教育质量监测评价研究的进展、热点及发展趋势

——基于国内外核心数据库文献的可视化分析

张 佳 谢 丹 谢泽慧

（成都师范学院，四川 成都 611130）

摘 要：本文采用文献计量分析法，对CNKI中CSSCI来源期刊和Web of science数据库中SSCI和A&HCI来源期刊中教育质量监测评价领域研究进行系统梳理和可视化分析。研究显示，该领域研究热度一直呈上升趋势，形成了以马里兰大学、密西西比大学、斯坦福大学、北京师范大学等国内外大学机构为核心的研究网络，研究热点围绕监测评价质量观、监测评价的同质性与异质性、规模化追踪化监测评价、评价效能等内容，延伸到人机动态交互评价等前沿，折射出人工智能驱动的教育监测系统构建的时代需求。我国应构建中国情境及特色的教育质量监测评价的理论和实践研究、采用纵向数据建立追踪研究体系、并协同人的经验与智慧做数据的预测、判断和干预。

关键词：教育质量；教育监测；教育评价；文献计量法

中图分类号：G40-034 **文献标识码**：A **文章编号**：1674-9227(2025)03-0186-10

引 言

当前，中国教育已进入高质量内涵发展的新阶段。建立健全新时代背景下教育质量监测评价体系，需要加强对教育评价与改革的内涵特征研究，以落实“为谁培养人、培养什么人、怎样培养人”的根本问题。《国家教育事业发展“十三五”规划》《深化新时代教育评价改革总体方案》等重要文件明确提出教育监测、评价具有科学育人的导向作用，是构建教育体系中“反馈—调节”的重要手段。我国教育评价在经历了以测量、描述、判断、建构为特征的四代评价之后，正在探索以价值判断、价值促成、价值创新为主要特征的第五代评价。对国内外教育质量监测评价研究的追踪和总结，有助于推动第五代“多元立体”教育评价体系研究的深入和拓展。

监测和评价两个词的内涵虽不一样，但在实践层面，教育监测通过对过程和结果的评估，监控教育发展的过程并发挥指挥棒的作用，而监测中也带有分析与判断，是一种“持续性的评价”。[1]本研究将监测评价界定为整体性概念，是教育监测与教育评价相结合的一种创新手段，是一个独立且连续的过程。

学者们对我国教育质量监测评价做了丰富的研究。一是从理论和实践视角梳理其发展历程。在理论层面，教育质量监测评价研究经历了三个阶段：间续发展阶段（1900—1977年），以西方的教育测量理论为基础，进行系列教育测验；

本文系四川省社科规划项目“新时代四川民族地区教育对口支援政策有效性评价及路径优化研究”（项目编号：SCJJ23ND324）；全国教育科学“十四五”规划教育部重点课题“新发展阶段民族地区教育对口帮扶评价与监测体系研究”（项目编号：DMA210300）的阶段性研究成果。

收稿日期：2025-04-07

作者简介：张 佳（1981—），女，四川遂宁人，博士，成都师范学院教育科学学院教授。研究方向：教育管理、民族教育。

谢 丹（1986—），女，四川南充人，成都师范学院教务处讲师。研究方向：民族教育。

谢泽慧（1972—），女，四川什邡人，成都师范学院副教授。研究方向：教育管理。

中国知网 https://www.cnki.net

理论积累阶段(1977—1985 年),引进大量海外教育评价研究;持续发展阶段(1985 年以后),大规模开展教育评价理论研究。[2] 在实践层面,20 世纪 70 年代末至 80 年代初开始了真正意义上的教育评价活动,经历了恢复和兴起阶段(1977—1983 年)、真正起步阶段(1984—1985 年)、全面研究与试点工作阶段(1986—1989 年)和工作正规化阶段(1990 年至今);[3] 二是对不同主题、不同数据库教育质量监测评价的文献分析。如基于 CNKI 数据库,段鹏阳、杜玲玲[4]对 1988—2021 年区域教育发展水平,檀慧玲、王发明[5]对 2006—2015 年教育政策评估,郭元祥、王秋妮[6]对 1998—2020 年增值评价研究进行梳理;基于 CSSCI 数据库,翟思卿[7]对 2000—2014 年教育评价研究进行梳理;三是对国外教育评价领域 SSCI 期刊中教育质量监测评价主题的发文探讨,如邱均平、欧玉芳[8]和王烁、李昂扬、苏君阳分别基于教育质量监测评价领域的 SSCI 期刊对 2000—2014 年和 2010—2019 年教育质量监测评价研究进行知识图谱分析。

综上可见,其一,现有研究综述多针对某一特定主题进行回顾,对国际学术研究动态探索不够;其二,从研究范围看,国内文献较多而国外文献不足;其三,从研究视角看,纵向分析研究较多,横向比较研究较少。基于此,本研究将厘清国内外教育质量监测评价研究进展,比较当前热点,展望未来趋势,为构建具有中国特色的教育质量监测评价体系提供理论和实践参考。

一、数据来源及研究方法

在 Web of Science (WOS) 数据库高级检索中,以 Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (A&HCI)为检索源,时间截至 2024 年 11 月 16 日,选择 article(含论文和综述类),按论文标题="education quality evaluation" 或者 "educational quality evaluation" 或者 "education monitoring" 或者 "educational quality monitoring" 或者"education quality assessment"或者"education assessment"或者 "education quality appraisal" 或者"education quality review"为关键词,共搜索到 4125 篇文献,阅读摘要后排除与教育质量监测评价主题不相关的文献,将数据导入 CiteSpace 中,选择"Remove Duplicate"进行去重处理,最终得到 3105 篇有效文献作为样本,时间跨度为 1968 年到 2024 年。在 CNKI 数据库中,选择 CSSCI 来源期刊,聚焦教育质量主题,以"教育质量评价、教育质量监测、教育质量评估、教育质量监测评价"为关键词,共搜索到 733 篇文献,剔除书评、短评、会议通知、期刊通知等无效信息,最终得到 715 篇有效文献作为分

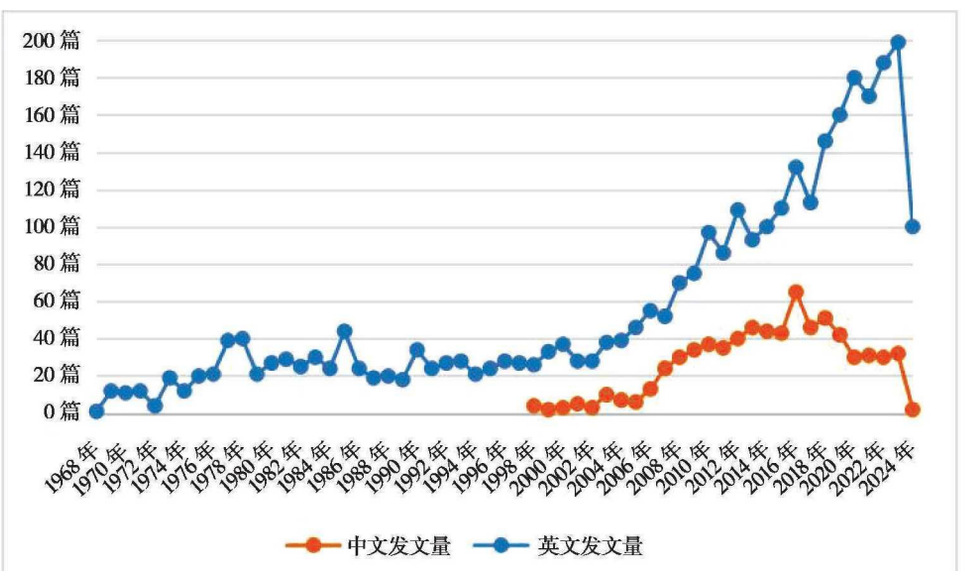


图 1 国际国内核心期刊"教育质量监测评价"相关主题发文量

中国知网 https://www.cnki.net

析样本。采用 CiteSpace 软件为分析工具，研究内容包括国内外教育质量监测评价的研究进展，热点和发展趋势。

二、教育质量监测评价研究进展

（一）文献时间分布特征

通过对发文量和关键词的历史统计，研究表明，国外、国内教育质量监测评价论文发表数量整体呈上升趋势（见图 1）。国内外教育质量监测评价研究呈现阶段式特点，1968 年国外有相关文献发表，1998 年国内有相关文献发表，2010 年之后，伴随着《国家中长期教育改革和发展规划纲要（2010—2020 年）》《中国教育现代化 2035》《深化新时代教育评价改革总体方案》等文件对教育评价改革提出新要求，国内教育质量监测评价的研究增速快，增长幅度明显，富有时代特征和中国特色的教育质量监测评价体系成为研究的主旋律。

（二）研究机构分析

CiteSpace 通过分析不同国家（地区）、机构和作者发文量，将研究机构的合作情况、主题文献发表数量等以节点和连接线的形式呈现出来。[9]研究以 1 年为时间切片，对文献发表的国家（地区）、机构以及作者合作情况进行可视化分析，得到了教育质量监测评价的文献空间分布可视化聚类图谱（见图 2）。如图 2 可见，中介中心性超过 0.1 的代表节点可称为关键节点，是连接网络的枢纽。其中，美国以发文量 938 篇排名第一，中介中心性为 0.8，远超其他国家，其次是澳大利亚（140 篇）和英国（111 篇）。根据节点之间的连线可以看出，形成了以马里兰大学、密西西比大学、斯坦福大学、宾夕法尼亚大学等世界知名高校为主要合作机构的网络。中国发文量（98 篇）排在第四位，中介中心性为 0.03，可以看出，近年来中国的教育质量监测评价研究开始在国际领域占有一席之地，但影响力相对有限。

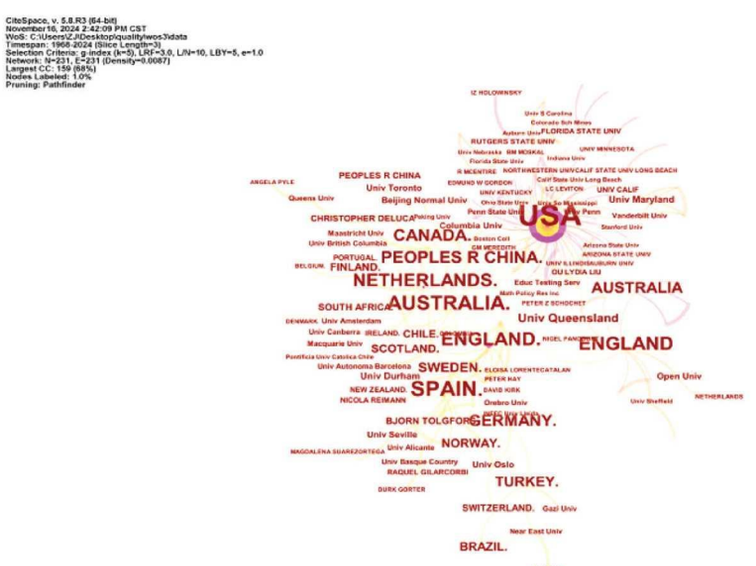


图 2　WOS 数据库中教育质量监测评价机构、国家（地区）网络聚类图谱

国内教育质量监测评价发文量占绝对优势的是北京师范大学（86 篇），其次是北京教育学院（9 篇）、清华大学教育研究院（9 篇）和西南大学教育学部（6 篇）；研究机构方面，形成了以中国基础教育质量监测协同创新中心（44 篇）为主的合作网络（见图 3）；作者方面，形成了以华东师范大学的陈玉琨教授、北京师范大学的辛涛教授、檀慧玲老师与多所高校、行政部门研究者为主开展的合作网络。

（三）发表期刊分析

研究领域有专业期刊的支持是该领域研究成熟且被重视的体现。国外 1897 篇文献（占 61%）集中发表于 *Assessment&Evolution in Higher Education*；*Educational Assessment Evaluation and Accountability*；*Applied Measurement in Education*；*Assessment*；*Studies In Educational Evaluation*；

中国知网 https://www.cnki.net

Evaluation and Program Planning; *American Journal of Evaluation*; *Educational Evaluation and Policy Analysis*; *Educational and Psychological Measurement*; *Evaluation Review*; *International Journal of Selection and Assessment*; *Journal of Educational Measurement*; *Applied Psychological Measurement*12 本期刊上。这 12 本期刊是教育质量监测评价领域的 SSCI 期刊,体现出国际教育质量监测评价研究起步早、成果丰硕、研究领域相对比较成熟等特点。

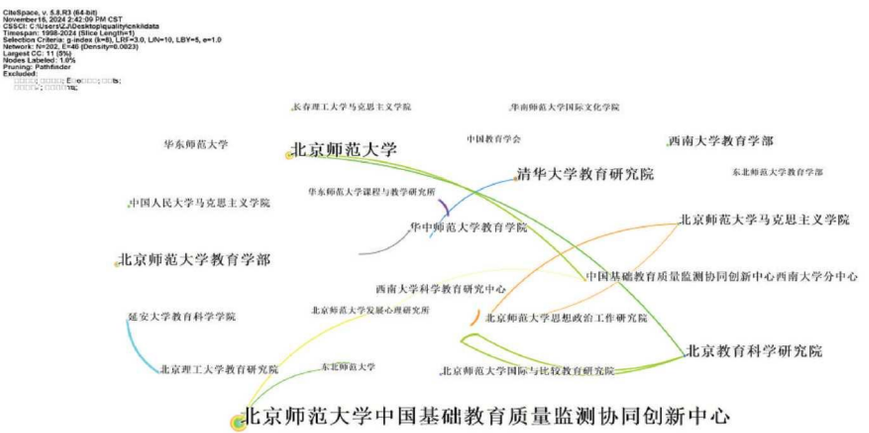


图 3 CSSCI 数据库中教育质量监测评价研究机构合作图

国内聚焦高等教育评价领域的研究成果多发表在《高教发展与评估》,这也是唯一一本质量监测评价领域的 CSSCI 来源期刊,其余文章均发表于 CSSCI 来源中的教育类综合性期刊或综合性大学学报,反映出虽然教育质量监测评价研究是政策热点,但国内学界在期刊的支持力度上还不够。

三、教育质量监测评价研究热点分析

CiteSpace 可以通过论文题目、关键词等提炼出名词性术语,对这些名词性术语进行特征词共线分析,展现相关研究热点。[10]研究设置 CiteSpace 时间切片为 1,选取每个时间切片长度 TOP50 的关键词,利用 Pathfinder 方法进行裁剪,形成国外、国内关键词共现网络图谱。国外教育质量监测评价研究的高频关键词有学生(student)、儿童(children)、教师(teacher)、影响因素(impact)、行为(behavior)、反馈(feedback)、高等教育(high school)以及同伴评价(peer assessment)等。国内教育质量监测评价研究的高频关键词有教学质量、高等教育、核心素养、义务教育及评价体系等。将

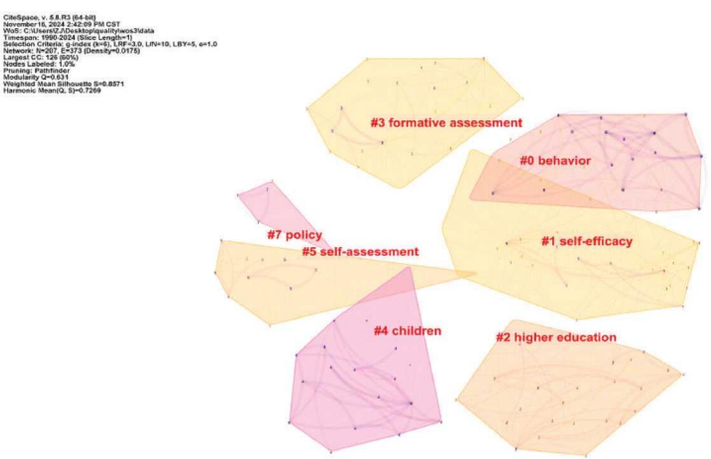


图 4 WOS 数据库中教育质量监测评价关键词、主题词共线网络聚类图谱

中国知网 https://www.cnki.net

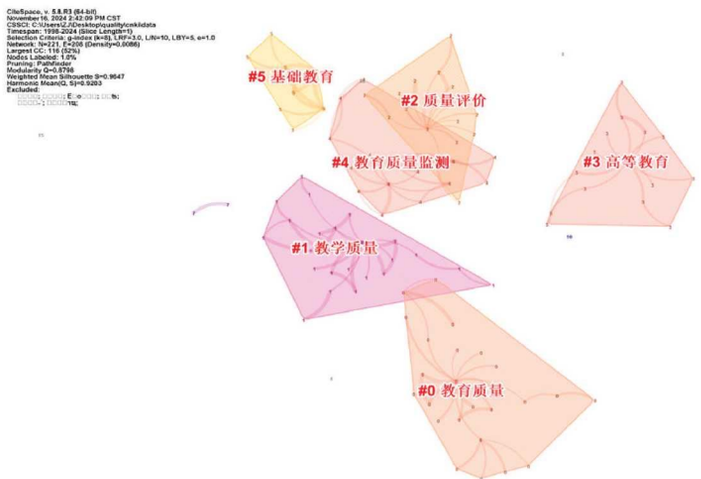


图 5 CSSCI 数据库中教育质量监测评价关键词共现网络图谱

国外和国内的关键词进行聚类分析，选择 LLR 算法，得到关键词聚类图谱（见图 4、图 5），其中，聚类图谱的模块值（Modularity Q）分别为 0.8798 和 0.631，均大于>0.3，平均轮廓值（Mean Sihouette S）分别为 0.9647 和 0.857，均>0.5，说明聚类结果是合理的。

通过对关键词共线图谱中高频关键词的归纳，结合关键词中心性，并考虑聚类图谱，根据同质性标准和文献二次探析，研究发现，国内外教育质量监测评价的研究热点存在一定的相似性，但各有侧重，具体可以划分为以下四个方面。

（一）理论建构从多学科理论的引入到评价质量观的建立

在世界教育领域，教育评价、教育基础理论和教育发展被誉为三大研究课题。国外教育质量监测评价理论研究源于对经济学、心理测量学等理论的借鉴和运用，国内的研究从引入国外理论到中国特色监测评价范式和质量观的树立。

国外教育评价监测理论研究起步较早，发展较为成熟。理论基础源于经典的教育生产函数理论，[11]增值评价理论、项目反应理论和认知诊断理论成为现代教育测量理论发展的主要推动力，这些理论分别从纵向和追踪、多维度能力值以及潜在结构测量等维度，应用于教育质量监测评价领域。

国内教育质量监测评价研究引入国外测量学开展实践，经过 30 余年的发展建立了我国教育质量监测评价理论和方法体系，形成了具有中国特色的教育质量监测评价制度框架和实践模式。教育评价范式从价值基础、模式重构、技术创新等方面实现转换。[12]评价质量观也逐步明确“为谁评价”“评价什么”和“怎样评价”的价值导向作用，并在结果运用上以促进教育发展为导向，兼顾问责性。

（二）评价主体、方式和手段从突破同质性到正视异质性

经过几十年的发展，国内和国外的教育质量监测评价在理论、工具、方法和技术等方面都取得了较大成就，也认识到教育质量监测评价要突破同质性，正视异质性。[13]国内外研究均显示出在不同学段，教育监测评价侧重点不同，在学前教育领域，聚焦课程实施和教学法、师幼互动、幼儿观察等国家和地方幼儿园保教质量评价；在中小学领域，聚焦学生学业过程性数据的采集，突出学习进阶、学业改进角度；在高等教育领域，聚焦 OBE 理念，开展双一流高校、应用型高校、职业院校等不同层次和类型学校整体性办学质量的过程性监测与评估。

国外研究中，政府、学校、社会三方都积极参与到教育质量监测评价领域，从关注目标导向的“结果性评价”转向了注重过程的“形成性评价和增值性评价”。比如，网络形成性评价模型（formative e-assessment）概括了与技术指标相关的形成性评估的理论和实践。[14]田纳西州增

中国知网 https://www.cnki.net

值评估系统(TVAAS)通过学生收益衡量教师教学有效性,[15] 利用分层线性模型和跨案例分析方法对比教师效能与课程观察数据,检验高效和低效教师课堂实践对教师整体效能的影响,[16] 通过建立大规模多变量纵向数据库的增值性评估模型,评估教师培训增值的实践问题等。[17]

国内研究中,监测评价以政府为主导,学界也积极探索社会及第三方评估机构参与的可能性和可行性。在监测评价手段和方式上,探索了 IO(投入—产出)二维模型、IEO(输入—环境—输出)三维模型、CIPP(背景—输入—过程—结果)和 CIPO(背景—投入—过程—产出)四维模型等,这些模型广泛地应用于高等教育质量监测评价、研究生教育质量监测评价、继续教育质量监测评价、基础教育质量监测评价等多个领域。

(三)评价范畴从单点标准化到多领域跨年度追踪型评价

以教育质量监测评价的理论和模型为构建基础,国内外都建立了不同学段的指标体系,对基础教育、高等教育等领域开展大规模、追踪式的监测评价。

国外而言,在基础教育领域,国外的基础教育质量监测通过智能课堂智能测量采集学生完成作业中生理和行为特征,数据的采集指向学生的连续性、多模态数据,从认知技能和非认知技能层面追踪学科素养的发展。较为知名的评价项目有国际经济与合作组织(OECD)于 2000 年开展的国际学生评价项目(PISA),国际教育成就评价协会(IEA)主导的国际数学和科学教育趋势研究项目(TIMSS)和国际阅读素养进展研究项目(PIRLS),美国国家评价管理委员会(NAGB)、国家教育统计中心等主导的国家教育进步评估项目(NAEP)、高等教育领域的 OBE(outcome-base education)项目[18]等。学者们对 PISA 项目的测试数据也展开了分析,如 Dincer 和 Uysal 以土耳其 PISA2006 年的科学素养测试结果为依据,运用广义最小二乘回归的方法分析土耳其学校学生成绩的影响因素。[19]Jehangir 和 Glas 利用多层线性模型的方法对 PISA2009 年的测评数据进行分析,探索父母的社会经济地位与学生阅读成绩之间的关系。[20]Cortez 和 Silva 采用回归树和随机森林等数据挖掘方法,将葡萄牙中学生的数学和阅读成绩与学生的个体特征进行关联分析。[21]标准化的评估结果被用于政策制定,确保在不同学习动机下,学生能实现收益最大化。

国内而言,虽然大规模的基础教育质量测评起步较晚,但随着中国基础教育质量监测协同创新中心的成立,基础教育领域监测评价的理论和实践研究已如火如荼开展,大规模的测试包括 PISA 在 2009、2012、2015、2018 年开展的四轮独立研究,以及教育部基础教育质量监测中心开展的中小学教育质量监测项目。在综合分析学生学业成绩的基础上,纳入了学生学习环境、教学环境、教师教学过程等因素。针对 2009 年在中国开始的 PISA 测试项目,国内学者认为 PISA 数据的国际性、跨年度可比性和评价的全面性对我国建立教育质量监测评价数据具有重要的借鉴意义。[22]在高等教育领域,有效支撑高等教育高质量发展是国内高等教育评估面对的复杂问题。在高等教育普及化阶段,经历了本科学士学位评估、审核评估之后,高校面临着生源特征转变、信息技术变革等机遇和挑战。在新一轮审核评估开展的背景下,高等教育质量评价范式亟待转变。[23]我国学者也从政策分析角度,探索创建世界一流大学的评估指数,平衡“中国特色”与“世界一流”两大特色。同时,也提出未来高等教育的评估工作要坚持动态性、开放性和交互性。[24]

国内各类教育政策的制定和实施是教育治理的主要方式,随着乡村振兴战略下民族地区教育对口帮扶政策力度的加大,对民族地区教育政策有效性的评价、学校教学质量的连续监测追踪成为国内学者重点关注领域,民族地区教育质量监测评价有其独特性,具体体现在监测目标、监测对象、监测内容、监测工具和结果反馈等五个方面。学者们从宏观层面的理论研讨、中观层面的框架构建和微观层面的实证分析三个层面着手,增加了语言使用、宗教与生活习俗差异、人口迁移、居住格局、交友、族际通婚和民族文化多样性等监测评价维度。针对我国特有的教育对口帮扶政策,学者们通过评估其政策的有效性,认为这是在“公平、优质、有效”的基础上促进民族地区教育均衡发展的重要保障,是提高民族地区教育社会认可度、增强民族地区学生自信心的有效途径。

(四)评价对象从课堂到教师教学和学生学习成效的落地

教师教学和学生学业成绩的质量监测评价也是国内外学者关注的热点。不管是学前教育、基础

中国知网 https://www.cnki.net

教育还是高等教育，对教师教学效能、学生学习成绩的有效性评价都是质量监测的落脚点。

就国外而言，教师教学质量监测评价基于“效能”“成长”“增值”的评价理念。评价工具和评估模型的引入对教师教学和学生学习评价的有效性具有重要作用。有学者在西班牙两所大学的实践课程中使用了 eportfolio 工具来评估学习的认知、态度和行为，推断出 eportfolio 作为个人发展学习工具是有效的。[25]有研究认为，摒弃传统的学习和评估模式，通过笔试测试学生定义、分析、解决真实问题的能力，通过同行评议测试学生的合作能力更符合学生认知。[26]针对美国学生选择不参与 k12 学校评估的人数逐年增加的情况，研究者评估了选择“退出模式”类型的影响差异。[27]欧洲高等教育领域研发出大学生学习能力融入就业市场需求的评估模型。[28]反馈是监测评估闭环系统的最后环节，但反馈的类型和方式会产生不同效果。[29]与教师反馈相比，同行反馈或评价会更及时，因此，鼓励教师将同行反馈或评价纳入学生的评价实践中。[30]如 21 世纪技能评估与教学项目（The Assessment and Teaching of 21st Century Skills），通过对协作解决能力的认知以及社会抽样工具的重新开发，纳入同行评议，确定学生内部和学生之间能力水平差异，为教师提供形成性反馈。[31]

国内研究对课堂评价、教师教学效能评价、学生学业成绩的评价经历了系列实验和改革。理念上，教—学—评一体化理念被认为有利于落实课程标准，[32] 教师教学效能和学生学业评价是学校对课堂教学活动价值进行判断、挖掘和提升的过程。[33]实践中，20 世纪 90 年代开展的“新基础教育”改革，课堂教学评价经历了从诊断式和常规式评价、原则性与比较性评价到全程整体性评价与阶段系统性评价三个阶段。[34]360 度绩效反馈评价法、弗兰德斯互动分析系统（Flanders Interaction Analysis System，FIAS）[35]等方法开启了多维数据评价方法的运用。

四、教育质量监测评价研究演进趋势

通过 Citespace 提供的时区图功能，可以展示研究主题随着时间演进的趋势。同时，采用 kleinberg 突现检测算法从文献记录的题目、摘要、关键词中抽取突现词，分析它们在不同区间内出现频次的突然变化，识别出代表研究前沿的若干名词术语，频次变化率高的突现词作为前沿术语。[36]国外研究的突现词包括“formative assessment”“university”“classroom”“perception”“science”，结合国外热点演进（见图 6）得出“academic performance”“educational policy”“assessment literacy”是国外研究前沿术语。国内研究的突现词包括，“核心素养”“质量评价”“教育评价”“基础教育”，结合国内热点演进（见图 7）得出“未来教育”“智慧监测”“质量观”是国内研究前沿术语。结合国内外前沿术语的支撑文献，可以了解到国内外教育质量监测评价研究的趋势。

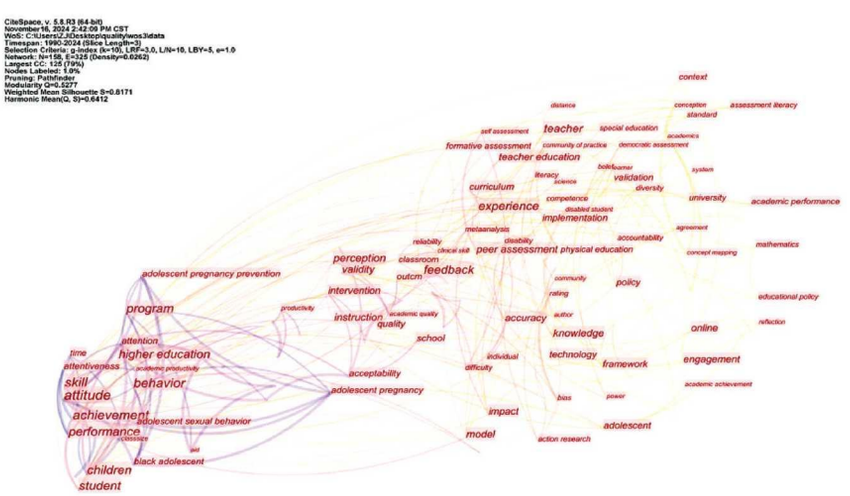


图 6 国外教育质量监测评价文献时区图谱热点演进图

中国知网 https://www.cnki.net

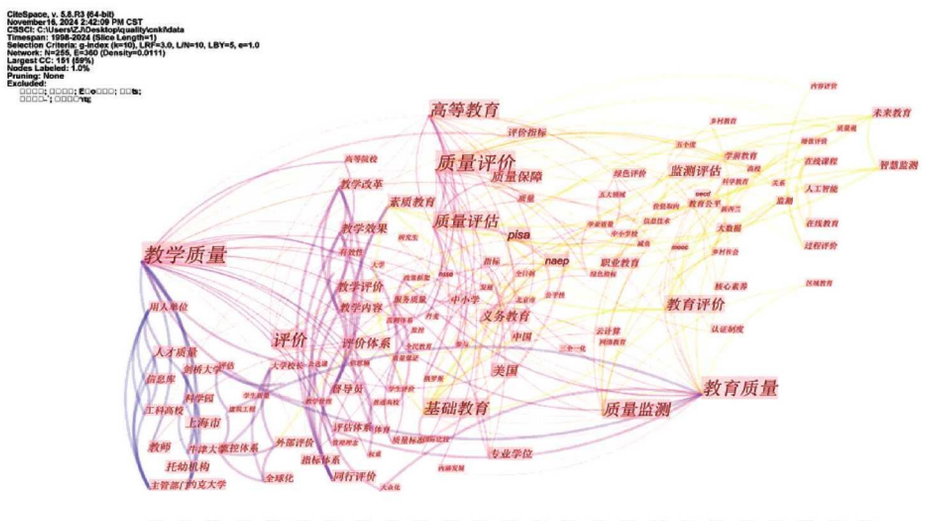


图 7 国内教育质量监测评价文献时区图谱热点演进图

（一）教育评价质量观的演进蕴含着不同国家评价文化的价值取向

各国对教育监测评价理论和实践层面探索和改革的目标，是通过评价改革监测效能，增强评价的有效性及应用，并通过应用反馈迭代监测评价工具，在量化评价和质性评价范式之间切换与融合，突出“育人为本”理念。学前教育、基础教育、职业教育、高等教育等不同类型的监测评价都蕴含着国家质量评价文化的价值取向。美国的监测评估通过引入第三方社会组织，形成了一种政府与民间合作，不依靠国家控制而依靠社会力量的模式。加拿大的监测评价体系遵循循证评价文化，突出标准化与质性的特点。我国新时代的教育质量评价是助推教育强国建设的时代需要，监测评价体系的建设、完善和更新伴随着国家教育发展战略进程，教育质量监测评价价值观在社会主义核心价值观的引领下，不仅发挥着“培养什么人”“如何培养人”的“指挥棒”作用，也是促进教育改革与发展的支点。

（二）核心素养理论与实践融合研究是监测评价的新形态

自 1997 年世界经济合作与发展组织实施素养遴选项目以来，世界主要国家和国际组织都先后启动了围绕“核心素养”的教育目标变革。国外和国内政府都充分认识到了素质、知识、能力和情感态度价值观是人全面发展中的关键要素，都将核心素养的培养作为推动教育改革和评价的重点，开展了诸如 PISA 等大规模测试来关注学生个体的全面发展。国内的质量监测评价研究倾向于构建中国特色的教育质量评价理论体系，从循证角度提供符合我国文化情境的依据，围绕核心素养与学科教学的融入，学校的教学变革将超越传统的班级授课和教学过程，课程评价也会超越知识本位和学科结构主义，教育质量的监测评价也将随之变革调整，这都将成为未来教育的主要特征。

（三）数字化技术的运用推进人机动态交互式智慧评价的发展

教育质量监测评价理论和统计测量理论的新进展、监测评价手段和工具的丰富改变了传统测验方式对“小数据”“事后数据”“回忆数据”的收取，量化方法与质性方法的融合运用，可以实现数据处理的规范化和多元化。智能技术的应用创设模拟真实、交互式问题情境，让教师和学生的教学、学习过程痕迹能够大量留存，智慧评价手段运用也愈发成熟。

（四）元评估标准的建立将加强评估的导向、诊断与激励作用

“教育评价的再评价”“谁来监测与评估教育政策”是教育政策开展的前提条件，这一研究正在引起学界和社会的重视。随着社会、学界和政府层面从管理到治理的认识潮流的兴起，教育治理、教育社会组织、第三方组织等概念日渐渗入教育研究领域，倡导社会参与教育政策监测与评估过程的呼声也从无到有。[37]如何保障多元主体参与到教育质量监测评价领域，切实发挥监测评价的价

中国知网 https://www.cnki.net

值和功能是新时代教育质量监测评价研究所面临的重大问题。元评估方法应势而生，任何教育评价的作用都是有限的，元评估作用在于知道评估自身的局限性，对于评价结果给予严格的界定。当前，世界范围内的元评估理论研究还比较薄弱，尚未建立起与现有监测评价体系相衔接的元评估制度，逐步开展教育质量监测评价的元评估理论和实践研究，丰富和完善我国教育质量监测评价的理论体系，是新时代教育质量监测评价成熟体系建设的客观要求。

五、研究启示

在基础教育普及面增大、高等教育大众化格局的背景下，各国教育质量监测评价聚焦发展内涵式高质量教育，形成技术赋能教育监测评价的新范式，这对我国开展教育质量监测评价具有一定的借鉴意义。

第一，构建具有中国情境和特色的教育质量监测评价的理论和实践研究。我国正处于百年未有之大变局情境下，中国教育具有特殊情境，需结合中国的区域分布、民族特色、制度等因素界定教育质量监测的标准，协同第三方评价社会组织，建立基于教育强国建设背景下教育评价改革研究框架。例如，在监测评价过程中，关注学生的社会情感、文化适应、正面反馈等重要变量，探究这些变量之间的关系和作用路径。

第二，实现纵向数据的教育监测评估，采用动态性视角开展增值性评价的追踪研究。信息技术赋能教育质量监测评价让评价的手段、方法更加多元化，例如，收集学生从入学到毕业全过程的学习成绩、行为表现、心理状态等多维度纵向数据，构建全面且精准的学生成长数据库。清晰呈现学生个体在知识掌握、能力提升以及素养发展等方面的动态变化轨迹。采用动态追踪研究，探讨教育质量监测评价改进机制，研究教育质量监测评价的差异性和动态性，能够丰富中国教育质量监测评价的翔实观察样本。

第三，丰富教育质量监测评价研究的视角和方法，融入高水平的算法和分析技术，贯通人工智能支持下个性化学习、课堂诊断与系统层面教育大数据的应用，提升数据决策意识，同时也要避免数据陷阱，将有效的预测、决策、干预和人的经验与智慧协同。

参考文献：

[1]Baker, F. Quality Assurance and Program Evaluation [J]. Evaluation & the Health Professions, 1983, 6(2): 149-160.

[2]陈玉琨，李如海.我国教育评价发展的世纪回顾与未来展望[J].华东师范大学学报(教育科学版)，2000(1)：1-12.

[3]金娣，王刚.教育评价与测量[M].北京：教育科学出版社，2007：44-48.

[4]段鹏阳，杜玲玲.我国区域教育发展水平评价研究的回顾与展望(1988—2021)——基于CiteSpace的可视化分析[J].教育测量与评价，2021(11)：10-18.

[5]檀慧玲，王发明.教育政策评估研究的关键词共现可视化分析[J].华南师范大学学报(社会科学版)，2017(4)：65-73+190.

[6]郭元祥，王秋妮.增值评价研究的知识图谱与前景展望[J].教育测量与评价，2021(7)：3-10.

[7]翟思卿.近十五年来我国教育评价研究的演进分析[D].郑州：河南大学，2014.

[8]邱均平，欧玉芳.美国教育评价研究的知识基础与热点——基于2000—2014年8种SSCI期刊的可视化分析[J].中国地质大学学报(社会科学版)，2016，16(2)：142-149.

[9]陈超美，陈悦，侯剑华，等. CiteSpaceII：科学文献中新趋势与新动态的识别与可视化[J].情报学报，2009，28(3)：401-421.

[10]李杰，陈超美.CiteSpace科技文本挖掘及可视化[M].北京：首都经济贸易大学出版社，2016：84-108，194-206.

[11]黄涛，王一岩，张浩，等.数据驱动的区域教育质量分析模型与实现路径[J].中国电化教育，2019(8)：30-36.

[12]戚业国，杜瑛.教育价值的多元与教育评价范式的转变[J].华东师范大学学报(教育科学版)，2011，29(2)：11-18.

[13]伍远岳.教育质量评价的异质性探析——基于PISA2015和PISA2018的数据分析[J].教育研究与实验，2021(6)：49-56.

[14]Pachler N, Daly C, Mor Y, Mellar H. Formative e-assessment: Practitioner cases [J]. Computers & Education, 2010, 54(3): 715-721.

[15]Ballou D, Sanders W, Wright P. Controlling for Student Background in Value-Added Assessment of Teachers[J]. Journal of Educational and Behavioral Statistics, 2004, 29(1): 37-65.

[16]Stronge J H, Ward T J, Grant L W. What Makes Good Teachers Good? A Cross-Case Analysis of the Connection Between Teacher Effectiveness and Student Achievement[J]. Journal of Teacher Education, 2011, 62(4): 339-355.

[17]Noell G H, Burns J L. Value-Added Assessment of Teacher Preparation[J]. Journal of Teacher Education, 2006, 57(1): 37-50.

[18]Spady W G. Outcome-Based Education: Critical Issues and Answers[M]. American Association of School Administrators, Arlington, 1994.

[19]Dincer M A, Uysal G.The determinants of student achievement in Turkey[J].International Journal of Educational Development, 2010, 30(6): 592-598.

[20]Jehangir K, Glas C, Stephanie V. Exploring the relation between socio-economic status and reading achievement in PISA 2009 through an intercepts-andslopes-as-outcomes paradigm [J]. International Journal of EducationalResearch, 2015(71): 1-15.

中国知网 https://www.cnki.net

[21]Cortez P. Silva A. Using Data Mining to Predict Secondary School Student Performance, Proceedings of 5th Annual Future Business Technology Conference[C].Porto:EUROSIS,2008:5-12.

[22]王蕾,景安磊,佟威.PISA 中国独立研究实践对构建中国特色教育质量评价体系的启示——基于 PISA2009 中国独立研究[J].教育研究,2017,38(1):114-123.

[23]钟秉林,王新凤.普及化阶段我国高校教学质量评价范式的转变[J].中国大学教学,2019(9):80-85.

[24] 孙成梦雪, 杜瑞军. 高等教育质量评价改革与发展新活力——中国高等教育学会教育评估分会 2020 年学术年会综述[J].高教发展与评估,2021,37(5):55-62.

[25]Lopez-FernandezO, Rodriguez-IlleraJL. Investigating university students´ adaptation to a digital learner course portfolio[J]. Computers & Education,2009,52(3):608-616.

[26]Segers M, Dochy F. New Assessment Forms in Problem-based Learning: The value-added of the students´ perspective[J]. Studies in Higher Education,2001,26(3):327-343.

[27]Marland J, HarrickM, SireciSG. Student Assessment Opt Out and the Impact on Value-Added Measures of Teacher Quality [J]. Educational and Psychological Measurement,2019,80(2):1-24.

[28]Gómez M, Aranda E, Santos J. A competency model for higher education: an assessment based on placements[J]. Studies in Higher Education,2016,42(12):2195-2215.

[29]Hattie J, Timperley H . The Power of Feedback[J]. Review of Educational Research,2007,77(1):81-112.

[30]Topping K J. Peer Assessment[J]. Theory into Practice, 2009,48(1):20-27.

[31]Griffin P , Care E . Assessment and Teaching of 21st Century Skills: Methods and Approach[M]. Dordrech: Springer, 2015.

[32]贾林芝."教—评一体化"视域下教师校内学业评价素养现状调研[J].教育发展研究,2020,40(20):53-61.

[33]杨清.学校课堂教学评价:价值的判断、挖掘与提升[J].教育科学研究,2021(11):61-65+71.

[34]叶澜,吴亚萍.改革课堂教学与课堂教学评价改革——"新基础教育"课堂教学改革的理论与实践探索之三 [J]. 教育研究, 2003(8):42-49.

[42]烷留一,叶海智,张瑾,等.基于多维数据的课堂教学行为评价研究[J].现代教育技术,2022,32(4):68-77.

[36]陈悦,陈超美,刘则渊,等.CiteSpace 知识图谱的方法论功能[J].科学学研究,2015,33(2):242-253.

[37]胡伶,范国睿.教育政策监测与评估主体的现状与发展建议——基于权力来源与向度的分析框架[J].教育发展研究, 2012,32(Z1):8-13.

(责任编辑 金子运)

Research on Progress, Hot Topics and Trend of Education Quality Monitoring and Evaluation ——A Bibliometric Analysis of International and Domestic Literature from Core Database

Zhang Jia Xie Dan Xie Ze-hui

Abstract: Using bibliometric analysis, this study systematically combs and visually analyzes research in the field of education quality monitoring and evaluation from CSSCI source journals in CNKI and SSCI/A&HCI source journals in the Web of Science database. The research shows that studies in this field have been on an upward trend, forming a research network centered on domestic and international university institutions such as the University of Maryland, the University of Mississippi, Stanford University, and Beijing Normal University. Research hot topics focus on concepts of monitoring and evaluation quality, homogeneity and heterogeneity of monitoring and evaluation, large-scale and longitudinal monitoring and evaluation, evaluation effectiveness, etc., extending to frontiers such as human-machine dynamic interaction evaluation, reflecting the contemporary need to build an artificial intelligence-driven education monitoring system. China should construct theoretical and practical research on education quality monitoring and evaluation with Chinese context and characteristics, establish a tracking research system using longitudinal data, and integrate human experience and wisdom to predict, judge, and intervene in data.

Key words: Education Quality; Education Monitoring; Education Evaluation; Bibliometric Analysis

中国知网 https://www.cnki.net