

基于国内外比较的省级教育质量监督 定位与发展趋势

马海燕

摘要：分析国际质量监督项目和国家质量监督项目的项目定位和实践重点，提出省域教育质量监督的定位与发展趋势：在监测功能上更突出学校中心，突出教学服务；在监测内容上更突出国测未测，实施专题监测；在监测过程上更体现数字导向，实现方式创新；在监测对象上更体现多元，实现领域创新；在监测结果上更关注社会导向，体现协同育人。

关键词：省级教育质量监督；国内外比较；定位；发展趋势

2022 年 4 月颁布的《义务教育课程方案》指出要开展国家、省两级课程实施监测，要求省级教育行政部门委托有关专业机构实施省级监测。比较国际，已有 PISA、TALIS、TIMSS 和 NAEP 等多个成熟项目，且都有清晰的功能定位。放眼国内，教育部中小学教育质量监督也对义务教育阶段课程实施情况开展了多年监测。作为省级教育质量监督机构，在国际监测项目和国家监测项目中该如何找到自己的定位，体现省级监测的功能价值？笔者通过对国际监测项目的比较和国家监测项目的分析，提出对省级监测项目的定位和发展趋势的思考。

一、基础教育质量监督的国际视野

基础教育作为人才培养的教育阶段，决定着一个国家未来的全民素质，基础教育质量已成为影响国家持续发展和提高国家竞争力的核心要素。因此，基础教育质量监督日益受到世界各国的高度重视，也逐渐成为各国教育研究者关注的焦点。据不完全统计，全世界迄今已有 148 个国家（地区）开展国家（地区）水平代表性基础教育质量监督。其中著名的跨国性大规模教育质量监督项目有经济合作与发展组织（OECD）组织的国际学生评价项目（Program of International Student

Assessment，简称 PISA）和教师教学国际调查（Teaching and Learning International Survey，简称 TALIS），国际教育成就评价协会（IEA）组织的国际数学和科学教育趋势项目（Trends in International Mathematics and Science Study，简称 TIMSS），国际阅读素养进展项目（Progress in International Reading Literacy Study，简称 PIRLS）。另外还有国家性的教育质量监督项目，如美国的国家教育进展评估项目（National Assessment of Educational Progress，简称 NAEP）。自 2007 年开始，我国也开始探索基础教育质量监督。这些项目在监测定位、监测内容、监测技术以及对监测结果的应用方面，都成为省域教育质量监督学习、效仿的榜样。

（一）PISA 项目：教育公平与质量的政策导向

PISA 项目是 OECD 经过多年研发、旨在测试各个国家或地区义务教育末期学生的相关“素养”，即测试学生们是否掌握参与社会所需要的知识与技能的项目。2000 年开展首次评估，每三年评估一次，重点评估完成基础教育的 15 岁学生在阅读、数学和科学领域的关键能力以及综合素养，如学生将所学知识和技能应用于日常沟通、理解、分析和推理，以及解决和解释各种不同情

作者简介：马海燕，女，浙江省教育厅教研室，高级教师。

境下复杂问题的能力。但 PISA 项目不是一种高利害性的考试，不评价学生学科知识的掌握情况，而是了解他们面向未来生活相关的基础知识和技能的掌握程度。其次 PISA 是基于对质量与公平的诉求做的监测。OECD 和各国（地区）发布以质量与公平为导向的研究报告，最终影响到各国（地区）教育政策的制定和改进。因此，PISA 是基于教育中对质量与公平的考量，利用发布报告的显性方式，为各国教育制定或调整教育政策提供参考和依据，以达到促进各国教育发展、提升教育质量的目的。

（二）TALIS 项目：重视专业发展的价值导向

教师教学国际调查（TALIS）是 OECD 于 2007 年策划并首次实施的项目，旨在通过收集和分析可靠、及时和可比的数据，帮助不同国家反思和制定高质量的教育政策，以构建高质量的教师专业队伍。TALIS 项目涉及四个领域：教师专业发展、对教师的评价与反馈、学校领导能力以及教学实践与教学观。TALIS 项目为我们展示了现代教师评价制度与理念，不仅重评价结果，更加重评价过程，强调参与式评价，其教师专业发展的价值导向和评价维度，对改进我国教师评价制度，还教师以专业发展的主体地位具有很大的启示。

（三）TIMSS 项目：关注学科达标的课程导向

TIMSS 项目由国际教育成就评价协会（简称 IEA）于 1995 年发起，每四年组织实施一次。该项目主要监测四、八年级学生的数学和科学成就及其影响因素，重点了解其对数学和科学的基本概念的掌握情况和推理应用情况。与 PISA 着眼未来发展的导向不同，TIMSS 评估的出发点和落脚点是课程，将当下所学课程的掌握程度当作评价的主要内容，即学生在正规的学校教育中被提供了什么课程，学生能够在多大程度上掌握被提供的课程。TIMSS 认为课程是影响学生学业成就最重要

的因素，并首创了课程评价三个层次：理想的课程、实施的课程、获得的课程。

（四）PIRLS 项目：关注阅读进步的过程导向

PIRLS 项目是由 IEA 发起的另一项关于国际阅读素养进步研究的教育质量监测项目，于 2001 年首次实施，每五年一次，监测对象为小学四年级学生。该项目重点监测学生的阅读成绩、阅读行为和阅读态度。PIRLS 非常关注阅读的目的与情境，以此将评价中的儿童阅读活动还原到真实的生活中，关注他们在学校和日常生活中的阅读活动。在 PIRLS 看来，阅读能力已经不再被局限于学科范围内，单纯地作为一种特定的学科能力，而是将阅读与生活相联系，将阅读融入现实生活的各个层面。与 PISA 关注面向未来生活的阅读素养监测不同，PIRLS 关注的是学生在学习的初期阶段所应具备的阅读能力，是被用来学习、创造和发展，使学生充分参与的一种重要能力，是学校各门学科学习的基础。

（五）NAEP 项目：多种功能揉合的发展导向

NAEP 项目是美国从 1969 年开始实施的国家教育进展评估项目，至今已经持续开展了 50 年的时间，是美国国内唯一的、最权威的、长期的中小学生学习成就全国测评体系，已成为衡量美国教育发展状况的一个重要指标，它为美国政府教育政策的制定和改进提供了可靠的依据，为世界各国所效仿。NAEP 的周期是四年，监测科目是数学和阅读，监测对象是 9、13、17 岁的所有适龄青少年包括非在校学生。美国国家教育进步评估包括三个方面，一个是主评估也就是全国性质量评估，一个是长期趋势评估即州评估，还有一个是专项评估即试验性学区评估，关注特定学生群体、进行专题领域等一些研究。迄今为止，NAEP 关注特定学生群体的研究，使用 NAEP 数据进行的专题领域研究，将 NAEP 结果与其他项目结果进行对接的研究都给其他国家带来很多的启示。

二、国家义务教育质量监测的实践影响

相比世界各国教育质量监测,我国的教育质量监测起步较晚,2007年,依托北京师范大学设立教育部基础教育质量监测中心,才开始中国基础教育质量监测的探索。2015年4月15日,教育部正式发布《国家义务教育质量监测方案》,标志着我国正式建立起国家义务教育质量监测制度。2020年10月13日,国务院印发《深化新时代教育评价改革总体方案》,提出构建政府、学校、社会等多元参与的评价体系,建立健全教育督导部门统一负责的教育评估监测机制,发挥专业机构和社会组织作用。这进一步明确了教育质量监测的工作定位。

国家义务教育质量监测,三年一个周期,一年2个学科领域,监测对象为四、八年级学生。它基于课程标准要求,考虑学生认知和学习能力发展的阶段性特征,关注学生对知识、技能的掌握程度、分析解决问题的能力以及相关影响因素。它既重视“基础知识”和“基本技能”,也关注学生终身发展和应对未来挑战所需的“基本素养”和“基本活动经验”。从2007年建立国家级监测中心,到2015年公布监测方案,到2021新一轮五育并举、全面质量的监测方案公布,标志着我国已建立了基础教育质量监测的理论体系,构建了完善的组织实践体系,彰显了诊断、激励、改进的价值体系。

(一) 监测内容更关注全面

相较于世界各国的坚持,我国的质量监测内容最全,关注五育并举、全面发展。在学科上除语文、数学和科学学科质量外,还监测体育、艺术和德育以及英语、心理健康、劳动教育。在影响因素上除了关注学生自身因素、学校管理因素、教师教学因素外还关注家庭成长因素,将学生放在社会、家庭与学校这个大环境中去全面评价,全面诊断。

(二) 监测功能更彰显整合

相较于PISA的政策导向、TIMSS的课程导向,我国的质量监测功能综合考虑了政策导向和课程导向,科学评估全国义务教育总体水平,客观反映义务教育质量相关因素基本状况,系统监测国家课程标准和相关政策规定执行情况。它既强调监测结果为政策所用,成为政策制定和调整的依据,也兼顾监测内容的课程体系,让“教、学、评一致性”更高。

(三) 监测过程更强调追踪

相较于PISA三年一测、TIMSS四年一测、PIRLS五年一测,我国的质量监测周期定为三年一测,既考虑三年内一轮多学科监测,形成义务教育阶段内学校教育的全面质量的评估,也兼顾了三年为一个周期的追踪发展变化的质量监测,如图1。

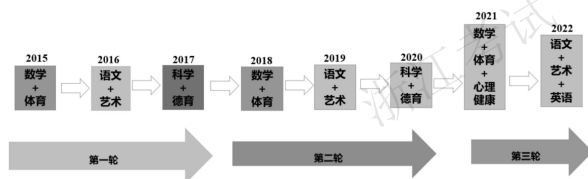


图1 国家三轮监测内容

(四) 监测方式更强调表现

目前世界各国的监测方式基本采用纸笔测验或计算机自适应测验。我国的教育质量监测方式充分考虑了国情的实际情况,采用多种手段融合的监测方式,在学科测试和调查问卷上采用纸笔测验,在体育、艺术监测上采用现场测试,既有体育学科的学生现场演示,也有美术绘画项目参与,更有音乐人机交往的评价。

三、省域教育质量监测的定位与发展趋势

随着国家对教育质量监测的重视,许多省份也加入教育质量监测的行列,比如上海2006年开始实施义务教育质量“绿色评价”,随后北京、重庆、浙江、江苏等都建立了自己的监测体系。审

视各个省域的质量监测，在监测内容上模仿较多、缺少省域定位，监测技术上借鉴较多、缺少自主创新，结果应用上考核较多、较少关注改进，导致省域教育质量监测缺少特点。鉴于国际、国家监测的功能、模式、过程以及结果使用的综合分析，结合省域定位，省级教育质量监测可以从以下几方面发展：

（一）在监测功能上更突出学校中心，突出教学服务

目前国家教育质量监测的功能已经非常清晰，就是要为国家教育决策制定和调整提供依据，为督导评估提供详细证据，最多服务到省。换言之，国家质量监测的学科意识、学校意识并不强，不能很好地为学科诊断、学校诊断提供依据。省测就可以错位发展，突出学校中心，加强学科监测，在监测结果上更注重学科层面、学校层面的分析，注重对学校的分类评价、学科的分层评价，更好地为学科改进服务，为学校发展服务。

（二）在监测内容上更突出国测未测，实施专题监测

国家教育质量监测的学科体系涉及语文、数学、德育，艺术、体育、英语、劳动和心理健康，除此之外，国家对信息技术、综合实践、地方课程等学科没有精力关注，因此省级教育质量监测可以关注这些实践性学科，引导学校开足开齐课程，落实课堂育人质量。其次国家对语文、数学等学科的监测，更多是通过纸笔测验的方式评价其学习结果。省域可以关注专题领域，比如语文的阅读素养、数学的问题解决、科学的实验探究。比如浙江就采用了“基础能力+专题监测”的模式实施省测。

（三）在监测过程上更体现数字导向，实现方式创新

虽然国家监测在信息化应用上做了长足的探

究，但是省域办学条件更均衡，在监测方式上更容易一致。因此省域教育质量监测，可以更注重互联网技术在互监测组织上的应用，多采用上机测试如英语的口语交际、信息技术学科的人机交往，也可以更注重技术在命题上的应用和研究，比如口语、实验的自动等级评价技术。

（四）在监测对象上更体现多元，实现领域创新

就像 OECD 有 TALIS 项目一样，省域更需要教师专业发展的诊断和评价。虽然国家监测中也有对教师教学方式、教师评价理念等信息的收集，但更多是作为影响学生发展的因素存在。目前我国还缺少专门的针对教师专业发展的评估与监测项目。上海引入 TALIS 项目，对上海市内的教师进行专业诊断，较好地促进了教师的专业发展。省域可以鉴于 OECD 和上海的经验，引入教师评价项目，为改进当地教师评价制度、促进教师专业发展做出有益尝试。

（五）在监测结果上更关注社会导向，体现协同育人

国家教育质量监测也很重视对监测结果的育人导向，每年都发布监测结果。但其结论过于简单，过程过于模糊，表达过于雷同，社会影响力虽然很大，但总是意犹未尽。省测可以在结果的报告形式上做更多元表达，比如浙江、苏州就在发布行政报告、县域报告、专题报告的基础上发布社会版报告，通过视频和蓝皮书的形式加大监测结果的宣传，让监测结果的应用更广泛多样。

总之，省域的监测部门，要多借鉴国际、国家教育质量监测模式与经验，立足于自主建构与研发的目标，逐步形成地方特色的教育质量监测模式，通过国际、国家、省三级监测体系的实践探索，进一步彰显五育并举、全面质量的理念，丰富循证改进的方式方法，营造良好的教育生态，促进学生全面而健康发展。