

PIRLS2021：国际阅读素养进展研究 新动向及启示¹

●张所帅 黄志军*

摘要：PIRLS国际阅读素养进展研究评价项目在2021年迎来了第五个评价周期。在保持阅读素养评价理念和框架结构稳定不变的前提下，PIRLS2021呈现出迈向更加数字化的评价方式、测试文本的类型更为多样、背景问卷的内容结构更加优化、采用群体自适应测试题册设计，以及ePIRLS与digitalPIRLS有机整合等新的发展动向。PIRLS2021的新变化给我们带来了要重视数字化阅读素养培养、丰富阅读教学评价文本类型、综合联动构建阅读学习环境、积极探索计算机自适应测试评价等有益启示。

关键词：PIRLS2021 国际阅读素养进展研究 阅读素养评价

中图分类号：G51 文献标识码：A 文章编号：1672-6175（2022）07-0072-09

国际阅读素养进展研究（Progress in International Reading Literacy Study，以下简称“PIRLS”）是由国际教育成就评价协会（International Association for the Evaluation of Educational Achievement，以下简称“IEA”）发起并组织实施的全球

性学生阅读素养跨国（或地区）比较评价项目，与评价数学和科学的TIMSS共同组成了IEA周期性研究的核心内容，定期向参与评价的各国和地区提供四年级学生有关阅读、数学和科学这三门基础学科的学业成就信息。

¹ 本文系2019年海南省哲学社会规划课题“国外大规模阅读素养评价项目PISA、PIRLS、NAEP研究”（项目编号：HNSK<ZC>19-06）阶段性研究成果。

* 张所帅，海南师范大学教师教育学院副教授；黄志军，华东师范大学教育学部副教授。

阅读素养是学业成功的重要基础。PIRLS阅读素养评价于2001年正式启动,每五年进行一次,为全球60多个国家和地区提供四年级学生阅读成就的发展趋势、国际比较及与阅读教育目标和标准有关的学生能力数据,被联合国教科文组织认定为“给监测可持续发展目标进展感兴趣的研究人员、教育工作者和决策者提供了坚实的证据基础”^[1],是监测教育新政或修订后的教育法规是否影响学生学业成就的重要工具,并为各国改善学生学习和教学成效提供了相关信息。

一、PIRLS2021阅读素养评价新动向

PIRLS阅读素养评价基于阅读理论研究成果和阅读素养评价理念,借助教育测量评价技术建构起全面的评价框架,具体包括阅读评价框架、背景问卷框架和阅读评价设计三个方面的内容。自2001年启动以来,IEA与各参与国以及TIMSS&PIRLS国际研究中心等相关单位和机构紧密合作,不断致力于PIRLS阅读素养评价内容和方法的发展与创新,在保持评价理念和框架结构稳定不变的前提下,每一个评价周期都会结合前瞻性的思考和现实中的需要不断调整、更新测试内容和方式,以期能更好地满足评价的目的。继2001年、2006年、2011年和2016年之后,PIRLS2021迎来了第五个评价周期。通过阅读学习IEA在其官方网站发布的PIRLS2021测试评价资料,可以发现五个方面的新变化。

(一) 迈向更加数字化的评价方式

随着信息技术和互联网的迅猛发展,网络生活已经成为人们日常生活的重要组成部分,网络阅读也日益成为学生参与校内外学习生活的重要途径之一。为了与时俱进地反映时代发展的需求,同时也为了更好地了解当下儿童网络阅读和学习信息的状况,PIRLS2021在既有纸质化评价版本的基础上,首次推出基于互联网(Web)的数字化评价版本——digitalPIRLS^[2],基于网络系统,借助电脑以数字化形式呈现PIRLS的阅读文章和题目。

作为21世纪评价阅读素养的先进方式,digitalPIRLS以一种吸引人的视觉形式呈现各种阅读文本,激发学生阅读、与文本互动,以及回答理解问题的积极性,提高作答反应的有效性;利用计算机评价的优势,极大地提高了翻译和翻译验证、数据输入和评分等环节上的操作效率,并且不需要进行文本材料的打印和运输,直接通过网络传送即可;可以与PIRLS在2016年启动的ePIRLS在线信息阅读评价融为一体,实现PIRLS阅读素养评价的全数字化变革,减少之前epirls需要额外一天进行评价的时间;同时,digitalPIRLS还可以通过灵活的方式将PIRLS文章和难度较低的PIRLS Literacy文章进行整合,从而使针对参与国家和地区学生群体成就水平来确定PIRLS2021评价难度的想法成为可能^[3]。

总的来说, 尽管PIRLS2021仍提供两种阅读素养评价版本, 一种是以传统纸质媒介为载体的版本, 一种是以现代数字化媒介为载体的版本, 但PIRLS已开始专注于向数字化评价方式过渡^[4], 其中约有一半的参与国家和地区借助电脑和网络传送PIRLS2021阅读素养评价材料。通过使用精简、易用的用户界面给学生以不同的阅读和作答体验, 充分体现了数字化时代的鲜明特征, 代表了未来阅读素养评价的发展趋势。

（二）测试文本的类型更为多样

基于对阅读素养内涵的界定以及对阅读理解理论的认识, 为确保阅读素养评价的目的能够实现, PIRLS在测试文本的选择上可谓是煞费苦心, 其中既包括传统写作形式的文本, 也包括现代数字形式的文本; 既包括一个文本类型单独呈现的形式, 也包括多种文本类型组合的方式, 并为此制定了明确的选文标准, 包括文本的清晰性和一致性, 跨越不同国家和文化的内容适当性, 对大多数学生来说有趣的、吸引人的内容, 具有足够评估阅读理解全过程的基础等^[5]。同时, 为了达到近似真实阅读经历的评价目标, PIRLS要求呈现给学生的阅读文章和在线材料必须具有典型性, 能够反映学生在校内外真实的阅读经历和生活体验。

随着评价方式向数字化环境过渡, PIRLS2021在测试文本的类型上更为丰富多样。例如, 在2016年PIRLS两种阅读目的的文

本类型中, “为了获得文学体验而阅读”所使用的文本类型主要是“叙事小说”, 考虑到参与国家课程和文化的不同以及文本类型的特殊性, “诗歌”和“戏剧”并不包含在内^[6]; 而在PIRLS2021的测试中, 这一阅读目的的测试文本中增加了“戏剧”文本类型。同时, 在“为了获得和使用信息而阅读”的文本类型中, PIRLS2021也有了一定程度的发展, 如信息可以用多种不同的格式来呈现, 即使是主要通过文本呈现的信息片段也有可能包括用于记录事实的表格或用于说明描述的图片; 无论是印刷材料还是网站, 都会通过列表、图表、图形和图示呈现大量的信息; 混合型文本虽然并不新鲜, 但是随着新媒体和数字文本带来的传播方式和模式的迅速变化, 混合型文本也有了新的发展态势^[7]。尤其需要说明的是, PIRLS2021对电脑视窗下、互联网背景中网站或页面信息模式的呈现方式着重进行了阐释, 表明其对ePIRLS在线阅读文本类型的高度重视, 反映了当下学生在阅读方式上的实际变化。

（三）背景问卷的内容结构更加优化

除阅读素养评价外, PIRLS在每个评价周期还会通过对学校、教师、学生、家庭及参与国家的调查问卷, 收集并报告一系列与学生阅读素养成就表现相关的教育背景信息, 各国可以利用阅读成就与调查问卷的数据结果做出基于证据的阅读教育决策判断, 并为未来改善阅读教学提供政策依据。特别是可以根据PIRLS近20年来系列评价周期的

发展趋势数据，将学生阅读成就的变化与教育政策或实践变化联系起来，使其成为确定已有政策或方法是否有效的证据来源。

与以往的每次评价一样，PIRLS2021背景问卷框架在已有研究基础上进行了更新优化，以便使其获得的调查数据在帮助教育决策者和实践者提高学生阅读素养方面更加有效。为维持和扩展发展趋势数据调查，PIRLS的背景问卷框架包括将继续收集一些被确认为优先考虑的事项信息，但这同时也意味着为确保跟上时代发展的步伐，还需要设法解决新的研究领域和政策的相关性问题^[8]。

PIRLS2021背景问卷的基本框架和调查对象将继续保持不变，但在具体调查的内容和结构上却发生了明显改变。与2016年相比，PIRLS2021在背景问卷的设计上有三点不同的变化：一是问卷的各部分内容呈现出层级化的结构特点，一个标题统领几个具体的方面，这样在后期数据处理时便于将单个条目的内容合并成相关话题的主题量表，进行彼此关联的数据分析和挖掘；二是问卷内容进行不同程度的增删、细化和调整，进一步凸显其实效性，如“家庭背景”问卷中取消了“早期的计算活动”和“家庭的阅读支持”，增加了“父母喜欢阅读”“学前教育”和“小学阶段的早期读写任务”等内容；再如“学校背景”问卷中将“学校的位置”“受资源短缺影响的教学”“学校对学术成就的重视”等内容进行调整，突出

“学校的规模和城市化程度”“阅读教学资源”“学校对阅读教学的重视”等范围限定，聚焦更为明确；同时，对“安全、有序、纪律严明的学校”内容进行细化，具体包括“学校的纪律”“安全有序的学校”和“霸凌”，并且还增加了“学校氛围”这一调查话题，内容更为丰富；三是问卷重心更为关注“家庭”“学校”“课堂”三个维度，尤其是“学校”和“课堂”扩充的内容最多，所占比重最大，变化也最明显。这种变化一方面是基于前面几个评价周期的已有研究基础，另一方面也体现了PIRLS致力于对影响学生阅读素养成就可塑性内容要素的探索研究。

（四）采用群体自适应测试题册设计

PIRLS测试评价的目标人群是代表正式学校教育第四年的学生（在大多数国家和地区是四年级学生），但是随着参与测试评价的国家和地区越来越多，彼此之间教育发展水平的差异性表现得越来越明显，有的国家和地区的大多数四年级学生仍处于发展基本阅读技能阶段，尚不能达到参与PIRLS阅读素养评价的认知发展水平。如何在更大范围内使测试题册难度与学生能力发展水平相匹配成为一个巨大的挑战。为应对这一挑战，PIRLS在2011年开发了prePIRLS，2016年开发了PIRLS Literacy，通过提供难度较低评价版本的办法来解决这一难题。此做法虽然解决了能力分布处在低端发展水平学生的能

力覆盖范围，但并没有解决能力分布处在高端发展水平学生对更具挑战性测试材料和题目的需求问题。同时另需单独的评价版本，既增加了测试的成本费用，也增添了评价的复杂程度。

为了更好地解决参与国家和地区评估难度和学生成就水平相匹配的问题，PIRLS2021的测试题册采用了群体自适应设计。所有参与国家和地区都采用统一的评价内容，使用相同的阅读材料和测试题目，但是不同难度的阅读材料和测试题目在一个国家和地区的分配比例将根据学生的阅读成就水平灵活确定。简而言之，新的题本设计基于难、中、易三个水平等级的阅读材料和测试题目，并被合并成难度较大和难度较小两个不同难度等级的题册。每个国家和地区都使用相同的题册进行测试，但难度较大和难度较小题册的分配比例会随着该国和地区学生阅读成就水平的变化而变化^[9]。

具体来说，PIRLS2021的测试题册由18篇文章组成，其中有12篇文章来自PIRLS2016（含PIRLS Literacy测试中的文章），作为链接用作对阅读发展趋势的测量，另外6篇文章是2021年最新开发的。这18篇文章根据难易程度被划分成难、中、易三个水平，每个水平等级分别包括3篇文学类文章和3篇信息类文章。文章难易程度的区分依据学生群体对其正确作答的平均百分比，理想中合理的平均百分比——水平等级“难”

组为40%，水平等级“中”组为60%，水平等级“易”组为80%。PIRLS2021新开发的文章有意识地呼应了以上难度级别要求，而之前已有文章的难度表现则逊色一些，特别是“难”和“中”水平等级的文章，其难度值高出不少。当然，这只是过渡时期的暂时表现，今后随着更多趋势性文章被更有针对性的新文章取代，这样的情况会有所改变。

PIRLS2021中的18篇文章将被编排成18本题册，每本题册均包含一篇文学类和一篇信息类文本。18本题册将被分成难度较大的题册（9本）和难度较小的题册（9本）。难度较大的题册由两篇难度大的文章或一篇中等难度的文章和一篇难度大的文章组成，难度较小的题册由两篇或一篇容易的文章和一篇中等难度的文章组成。18本题册会分发到每个国家和地区，以确保所有参与评价的国家和地区测试的内容相同。根据PIRLS先前周期测试结果或首次参与国家和地区实地测试表现中学生的平均阅读能力，难易程度不同的题册在分配中所占的比例也不同，使每个国家和地区的测试难度与其学生的阅读能力更好地匹配。

根据阅读成就量表得分情况，PIRLS把所有参与测试的国家和地区分成高、中、低三个水平层级，以此确定不同难度等级题册的分配比例：总体平均阅读成就约为550分或者更高的国家和地区，将随机分配更多难度较大等级的题册(70%)，较少难度较小等级的

题册(30%);总体平均阅读成就在量表中点、约为500分的国家和地区,将分配相同比例难度较大等级的题册(50%)和难度较小等级的题册(50%);总体低于平均表现水平、大约450分或更低的国家和地区,将分配较少比例难度较大等级的题册(30%)和更多比例难度较小等级的题册(70%)。

测试题册群体自适应设计可以说是PIRLS2021阅读素养评价的最大亮点。这种创新的、自适应的题册设计将彻底改变PIRLS之前评价周期对不同阅读水平国家和地区各类阅读水平测试的做法,将所有的国家和地区统一到相同的模板中来,提高了测试内容的灵活性、测试学生的积极性和阅读水平的匹配性,既适用于国家与国家之间,也适用于国家内部不同群体之间,值得称道和推广。

(五) ePIRLS和digitalPIRLS有机整合

PIRLS2021首次推出基于互联网的数字化评价版本——digitalPIRLS,这样一来可以将其与PIRLS在2016年启动的ePIRLS在线信息阅读评价融为一体,进行无缝对接^[10]。作为digitalPIRLS评价版本的重要组成部分,2021年ePIRLS在线信息阅读评价包括5个阅读任务,其中的3个阅读任务作为2016年ePIRLS的一部分已经实施,主要用于学生阅读素养发展趋势的测量,另外两个阅读任务是2021PIRLS新开发的。

作为未来整合PIRLS和ePIRLS的第一步,在实行digitalPIRLS评价版本的国家

和地区中,ePIRLS的阅读任务也会包括在题册分配的循环方案中,以混合题册面貌出现,即由一篇digitalPIRLS信息类文章和一个ePIRLS阅读任务组成,这意味着参加digitalPIRLS评价版本的学生可能会收到三种题册类型中的一种:包含两篇digitalPIRLS文章的普通题册,或包含两篇ePIRLS阅读任务的ePIRLS题册,或包含一篇digitalPIRLS信息文章和一篇ePIRLS阅读任务的混合题册。

DigitalPIRLS评价版本题册的组合方式与paperPIRLS评价版本的题册组合方式是完全一致的,与纸质版本一样,有18篇相同的文章分配在同样的18本普通题册中,所不同的是digitalPIRLS还包括了ePIRLS题册和混合题册。与2016年所使用的ePIRLS分配方案类似,2021年ePIRLS的5个阅读任务中的每一个都与其他阅读任务进行配对,把每个阅读任务在题册中出现的先后顺序不同考虑在内会产生20本不同的ePIRLS题册。在混合题册中,digitalPIRLS的文章总是在ePIRLS阅读任务之前,9篇digitalPIRLS信息类文章中的每一篇都与5个ePIRLS阅读任务中的每一个进行配对组合,最终会产生45本混合题册。如此一来,将有18本digitalPIRLS普通题册、20本ePIRLS题册和45本混合题册,共计83本题册可供被试学生轮换使用。但是,如果以上每一种题册类型都以同样的比率进行轮换使用的话,digitalPIRLS普通题册的

数量会显得相对较少，特别是与实施纸质版本评价的国家和地区相比更是如此。为了避免这种不平衡，digitalPIRLS仅使用其中的27本题册，包括全部18本digitalPIRLS普通题册、6本ePIRLS题册和3本混合题册。每个学生的题册由IEA的WinW3S校内抽样软件进行分配，以确保轮换方案的准确实施。

WinW3S随机选择题册的第一轮循环包括所有的18本digitalPIRLS普通题册，20本ePIRLS题册中的1~6本，混合题册中的1~3本；第二轮循环再次包括18本digitalPIRLS普通题册，ePIRLS题册中的7~12本和混合题册中的4~6本；接下来第三轮循环选择，仍包括18本digitalPIRLS普通题册，13~18本ePIRLS题册和7~9本混合型题册，全国各地所有学生的样本都是如此。这种题册分配方案的结果是，2/3的学生将对digitalPIRLS的题册进行作答，2/9的学生将对ePIRLS的题册进行作答，1/9的学生将对混合型的题册进行作答，由此实现ePIRLS和digitalPIRLS的有机整合。

二、PIRLS2021阅读素养评价的启示

PIRLS2021阅读素养评价的最新发展动向为我国中小学的阅读素养教学及评价提供了有益启示，主要体现在四个方面。

（一）应重视数字化阅读素养培养

随着信息技术的迅猛发展，电脑、手机等数字设备的普及，儿童和青少年有更多的机

会进行数字化阅读^[11]。与传统的纸媒阅读相比，数字化阅读具有跳跃性、交互性和可扩展性（即读者可以通过超文本链接到其他相关网络文本以获取更多的信息）等特点^[12]，在为儿童和青少年学习提供便利的同时，也带来了很大挑战。面对网页呈现的各种纷繁复杂信息，明确自己的阅读目的并精确定位，有效监控自己的阅读理解和过程，评判网页信息的科学性和准确性，以及对多个链接文本的信息进行梳理和归纳，是阅读者在数字化阅读环境中所应具备的基本素养。

尽管数字化阅读已经成为一种发展趋势，但在学校教育层面，有关数字化阅读的相关教学和指导却较为匮乏。PIRLS阅读素养评价项目从2016年开始测试学生的数字化阅读素养，在全球范围内引起越来越多的教育者对儿童和青少年数字化阅读素养的关注和重视，建议我国在教育方案制定、学校课程管理、课堂教学实施等环节加强这方面的引导和实践。

（二）丰富阅读教学评价文本类型

PIRLS阅读素养评价将测试文本划分成“为了获得文学体验而阅读”和“为了获得和使用信息而阅读”两种类型，并在2021年首次增加了“戏剧”这一文学性文本类型。实际上，PIRLS阅读素养评价涵盖的文本类型非常多元，包括新闻报告、小说、自传、书信、图表、表格、地图、广告等等。相比而言，我国中小学阶段现行的语文教科书和各类有关阅读测试的考试评价中，在文

本类型的选用上还不够多元，实用性还不够突出，与社会发展的要求还有很大差距，需要结合语文课程的目标和功能定位，逐步调整，不断优化。

此外，戏剧作为一种文学类文本类型，在培养学生审美情趣和艺术品味等方面具有独特的优势，加之戏剧文本本身直接为舞台表演服务，更容易引发学生的学习兴趣。因此，建议在我国中小学语文教学和测试评价中，可以考虑融入戏剧素材，进而探索更为有效的戏剧教学和评价方式。

（三）综合联动构建阅读学习环境

中小学生的阅读素养会受到多种因素的影响，需要国家、社会、学校、家庭相互配合，协同发力。长期以来，我国小学生阅读素养发展所涉及的国家及社区、家庭、学校、课堂、学生特征与态度等综合背景因素被忽视，割裂了小学生阅读活动的整体性^[13]。

PIRLS2021提供的背景问卷内容更为丰富和精细，也为我们思考如何更好地培育学生的阅读素养提供了一个多维的科学、合理的理论框架。无论是家庭、学校、社会（社区），还是国家，都可以从这份框架中获得学生阅读指导的重要信息。例如，作为一名学校管理者，要想进一步提高全校学生的阅读素养，除了重点关注学校阅读环境的打造、阅读课程教学的质量外，还要关注学校资源、学校氛围、学校纪律和安全等多个可能影响学生阅读素养培育的要素。作为一名语文教师，要关

注学生对阅读的投入、教师对学生阅读理解技能和策略的培养、班级图书馆的建设和维护、促进阅读素养的课堂评价等，同时还要关注学生在线学习中的数字化阅读指导，加强班级管理 and 自身的专业发展。总之，只有树立综合、联动的阅读素养培育观念，才能更好地促进青少年学生的阅读素养发展。

（四）积极探索计算机自适应测评评价

计算机自适应测试是基于项目反应理论的一种测试方式，测试自行适应被试者水平，灵活施测与被试者能力水平相匹配的题目，实现测试的准确与高效^[14]。由于计算机自适应测试具有多种潜在优势，不少国家或地区都已纷纷探索计算机自适应技术在教育监测、人才选拔考试中的应用。PIRLS2021测试题册采用了群体自适应设计，提高了测试内容的灵活性、测试学生的积极性和测试水平的匹配性，是PIRLS阅读素养评价的一个重大突破。实际上PISA国际学生评估项目已经于2018年开始引入多阶段自适应测验（Multistage Adaptive Test）技术来测量学生的阅读素养水平^[15]。相信在不久的将来，计算机自适应测评技术也将会应用于国内大型的阅读素养评价项目中，以精准测查学生的阅读素养表现。

三、PIRLS2021阅读素养评价的创新与发展

PIRLS自2001年正式启动以来，一直致

力于阅读素养评价内容和方式的创新与发展。2006年,测试材料由8篇文章扩展到10篇文章,除了对整体阅读素养和每种阅读目的的结果进行报告之外,还对阅读理解过程的评价结果进行了报告;2011年,PIRLS和TIMSS的评价周期相遇,这为各国和地区同时收集四年级学生在阅读、数学和科学上的成就数据提供了难得的机会,除了背景问卷以外,还扩大了PIRLS的测试范围,包括鼓励一些国家和地区的五、六年级学生参加测试,推出prePIRLS测试版本,以满足大多数四年级学生仍处在发展基本阅读技能的国家 and 地区需求;2016年,PIRLS的测试材料增加到20篇,开发了一种难度相对较低的阅读评价选择——PIRLS Literacy和一种全新的阅读评价方式——ePIRLS,积极面对现实诉求,回应时代发展潮流,直至2021年第五个评价周期,其发展进步仍然是有目共睹的。

从PIRLS20年的发展历程中可以发现,变与不变是其始终坚持的追求,我们应在准确把握其不变的精髓前提下,及时关注其最新发展动态,为我国基础教育学生阅读素养的培育和评价提供有益启示。❖

参考文献:

- [1]PIRLS—Progress in International Reading Literacy Study [EB/OL].<https://www.iea.nl/studies/iea/pirls#section-123>,2021-02-23.
- [2]PIRLS Assessment Options—digitalPIRLS [EB/OL].<https://www.iea.nl/studies/iea/>

pirls#section-120,2021-02-23.

[3][7][8][9][10]IEA.PIRLS2021 Assessment Framework [EB/OL].http://pirls2021.org/wp-content/uploads/sites/2/2019/04/P21_Frameworks.pdf,2021-02-24.3,18,27,57-58,65-66.

[4]PIRLS 2021: Transitioning to Digitally Based Assessment [EB/OL]. <https://timss.bc.edu/pirls2021/index.html>,2021-02-24.

[5][6]IEA.PIRLS2016 Assessment Framework, 2nd Edition[EB/OL].https://timssandpirls.bc.edu/pirls2016/downloads/P16_Framework_2ndEd.pdf, 2021-02-26.27,16.

[11]黄志军,王晓诚.PISA和PIRLS数字阅读素养测评框架探析[J].外国中小学教育,2018,(01):1-7.

[12]王晓诚.PISA数字化阅读素养测评及其启示[J].外国中小学教育,2018(05):1-8.

[13]宋乃庆,肖林,程浩.小学生阅读素养的背景因素探析——基于国际阅读素养进步研究视角[J].中国教育学报,2017(02):61-66+104.

[14]漆书青,戴海琦,丁树良.现代教育与心理测量学原理[M].北京:高等教育出版社,2002:82.

[15]孙小坚,宋乃庆,辛涛.PISA测试中多阶段自适应测验的实施及启示[J].现代教育技术,2021,31(06):72-78.

(责任编辑 王芳)