺贝尔奖和菲尔兹奖，及高被引科学家数这两项师资质量的评价指标，占其总权重的 40%。

产出环节是对大学教育质量的测算与评估，包含了隐性指标（如声誉、同行评议等）的定性评价和显性指标（如论文数、奖励数等）的定量评价。这四个指标体系都注重了对产出环节的评价。其中，THES 和 QS 均设定了对隐性指标的考察，并分别设定了 33%和 50%的权重值。而科研成果指标则是这四个指标体系均注重的显性指标，只是各自的侧重点不同：THES 是对科研收入和科研产出量的师均数的考量，RCCSE 和 QS 记录的是学术论文的总体量值，而 ARWU 则是强调对高级别奖励和学术论文的总量统计。此外，RCCSE 指标体系中还包含有发明专利数和网络排名这两个二级指标，但权重值较小。

三、世界大学评价排行榜的稳定性比较

从上述分析可以看出，四个世界大学评价方法既存在差异又具有相同之处。这种差异是否会导致其排行榜产生较大差异，下面将对其排行榜从纵向稳定性和横向稳定性两个维度进行定量分析。

（一）纵向稳定性

纵向稳定性是指同一个评价方法的历年排行榜没有太大的变动，即时间维度的稳定性。具体做法是选取这四种排行榜的 2014 年前 100 名作为参比对象，分别生成其在 THES、RCCSE、QS 和 ARWU 排行榜中近五次的排名矩阵，在通过 SPSS 进行相关性分析，结果如表 2 至表 5 所示。

表 2　英国 THES“World University Rankings”排名数据的相关性

		2014	2013	2012	2011	2010
2014	Pearson Correlation	1	0.967	0.927	0.804	0.720
	Sig. (2-tailed)	.	0.000	0.000	0.000	0.000
	N	99	97	97	97	96
2013	Pearson Correlation	0.967	1	0.929	0.816	0.740
	Sig. (2-tailed)	0.000	.	0.000	0.000	0.000
	N	97	97	97	97	95
2012	Pearson Correlation	0.927	0.929	1	0.883	0.767
	Sig. (2-tailed)	0.000	0.000	.	0.000	0.000
	N	97	97	97	97	95
2011	Pearson Correlation	0.804	0.816	0.883	1	0.868
	Sig. (2-tailed)	0.000	0.000	0.000	.	0.000
	N	97	97	97	97	95
2010	Pearson Correlation	0.720	0.740	0.767	0.868	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	.
	N	96	95	95	95	96

**. Correlation is significant at the 0.01 level (2-tailed).

表 2 中整体相关性度较高，但出现比较大的波动是在 2011 年和 2010 年。产生波动的主要因素

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

是指标权值的变化：其中变化最大的是科研竞争力，2010 年其二级指标分别设置为学术声誉调查（权值为 19.5%）、科研收入（权值为 5.25%）、师均论文产出量（权值为 4.5%）和公共研究收入（权值为 0.75%）；考察篇均论文被引次数的科研影响力指标，权值由 2010 年的 32.5%下降到 2011 年的 30%，其后保持不变；将 2010 年的国际化程度指标由国外教师（权值为 3%）和国外学生（权值为 2%），调整为 2011 年的 3 项二级指标，包括国际国内学生比、国际国内师资比、国际化学术合作程度与论文数，其权值各为 2.5%。

表 3 和表 4 显示 QS 和 ARWU 的近 5 次世界大学排行榜相关系数均在 0.9 以上，呈现高度相关。

表 3　英国 QS "World University Rankings" 排名数据的相关性

		2014	2013	2012	2011	2010
2014	Pearson Correlation	1	0.977	0.952	0.912	0.901
	Sig. (2-tailed)	.	0.000	0.000	0.000	0.000
	N	96	96	96	96	96
2013	Pearson Correlation	0.977	1	0.980	0.952	0.936
	Sig. (2-tailed)	0.000	.	0.000	0.000	0.000
	N	96	96	96	96	96
2012	Pearson Correlation	0.952	0.980	1	0.979	0.959
	Sig. (2-tailed)	0.000	0.000	.	0.000	0.000
	N	96	96	96	96	96
2011	Pearson Correlation	0.912	0.952	0.979	1	0.967
	Sig. (2-tailed)	0.000	0.000	0.000	.	0.000
	N	96	96	96	96	96
2010	Pearson Correlation	0.901	0.936	0.959	0.967	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	.
	N	96	96	96	96	96

**. Correlation is significant at the 0.01 level (2-tailed).

表 4　ARWU "世界大学学术排名" 数据的相关性

		2014	2013	2012	2011	2010
2014	Pearson Correlation	1	0.977	0.967	0.968	0.945
	Sig. (2-tailed)	.	0.000	0.000	0.000	0.000
	N	78	78	78	78	78
2013	Pearson Correlation	0.977	1	0.995	0.991	0.963
	Sig. (2-tailed)	0.000	.	0.000	0.000	0.000
	N	78	78	78	78	78
2012	Pearson Correlation	0.967	0.995	1	0.994	0.965
	Sig. (2-tailed)	0.000	0.000	.	0.000	0.000
	N	78	78	78	78	78
2011	Pearson Correlation	0.968	0.991	0.994	1	0.975
	Sig. (2-tailed)	0.000	0.000	0.000	.	0.000
	N	78	78	78	78	78
2010	Pearson Correlation	0.945	0.963	0.965	0.975	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	.
	N	78	78	78	78	78

**. Correlation is significant at the 0.01 level (2-tailed).

因前述 RCCSE 在 2014 年的评价指标中去除了高被引论文占有率，对排行榜产生了一定的影

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

响，因而导致选取的前 100 名参比对象与之前的排行榜有一定的差异，故而表 5 显示出相关系数偏低，呈现中度相关性，但从另一角度体现出 RCCSE 的世界大学评价体系的动态性。

表 5　RCCSE“世界大学科研竞争力”排名数据的相关性

		2014	2013	2012	2011	2009
2014	Pearson Correlation	1	0. 524	0. 709	0. 446	0. 535
	Sig. (2-tailed)	.	0. 000	0. 000	0. 000	0. 000
	N	99	88	88	90	56
2013	Pearson Correlation	0. 524	1	0. 616	0. 133	0. 440
	Sig. (2-tailed)	0. 000	.	0. 000	0. 219	0. 001
	N	88	88	87	87	50
2012	Pearson Correlation	0. 709	0. 616	1	0. 319	0. 450
	Sig. (2-tailed)	0. 000	0. 000	.	0. 002	0. 001
	N	88	87	88	88	51
2011	Pearson Correlation	0. 446	0. 133	0. 319	1	0. 503
	Sig. (2-tailed)	0. 000	0. 219	0. 002	.	0. 000
	N	90	87	88	90	53
2009	Pearson Correlation	0. 535	0. 440	0. 450	0. 503	1
	Sig. (2-tailed)	0. 000	0. 001	0. 001	0. 000	.
	N	56	50	51	53	56

**. Correlation is significant at the 0. 01 level (2-tailed).

（二）横向稳定性

横向稳定性是同一个大学在不同的评价方法的变动情况，是空间维度的稳定性。分别选取 2014 年 THES 排行榜中世界排名前十和中国排名前十的大学进行统计，分析其在这四个排行榜中排名的变化情况（如图 1、图 2 所示）。

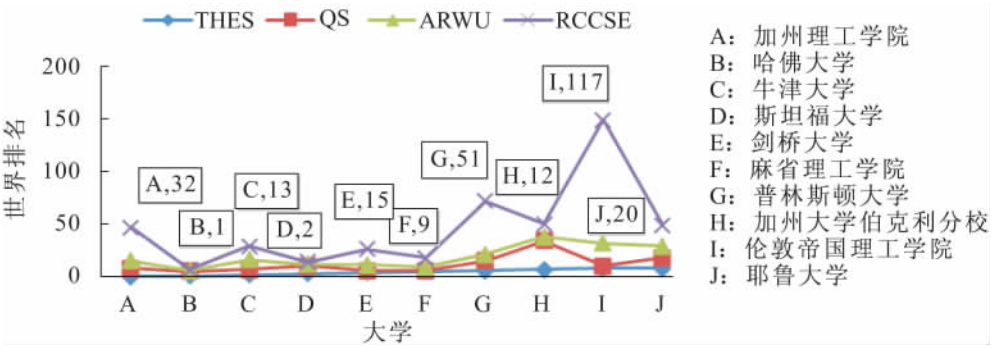

图 1　2014 年 THES 世界排名前十的大学横向稳定性比较

图 1 中所示的这十所排名前十的世界大学在四个排行榜中出现小幅波动。在 RCCSE 排名中，加州理工学院位列 32、牛津大学位列 13、剑桥大学位列 15、普林斯顿大学位列 51、耶鲁大学位列 20。出现波动最大的是伦敦帝国理工学院，排在 117 位，影响其排名的主要原因是科研生产力和科研创新力两个指标得分较低，从而影响了总排名。而剩余其他大学的排名均只在前十名内波动，可见排名前十的大学在四个评价方法的排行榜中的位置是相对稳定的。

图 2 中 2014 年 THES 世界排名中国前十的大学在这四个排行榜中的波动非常大。香港科技大学在 THES 和 QS 中分别位列 51 和 40，而在 ARWU 和 RCCSE 中则分别排在 249 和 289 位。香港

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

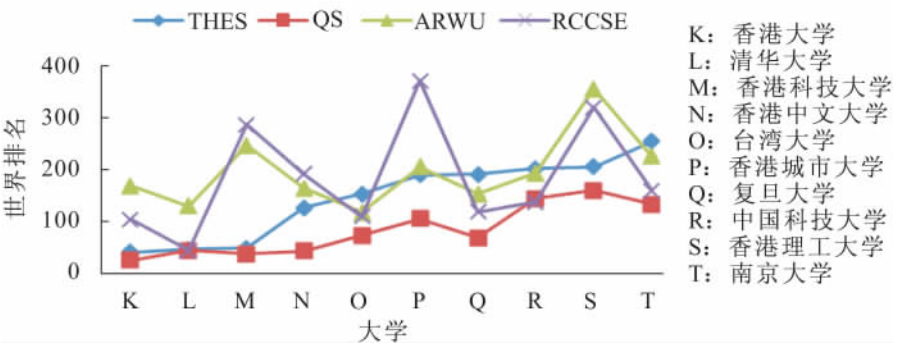

图 2　2014 年 THES 世界排名中国前十的大学横向稳定性比较

城市大学在 THES、QS 和 ARWU 中的排名分别为第 192、108 和 208 位，而在 RCCSE 的排名为第 373 位。香港理工大学在 THES 和 QS 中的排名分别为第 208 和 217 位，但在 ARWU 和 RCCSE 的排名分别为 357 和 322 位。

四、结　语

世界大学的评价和排名有着一系列的社会和商业影响，对于其评价方法和对应排行榜的比较研究有助于揭示世界大学排名方法的特点和规律，也对我国建设世界一流大学具有指导意义。本文的主要研究结论如下：

一是评价方法的多样性。四种世界大学评价体系有着各自鲜明的特点，不同的世界大学评价方法在评价对象选择、评价体系和权值分配，以及如何获取数据源等方面存在很大差异，体现出世界大学评价体系的多样性，也为中国建设世界一流大学提供了综合性的参考，如同行评议，可以引导大学、政府和社会重视和推进大学的内涵建设；又如网站排名，通过大学英文网站可以进一步促进大学拓展国际化程度和公开度，而这一点正是中国大学目前在迈向世界一流大学过程中最为需要却又最为欠缺的部分。

二是评价理念的差异导致评价方法的差异。用一定价值观对高等教育现象或活动的状态进行价值判断，实现评定效益、帮助决策、促进改善和发展。任何独树一帜的评价方法都不可能尽善尽美，多元化、发展性的评价方法为大学发展和公众选择提供了更多的参考，也促进了大学排名方法本身的发展与进步[15]。

三是评价结果的稳定性。尽管评价方法存在不小的差异，但评价结果却有着相对稳定性。纵向稳定性体现了符合高等教育发展规律的稳定变化模式，且这种稳定性随着排行榜次序的下降也依然保持稳定；而横向稳定性则体现出世界一流大学在这四种评价方法中依然保持着优势与稳定，但这种稳定性却随着排名次序的下降而逐渐丧失[16]。

参考文献

[1] Van Vuaght, F., F. Ziegele. (eds.). *Multidimensional Ranking: The Design and Development of U-Multirank*[M]. Dordrecht: Springer, 2012.

[2] 梁爽. 世界大学排行质疑[J]. 煤炭高等教育, 2007, (4).

[3] 殷之明. 大学评价指标的国内外差异分析[J]. 科技进步与对策, 2008, (3).

[4] 叶赋桂. 大学评价和排名：最新的发展及其对大学的意义[J]. 清华大学教育研究, 2008, (2).

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

[5] Usher, A., M. A. Savino. *World of Difference: A Global Survey of University League Tables*[Z]. Toronto, ON: Educational Policy Institute, 2006.

[6] 王博闻，赵伟. 世界一流大学评价的比较分析——以《泰晤士报高等教育增刊》THES 排行榜和上海交大 ARWU 排行榜为例[J]. 当代教育科学，2011，(17).

[7] Times Higher Education. *World University Rankings* 2014—2015 *Methodology*[EB/OL]. http://www.timeshighereducation.co.uk/world-university-rankings/2014-15/world-ranking/methodology, 2014-12-05.

[8] 王珊. USNWR、THES、ARWU 三大大学排行榜指标体系比较[D]. 长沙：湖南师范大学，2013.

[9] QS World University Rankings. *Methodology*[EB/OL]. http://www.topuniversities.com/university-rankings-articles/world-university-rankings/qs-world-university-rankings-methodology, 2014-12-05.

[10]刘强，丁瑞常. QS 大学排名体系剖析[J]. 比较教育研究，2013，(3).

[11]戚巍，缪亚军. 世界大学排名方法解构与比较研究——基于 THE 与 QS 的分析[J]. 研究生教育研究，2012，(3).

[12]ARWU2014 世界大学学术排名. 排名方法. [EB/OL]. http://www.shanghairanking.cn/ARWU-Methodology-2014.html, 2014-12-05.

[13]占侃. ARWU 视角下我国重点高校国际竞争力分析[J]. 中国高等教育评估，2014，(3).

[14]赵蓉英，王嵩，柴雯，等. 发展与梦想——2014—2015 世界一流大学及学科竞争力评价与结果分析[J]. 重庆大学学报(社会科学版)，2015，(1).

[15]上官剑. 稳定性与差异性："QS"与"ARWU"世界大学排名的实证分析[J]. 比较教育研究，2013，(11).

[16][加]亚历克斯·埃舍尔，马斯莫·萨维诺. 差异的世界：大学排名的全球调查[J]. 清华大学教育研究，2006，(10).

（责任编辑　朱　蓓）

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net