

全国教育科学“十二五”规划教育部重点课题成果

基础教育质量综合评价 理论与实践研究

王建华 卢鸿鸣 缪雅琴
—— 著

CIS

湖南教育出版社



序

教育，国之大计，党之大计。在 2018 年全国教育大会上，习近平总书记提出，要“扭转不科学的教育评价导向，坚决克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾”，“从根本上解决教育评价指挥棒问题”。历史和现实证明，教育评价对教育发展具有重要的引导、推进功能，是教育综合改革的关键环节。2019 年全国教育工作会议提出“把教育评价改革作为最硬的一仗”，评价制度改革作为教育综合改革的突破口，在深化教育体制改革、健全立德树人落实机制中具有战略性地位，这无疑对教育评价改革提出了新的要求。我国教育评价正处在发展和转型的关键时期，如何从以往比较单一、局限的评价，发展为更加多样化、专业化的评价？如何实现评价的科学转型，以求更大限度地发挥其多种功能？这不仅是教育评价领域，也是全面深化教育改革过程中，需要从理论和实践层面进行探索的现实课题。长沙市作为全国教育综合评价改革实验区，坚持科研引领，运用增值评价理念，构建测评体系，自主开发测评工具，挖掘已有教育大数据，有序推进评价改革，打破了教育评价导向的片面性和评价机制的局限性，创新性地形成了综合考量学生、教师和学校发展的中小学教育质量综合评价改革“长沙模式”。其申报的课题“基础教育质量发展性评价体系构建的理论与实践研究——以湖南省长沙市为例”获批为教育部

2013 年度重点资助课题。

长沙市紧扣教育部教育质量综合评价指标框架，结合本地实际设立了立体多维的评价指标体系。在评价内容上既关注学业水平状况，又注重学生的综合素质和个性发展；在评价方式上既关注学生学习结果的体现，又注重学生的学习过程和心理品质；在评价主体上既关注学生个体的成长，又注重教师的师德、专业素养以及学校的办学管理水平。全面评价学校办学水平，形成涵盖学生、教师和学校发展的“体检表”，引导全社会转变质量观和人才观，为学校办学营造了良好的育人环境。

教学质量是社会评价学校办学的重要指标，如果仅关注学校最终的升学率，就难以全面客观反映学校办学水平的进步幅度。为改变单纯强调结果而忽视进步程度的倾向，充分发挥评价的诊断、改进、激励等功能，长沙市采用增值评价，“从入口看出口、从起点看变化”，重视学生接受学校教育后的进步程度，关注学生成长状况，客观评价学校办学水平，激励学生和学校尽最大努力，做最好的自己。可以肯定地说，增值评价以发展的、全面的视角来评价学生与学校，更合理、更公平、更科学，不仅为学生的终身发展奠定基础，同时对推动中小学提高教育教学质量、办出特色、走可持续发展的道路具有积极的意义。

“工欲善其事，必先利其器”，实施增值评价需要大数据支撑，开发整套有信度和效度的测评工具显得非常重要。长沙市组织专家团队自主研制系列学业质量试卷和影响学业质量相关因素测评问卷，采集七年级新生分班考试、中考、学考、高考等人民群众非常关注的学业数据以及问卷调查、国家学生体质健康监测、综合素质评价等学业相关因素数据，同时以“人人通”为载体搭建教育质量综合评价管理平台，将在线学习中心、综合素质评价、体质健康监测等系统互联互通，打破信息孤岛，实现数据融合。采集客观真实、类型丰富的全市全样本数据，为实现评价改革提供了保障。

生逢其时，重任在肩。教育评价的目标要从狭窄转向综合，教育评价的方法要从陈旧转向创新，教育评价的主体要从单一转向多元，评价结果的呈现方式要更加丰富、更加有效，需要教育改革者积极实践、开拓创新。长沙市教育改革者紧紧抓

住评价的“牛鼻子”，不断探索教育质量综合评价方法、手段和技术，尤其是增值评价技术从采用 K 均值聚类法到采用多水平分析，立足实际不断优化增值评价技术的使用，为其他地区教育质量综合评价改革提供了很好的借鉴。

这本著作凝聚了长沙市教育改革者的集体智慧，是遵循教育规律、人才成长规律，大力推进教育体制改革的有效创新。全书全面总结了五年来教育质量综合评价改革成果，对长沙市教育质量综合评价改革的背景、理念、操作和结果运用进行了详细阐述，不仅有理论探讨，更有实践操作，对一线教育工作者很有借鉴意义。

辛 涛

教育部基础教育质量监测中心副主任

北京师范大学教授、博士生导师

目 录

第一章

基础教育质量综合评价概述—— 001

第一节 基础教育质量综合评价背景分析—— 003

第二节 基础教育质量综合评价的目的与意义—— 017

第三节 基础教育质量综合评价的研究路径—— 022

第二章

基础教育质量发展性评价理论及运用—— 027

第一节 增值评价的内涵、特点与功能—— 029

第二节 常用的几种增值评价技术—— 033

第三节 国内外增值评价实践—— 038

第四节 基础教育质量增值评价探索与实践—— 043

第三章

基础教育质量综合评价体系的构建—— 051

第一节 基础教育质量综合评价的价值取向—— 053

第二节 基础教育质量综合评价的指标体系—— 055

第三节 基础教育质量综合评价的方法—— 061

第四节 基础教育质量综合评价的机制保障—— 071

第四章

基础教育质量综合评价工具的研发与施测———— 077

第一节 基础教育质量学业水平测验的研发———— 079

第二节 学业相关因素测评问卷的研发———— 094

第三节 基础教育质量综合评价测评工具的施测及数据处理———— 119

第五章

基础教育质量综合评价报告的撰写与结果应用———— 135

第一节 基础教育质量综合评价报告的类型、特点与基本结构———— 137

第二节 基础教育质量综合评价的结果应用———— 161

第三节 基础教育质量综合评价报告应用案例———— 165

第六章

基础教育质量综合评价的成效与展望———— 233

第一节 基础教育质量综合评价的成效———— 235

第二节 基础教育质量综合评价的展望———— 239

附录

2019 年长沙市初中学业质量评价标准 语文———— 243

2019 年长沙市初中学业质量评价标准 数学———— 272

2019 年长沙市初中学业质量评价标准 英语———— 296

2019 年长沙市初中学业质量评价标准 理科综合（物理科）———— 323

2019 年长沙市初中学业质量评价标准 理科综合（化学科）———— 338

2019 年长沙市初中学业质量评价标准	文科综合（道德与法治科）	————— 353
2019 年长沙市初中学业质量评价标准	文科综合（历史科）	————— 368
2019 年长沙市初中学业质量评价标准	地理	————— 397
2019 年长沙市初中学业质量评价标准	生物	————— 420

第一章

基础教育质量综合评价概述

第一节	基础教育质量综合评价背景分析———	003
第二节	基础教育质量综合评价的目的与意义———	017
第三节	基础教育质量综合评价的研究路径———	022

当前我国教育改革已经进入深水区，教育质量评价是教育综合改革的关键环节，推进中小学教育质量综合评价改革，是推动中小学全面贯彻党的教育方针、全面实施素质教育、落实立德树人根本任务的重要举措，是引导社会和家长树立科学的教育质量观、营造良好育人环境的迫切需要，是基本实现教育现代化、加强和改进教育宏观管理的必然要求。

第一节 基础教育质量综合评价背景分析

当前,科学技术以空前的发展速度和影响力左右着人类前进的步伐,世界经济和社会的进步越来越依靠知识、技术和能力,人力资源已经成为各国在全球市场竞争中的决定因素。基础教育作为所有儿童、少年进入社会前必须经历的学习阶段,决定着一个国家未来公民的生命质量与价值,也决定着未来公民的品格与行为高度。党的十九大报告提出,优先发展教育事业、加快教育现代化、建设教育强国,要求办更加公平更加优质的基础教育。基础教育是国民教育体系的根基,当前各国越来越重视基础教育质量,纷纷从国家战略的角度出发建立完备的基础教育质量监测与评价体系,作为提高基础教育质量的重要举措。

一、基础教育质量评价的理论研究

提高教育质量是办人民满意教育的内在需求。科学评估基础教育质量,需要对教育质量、教育质量评价、评价标准等概念内涵进行分析,形成基本共识。

(一) 教育质量

对于教育质量,许多学者都有各自精辟的见解。德国著名教育家赫尔巴特认为,“道德”在教育质量中占有重要的地位。认知心理学家则认为,知识是教育质量中的重要组成部分。用学者菲佛和库德的话来说,“‘质量’这个词意义多变,对不同的人而言指的是不同的事,是个必须小心处理的概念”^①。“质量”一词最初来源于管理学,主要反映产品或工作的优劣程度。而在教育中,教育质量反映教育水平的高

^① Edward Sallis. 全面质量教育 [M]. 何瑞薇,译. 上海:华东师范大学出版社,2005.

低和教育效果的优劣，其内涵具有多层次性，可以从多个角度予以评判：从教育主题是育人的角度来看，教育质量体现为学生个体的综合素养；从教育的具体载体是知识与技能传授的角度来看，教育质量则体现为教育活动所包含的知识深度、广度和强度必须达到国家规定要求的层次、类型和规格；从教育客体是服务社会的角度来看，教育质量则体现为满足社会需求的情况，即教育产品的市场检验合格率和进入社会后的持续竞争力。^①《教育大辞典》从宏观和微观层面对教育质量进行了解释。宏观教育质量又称“体系质量”，指整个教育体系的质量，实质是指其与系统规模、结构和效益等之间的协调状况。也就是说，体系质量以系统内部各要素之间是否协调一致为衡量标准，当系统内部各要素之间协调一致时，就表现出较高的体系质量。从微观层面看，教育质量体现在培养对象的质量上，是指教育水平高低和效果优劣的程度，衡量标准是教育目的和各级各类学校的培养目标。前者规定受培养者的一般质量要求，亦是教育的根本质量要求；后者规定受培养者的具体质量要求，是衡量人才是否合格的质量规格。^②

虽然不同教育价值观下对教育质量内涵的理解有所不同，但都不能否认教育是培养人的活动，教育活动的核心产出是学生。从事基础教育的每所学校、每位老师开展的工作都是为了确保以学生为主体的教育教学活动正常进行，因此，基础教育质量最终体现于所培养的学生的质量。我国教育方针对基础教育所培养的学生质量有明确要求，对如何在教育方针的指导下实施素质教育也提出了一些基本要求。鉴于此，界定基础教育质量的内涵，需把握以下几个要点：对教育提供者来说，教育质量应指教育的成效或结果达到国家教育目标所规定标准的程度与水平；对受教育者来说，教育质量应指学生在知识与技能、过程与方法、情感态度与价值观等方面的发展状况，达到教学大纲、课程标准规定的程度与水平；对教育评价者来说，教育质量应指对全体学生全面发展的综合测量与评价，以及对由每一个学生个体汇聚起来的每一所学校学生群体状况的衡量与比较。这种综合测评，需对照实施素质教

① 许世红，黄小平，王家美. 基础教育质量监测研究 [M]. 广州：广东高等教育出版社，2016.

② 顾明远. 教育大辞典（增订合编本）[M]. 上海：上海教育出版社，1998.

育的规定与全面发展的要求，应该包括学生的品德行为、学业水平、身心健康、艺术素养和社会实践等多方面的水平。

当前，仍有不少人存在对教育质量内涵理解上的误区，例如将基础教育质量等同于生源质量、等同于教学质量、等同于结果质量等。^① 这些对教育质量的不准确理解使得很多地方出现生源竞争现象、“揠苗助长”现象，摒弃了教育的价值引导，扭曲了教育的本质，既影响了学生的全面而有个性的发展，又不利于教育的可持续发展。

（二）教育质量评价

《国家中长期教育改革和发展规划纲要（2010—2020年）》明确提出，“把育人为本作为教育工作的根本要求”，“把提高质量作为教育改革发展的核心任务”；党的十八大报告提出，“全面实施素质教育，深化教育领域综合改革，着力提高教育质量，培养学生社会责任感、创新精神、实践能力”；党的十九大报告进一步要求“努力让每个孩子都能享有公平而有质量的教育”。开展基础教育质量评价，对科学评估基础教育质量总体水平、摸清中小学生发展现状、客观反映影响基础教育质量相关因素基本状况、改进基础教育学校教育教学、提升基础教育质量等具有重要意义，目前已成为我国国家层面的一项重大战略任务，是教育治理体系与治理能力现代化的重要内容，是引领教育改革和发展的重要手段，是教育综合改革的关键环节。

基础教育质量评价具有多重功能，主要可以概括为五个方面：诊、咨、督、促、导。“诊”，即通过评价，对学生各个阶段的德智体美劳发展情况进行全面“体检”，科学诊断存在的问题，揭示问题背后的深层次原因；“咨”，即基于评价数据，为国家 and 地方教育决策的制定提供科学的咨询建议，推动教育管理从“粗放式”向“精细化”转变，推动教育决策从“经验型”向“科学化”转型；“督”，即通过评价，对学校和地方教育行政部门办学行为进行针对性督查，适时公布评价结果，督促整改不良办学行为；“促”，即通过评价，有效促进教师的教学行为改善，促进教育质

^① 柳文华. 中小学教育质量评价研究 [J]. 教学与管理, 2018 (30): 118—121.

量整体提升;“导”,即通过评价,引导社会、学校、家长树立科学的教育质量观。通过发挥诊、咨、督、促、导多重功能,让教育质量评价成为“促进学生全面发展”的教育“体检仪”和新的“指挥棒”。^①

世界各国越来越重视基础教育质量评价工作,美国、英国、德国、日本等发达国家纷纷成立专门的教育质量测评机构,建立本国完备的基础教育质量评价体系,即使像越南、柬埔寨等发展中国家也在建立自己的基础教育质量监测评价体系。^②此外,越来越多的国家参与到国际学生评估项目(Programme for International Student Assessment, PISA)、国际数学和科学评测趋势研究(the Trends in International Mathematics and Science Study, TIMSS)等国际跨国教育质量测评项目。

(三) 教育质量评价标准

教育质量标准的选择与质量的判断对各级各类教育机构的发展具有导向性、基础性和全局性的影响与作用。^③“世界多国教育发展到较高水平的重要特征之一即是制定教育质量国家标准,建立健全教育质量保障体系。”^④很多学者从宏观角度对教育质量评价标准进行了探讨。学者王彦明将基础教育的质量标准概括为全面基础、底线合格和差异评价。他认为“为美好人生奠定基础是基础教育质量的核心内核,高质量的基础教育要能够为每个学生的美好未来奠定全面、扎实和牢固的基础”^⑤;“基础教育的底线是培养具有健康人格、良好素质的合格公民……而基础教育的顶线是培养适合高等职业教育和高等专业教育选拔规格的人才”^⑥;同时强调不同地区之间、城市之间因当地的社会、文化和经济背景等因素的不同,应采用不同的评价标准进行差异评价。

学业质量水平是衡量教育质量高低的重要指标之一,构建基础教育学业质量评价标准是当前研究的热点。学者原克学、郝华杰认为,“学业质量评价标准研制需厘

①② 林崇德. 对未来基础教育的几点思考 [J]. 课程·教材·教法, 2016, 36 (3): 3—10.

③⑤ 王彦明. 基础教育质量之感 [J]. 教育理论与实践, 2013, 33 (16): 53—56.

④ 中国教科院教育质量标准研究课题组. 教育质量国家标准及其制定 [J]. 教育研究, 2013 (6).

⑥ 杨启亮. 合格性评价: 基础教育评价的应然选择 [J]. 教育研究, 2006, (11): 11—17.

清学生发展核心素养、国家课程标准、教育质量综合评价的指标与学业质量评价标准的关系”。^① 学者乐毅认为,“学科教育质量评价标准需要跨越不同学科领域,对基础教育阶段所预期的教育结果有一个整体构思,在基础教育阶段结束时,明确每个学生作为未来社会的国家公民所应具有的核心素养”。^② 学者曾沛权强调,“义务教育阶段学业教育质量评价标准是课堂教学与学科教育质量监控间的支点,是教师在日常教学中进行自我监控的教学建议,是学科教育质量监控中测试命题的依据”。^③ 国内已有多个地区开展学业质量评价标准研制,例如:2009年11月,广州市发布《义务教育阶段学科学业质量评价标准》;2011年2月,上海市闵行区启动《学科学业质量标准》制订工作;2011年浙江省启动义务教育阶段学科学业质量标准编制工作;2011年11月,武汉市出台《义务教育学生学业质量监测评价标准》;2018年3月,长沙市发布《2018长沙市初中学业质量评价标准》。^④

(四) 教育质量评价指标

基础教育质量评价指标体系是对“评价什么”的具体描述和说明,是开展教育质量测评的基础和前提条件,也是测评工具开发的依据。联合国教科文组织在其每年的年度出版物《全球教育统计摘要》中用七大类指标来衡量学校教育质量,即教育机会、教育效率和成果、教师资源、教育经费、教育结构、教育公平、教育国际化;经济合作与发展组织在其年度报告《教育概览 2013: OECD 指标》中用四大类指标来评价教育质量,即教育机构的产出及学习的影响,教育中的财力与人力资源投入,教育机会、参与与过渡,学习环境与学校的组织。^⑤

① 原克学,郝华杰.对研制义务教育阶段学业质量评价标准的思考[J].教育理论与实践,2017,37(29):22-24.

② 乐毅.试论制定国家学业质量标准的若干基本问题[J].教育研究,2014(8):40.

③ 曾沛权.浅析《广州市义务教育阶段学科学业质量评价标准》在课堂教学和测试命题中的导向[J].新课程(小学),2015(10):94-95.

④ 孙智明,刘正华.基于学生发展核心素养的学业质量评价标准编制与应用——以《2018长沙市初中学业质量评价标准》为例[J].当代教育论坛,2018(4):26-32.

⑤ 刘晓玫,刘瑶,宋庆莉.区域内学校教育质量评价指标体系构建[J].教育理论与实践,2016,36(17):10-12.

近年来,诸多学者对宏观层面的教育公平、教育发展等问题以及微观层面的学校教育质量及其评价等问题进行了探讨。如刘晓玫、刘瑶、宋庆莉通过德尔菲法构建区域学校教育质量评价指标体系,通过层次分析法确定评价指标权重,将定性分析与定量分析相结合。^①董博清、霍素君构建了涵盖学习态度、学习兴趣、学习习惯、学习方法、对学习目的与意义的认识等方面的“学习素养”指标,来评价学生的学习品质,补充和完善教育质量评价的指标内容。^②陈惠英构建了包括学校工作质量、学生素质发展水平两方面共11项二级指标、30项三级指标的初中学校教育质量自我监控与评价指标体系。^③2016年9月,由教育部委托北京师范大学历时3年完成的《中国学生发展核心素养》研究成果发布。核心素养以培养“全面发展的人”为核心,分为文化基础、自主发展、社会参与三个方面,综合表现为人文底蕴、科学精神、学会学习、健康生活、责任担当、实践创新六大核心素养,具体细化为国家认同等18个基本要点。“核心素养不是一个理论性的抽象框架,而是要根据学生发展的连续性和阶段性特点,把核心素养细化为不同教育阶段的培养目标。”^④构建具有操作性的基于核心素养框架的教育质量测评指标体系也是当前的研究热点。学者张莹、冯虹构建了基于核心素养的教育质量评价指标体系,“评价指标体系由学科素养、生存素养和精神素养三个领域的学业成绩、知识应用、工具使用、人际交往、健康审美情趣以及情感态度与价值观六个维度构成,以希望建立适应不同年龄阶段的‘一通多专’的评价系统”^⑤。学者孙智明、刘正华从学科核心素养出发,建构了语文、数学、英语、道德与法治、历史、物理、化学、生物、地理等9个学科的评价内容,分层分类考察学生的综合运用所学知识分析和解决实际问题的

① 刘晓玫,刘瑶,宋庆莉.区域内学校教育质量评价指标体系构建[J].教育理论与实践,2016,36(17):10-12.

② 董博清,霍素君.学生学业发展水平评价体系的研究与思考——以河北省义务教育评价为例[J].河北师范大学学报(教育科学版),2018,20(4):123-128.

③ 陈惠英.初中学校教育质量自我监控与评价指标体系的建立与思考[J].教育测量与评论(理论版),2011,(7).

④ 褚宏启,张咏梅,田一.我国学生的核心素养及其培育[J].中小学管理,2015(9).

⑤ 张莹,冯虹.基于核心素养的教育质量评价指标体系的构建与应用[J].教育探索,2016(7):60-64.

能力。^①

2013年,教育部出台《中小学教育质量综合评价指标框架(试行)》,把学生的品德发展水平、学业发展水平、身心发展水平、兴趣特长养成、学业负担状况等方面作为评价学校教育质量的主要内容,要求各地要在涵盖以上5个方面评价内容的基础上,对照20项关键性指标,按照小学、初中和普通高中教育的不同性质和特点,细化评价指标、考查要点和评价标准的内容要求,完善综合评价指标框架。随后,上海市构建了中小学学业质量综合评价“绿色指标”,包括学生学业水平指数、学生学习动力指数、学生学业负担指数、师生关系指数、教师教学方式指数、校长课程领导力指数、学生社会经济背景与学业成绩相关指数、学生品德行为指数、学生体质健康指数和跨年度进步指数10项指标。广州市构建了中小学教育质量阳光评价指标体系,包括学生的品德与社会化水平、学业发展水平、身心发展水平、学习潜能、学业负担状况、对学校的认同六个方面共22项指标。长沙市构建了中小学教育质量综合评价指标体系,涵盖学生成长、教师发展、学校管理三方面共11项一级指标、45项二级指标。

当前,对基础教育质量监测与评价的理论研究除了集中在以上几个方面,还有学者对教育质量测评工具的开发、测评数据的分析及结果运用进行了探讨。例如,学者秦建平、张惠、陈飞鹏提出中小学生公民素养测评维度结构假设,运用结构方程模型对假设进行检验和矫正,开发中小学公民素养测验工具。^②学者陈荣荣针对目前教育质量测评数据挖掘和应用中存在“数据质量不高,缺乏系统化、结构化的数据库;数据挖掘方法单一,尚停留在数据描述分析阶段;数据挖掘与应用意识淡薄,主要停留在结果公布阶段”等问题,明晰数据挖掘的五个层次,构建基于数据驱动的教育教学改进的行动路径,即“基于数据寻找‘真问题’、拓宽思路开展‘真研究’、综合判断实施‘真改进’”。^③

① 孙智明,刘正华.基于学生发展核心素养的学业质量评价标准编制与应用——以《2018长沙市初中学业质量评价标准》为例[J].当代教育论坛,2018(4):26—32.

② 秦建平,张惠,陈飞鹏.中小学生公民素养测评工具研究[J].上海教育科研,2018(7):5—10.

③ 陈荣荣.教育质量评价数据的挖掘思路与应用路径[J].上海教育科研,2017(6):54—60.