



教育部考试中心
NATIONAL EDUCATION EXAMINATIONS AUTHORITY



PISA2015中国国家中心
PISA2015 National Center of China

PISA

中国PISA2015

测试实践指导

Programme for International Student Assessment

教育部考试中心 ★ 编



广东高等教育出版社
Guangdong Higher Education Press

中国PISA2015

测试实践指导

ZHONGGUO PISA2015 CESHI SHIJIAN ZHIDAO

教育部考试中心★编



广东高等教育出版社
Guangdong Higher Education Press

• 广州 •

图书在版编目 (CIP) 数据

中国 PISA2015 测试实践指导/教育部考试中心编. —广州: 广东高等教育出版社, 2016. 7

ISBN 978 - 7 - 5361 - 5647 - 0

I. ①中… II. ①教… III. ①学生 - 教育评估 - 研究 - 中国
IV. ① G449. 7

中国版本图书馆 CIP 数据核字 (2016) 第 126607 号

出版发行	广东高等教育出版社
	地址: 广州市天河区林和西横路
	邮政编码: 510500 电话: (020) 85250745
	http://www.gdgjs.com.cn
印 刷	佛山市浩文彩色印刷有限公司
开 本	787 毫米 × 1 092 毫米 1/16
印 张	18.5
字 数	342 千
版 次	2016 年 7 月第 1 版
印 次	2016 年 7 月第 1 次印刷
定 价	46.00 元

本书系国家教育体制改革领导小组办公室委托项目
“研提 OECD 组织 PISA 测试项目经验启示”
的研究成果之一



《中国 PISA2015 测试实践指导》编委会



主 编：李光明 王 蕾
副 主 编：汤贞敏 李文郁
编 委：佟 威 王羽佳 丁慧明
 符华均 焦丽亚 许世红
 张伟民 孟 彦 景安磊

前言

国际学生评估项目（Programme for International Student Assessment, 简称 PISA）是由经济合作与发展组织（Organization for Economic Co-operation and Development, 简称 OECD）于 2000 年发起的对基础教育质量和公平进行的跨国、跨文化的比较研究项目，主要通过测评学生运用知识和技能解决现实问题的能力，同时通过问卷调查的方式收集学生、教师和学校的背景信息来分析影响学生成绩的因素，形成对整个教育体系的评价报告，从而帮助参与国家和地区了解其在国际比较中的相对位置并发现现行教育体系的优、劣势，为改进和完善现行教育政策提供借鉴。PISA 凭借其引领性的测评理念、科学的方法和先进的考试评价技术，已经成为最具国际影响力的教育评价项目。

当前，我国正处于重启教育改革议程的关键期，学习借鉴国际先进的教育理念、教育方法、评价体系等内容，有助于推动我国教育模式的创新及教育观念的改变。PISA 作为当今最具国际影响力的教育评价项目，有很多值得我们学习借鉴的地方。一是 PISA 从关注学生能否记忆和再现知识内容转移到学生能否灵活拓展和应用知识及技能，测评内容超越了学校的课程标准，不以学科知识为中心，反映了社会生活环境对学生的实际要求。PISA 同时深入分析学生成绩的多元影响因素，从理论和实践上论证了“公平有质量”的教育才是真正卓越的教育，质量和公平是完全有

可能统一的，强调卓越的教育需要同时具备较高的学生素养和较低的地区、校际差异，这些也都是我国教育改革发展实践正在践行的理念。二是《国家中长期教育改革和发展规划纲要（2010—2020年）》颁布实施以来，贯彻落实教育规划纲要的切入点是什么，着力点在哪，在纷繁复杂、看法不一的情况下怎么形成共识，在要说的、要做的事情特别多的情况下，怎么切实抓纲带目，有序有效有力指示各地各校静下心来，扎扎实实抓教育公平、教育质量，PISA为解决上述问题从而促进教育改革发展提供了框架。同时，面临全球化进程的加速，我们也需要借助国际视野去审视国内基础教育质量，探索用国际化的教育质量标准 and 科学的测评手段评价基础教育阶段的学生学业质量，PISA测试对中国教育质量标准的制定、监测和发布有很大的借鉴作用。三是PISA的证据文化利于提高教育决策科学化水平。在知识洪流和数字化时代，以量化数据为基础的证据文化在公共管理和决策过程中发挥了重要作用，没有量化数据的支撑，管理和决策将无所适从，对真实的量化数据的掌握程度将影响管理和决策的质量。但是，大量的数据和繁杂的信息需要过滤程序的整理、比较和分析，从而挖掘隐于其“后台”的规律和特征，才能为管理和决策服务。当前，对许多教育难题，国内不乏有知之士、有识之士、有智之士，可惜这些教育研究和教育行政工作常常是脱节的，甚至有些同志对某些教育研究是不屑的。这很不好！PISA基于2000年、2003年、2006年、2009年和2012年的五轮测评，构建了大规模教育测评数据库，其服务于教育决策改进的成功经验为我们提供了借鉴。PISA的参与成员通过横向国际比较和纵向发展比较，实现了数据库资源的共享，在数据挖掘技术的支持下，利用PISA数据库监测本国或本地区教育发展状况，及时发现存在问题，大大提高了教育决策的科学性，促进了教育改革发展。因此，教育质量监测和评价要求培育证据文化，建立量化信息数据库，提高教育决策科学化水平。四是PISA的测试理念、命题方式、阅卷评分、机考模

式、数据挖掘等为剖析当前教育考试评价领域中存在的诸多问题提供了可行路径。当前,我国教育考试评价领域中的“唯分数论”影响学生全面发展,一考定终身使学生学习负担过重。PISA 以评价学生的核心素养为根本宗旨,尤为重视学生的批判反思能力和解决实际问题的能力。此外,PISA 的命题方式、阅卷评分、机考模式、数据挖掘等也可以为我国的教育考试制度改革提供较好的借鉴。当然,我们也需要辩证地去看待 PISA 测试,以达到为我所用的目的。

正是基于上述原因,教育部考试中心从 2006 年起“自备干粮”,先是以科研名义,以学习先进的考试评价技术为目的开展了 3 个轮次的“PISA 中国独立研究”,组建了专业化的研究团队,提交的《PISA 中国独立研究报告》也得到了教育部领导的高度评价和重视,报告还曾先后呈送国家教育咨询委员会和刘延东同志。2009 年和 2012 年,上海市正式参加 PISA2009 和 PISA2012 测试,并两次取得全球第一的优异成绩,引起国际广泛关注,我国教育的国际影响力进一步提升。2014 年,为在国际舞台上充分展示我国教育改革取得成就,坚持以开放促改革、促发展,教育部与 OECD 签订协议,选取北京、上海、江苏、广东等 4 个省(市)正式参加 PISA2015 测试,并将“PISA2015 中国国家中心”设在教育部考试中心,具体负责测试管理和组织实施工作。2015 年 4—5 月,在历经学校、教师和学生的信息收集、试题和问卷的翻译改编、试测反馈、正式抽样等几个阶段的工作后,我国完成了 PISA2015 正式测试和阅卷编码工作。目前,PISA2015 中国国家中心正在开展数据清理、专题研讨、报告撰写等工作,OECD 将于 2016 年 12 月公开发布国际测试结果,我国也将同步发布国内测试结果和研究报告。

与以往几个轮次的测试相比,PISA2015 在测试内容和形式上都出现了新的进展,内容上除了常规的阅读、数学、科学和财经素养等,还新增了对协作问题解决能力(Collaborative Problem

Solving) 的测试;在形式上,PISA2015 摒弃了传统的纸笔测试,采用了基于计算机化的测试模式,学生可以通过计算机完成试题作答和问卷填写,校长和教师也是在线填写调查问卷。

我国 PISA2015 测试严格执行国际统一标准和技术规范,为了向国内同行系统介绍 PISA 测试的理念和具体操作过程,我们团队编写了这本书,以期推进我国教育领域的改革发展。本书凝聚了许多研究人员的集体智慧。主要分工如下:第一、七章由丁慧明撰写,焦丽亚负责统稿;第二、三章由佟威撰写,王蕾负责统稿;第四、五章由王羽佳撰写,景安磊负责统稿;第六章由符华均、孟彦撰写,焦丽亚负责统稿;第八章由许世红撰写,汤贞敏、李文郁、张伟民负责统稿。全书由李光明、王蕾、景安磊统稿。

本书在编写和出版过程中得到了很多单位和专家的大力支持,在此一并致谢。同时感谢广东省教育研究院和广东高等教育出版社对本书的大力支持。

由于各方面的原因,本书可能还存在一些欠妥之处,恳请广大读者不吝指正。

编 者
2016 年 1 月

目录

第一章 PISA2015 抽样	1
一、PISA2015 抽样任务与技术标准	1
(一) 抽样任务	1
(二) 技术标准	2
二、PISA2015 试测抽样设计	4
(一) 定义目标总体	5
(二) 确定抽样设计	5
(三) 收集学生分布信息	6
(四) 收集学校信息与试测排除类别信息	6
(五) 收集试测样本学校及其替补学校信息	7
(六) 收集试测学生、教师名单	7
(七) 抽样软件及试测名单导入	11
(八) 试测校内抽样	14
(九) 更新试测抽样数据	15
(十) 校验与提交试测抽样数据	17
三、PISA2015 正式测试阶段抽样工作	18
(一) 收集期望目标总体和确定目标总体信息	18
(二) 收集正式测试抽样框	19
(三) 抽取正式测试样本学校及其替补学校	20
(四) 收集正式测试学生、教师名单与校内抽样	20
(五) 更新正式测试抽样数据	22
(六) 校验与提交正式测试抽样数据	23
四、体会与启示	23

第二章 PISA2015 试题开发、翻译和改编	25
一、试题开发	25
二、试题的翻译改编审核程序	26
(一) 初始步骤	27
(二) 设置翻译人员	27
(三) 整合	27
(四) 校对审核	27
(五) 审核鉴定	27
(六) 本国复审	28
(七) 布局改编	28
(八) 跟进审核	28
(九) 最终确认	28
三、PISA2015 中国国家中心关于翻译改编工作的实施情况	29
(一) 团队的招募和组建	29
(二) 文件管理	29
(三) 翻译改编工具及操作方法	29
(四) 翻译改编注意事项	31
(五) 正式测试翻译改编的修订	34
四、体会与启示	41
(一) 与时俱进的试题	41
(二) 规范的翻译改编流程	41
(三) 严格的试题筛选标准	42
(四) 协商的工作机制	42
第三章 机考测试系统	43
一、计算机诊断程序	43
(一) 计算机诊断程序操作简述	44
(二) PISA2015 中国国家中心在试测阶段的检测工作	48
(三) PISA2015 中国国家中心在正式测试阶段的检测工作	50
二、学生机考系统	52
(一) PISA2015 试测阶段的学生机考系统	53
(二) 正式测试阶段的学生机考系统	61

三、体会与启示	68
(一) 关于计算机化考试的宣传和引导	68
(二) 关于计算机化考试的管理	69
(三) 关于信息化时代的大数据	70
第四章 阅卷编码	71
一、阅卷编码的理念	71
(一) 充分理解学生的作答思路, 进行有利推断	71
(二) 忽略微差错, 发现闪光点	72
(三) 矛盾要素和不相关要素	72
二、主观题编码设计	73
(一) 分数等级标签	73
(二) 代码编号	74
三、关于工作坊材料的说明	75
四、本国样例的说明	75
五、其他资料	75
六、多步骤的编码解决方案	75
(一) 严格忠实于编码指南	76
(二) 咨询培训专家	76
(三) 国际咨询和仲裁	76
七、试测阅卷编码	76
(一) 测试工具箱整理	77
(二) 阅卷编码前期准备	79
(三) 纸笔测试阅卷编码过程	80
八、试测机考编码流程	82
(一) 数据流转	82
(二) 编码员编号和数据拆分	84
(三) PDF 文件	84
(四) 编码流程	85
(五) 质量检验	86
九、正式测试机考编码	88
(一) 样例更新	88

(二) 编码员分组	90
(三) 编码质量监控	91
(四) PISA 正式测试质量监控操作规范	94
十、职业编码及学科编码	94
(一) 职业编码依据	96
(二) 职业编码人员	97
(三) 职业编码工作流程	97
(四) 科学课程名称编码	101
十一、体会与启示	102
(一) 理念上从传统的评分提升到现代的编码	102
(二) 在不同省份设置共同作答组, 监控各省份的阅卷质量	103
(三) 充分借助信息技术, 更大程度发挥网上阅卷的潜在优势	103
第五章 学校、教师和学生问卷	104
一、PISA2015 问卷设计	104
二、PISA 试测问卷的模块主题	104
(一) 优先关注模块	105
(二) 中度关注模块	107
(三) 剩余模块	107
三、问卷的翻译改编	108
(一) PISA2015 问卷的基本情况	108
(二) 问卷的翻译改编	109
四、测试实施——在线问卷的作答	119
(一) 学校问卷、教师问卷作答账号管理	119
(二) 监控在线问卷作答进展	120
五、体会与启示	121
(一) 高效的管理运行模式	121
(二) 系统的信息收集分析	122
第六章 PISA2015 考务管理	123
一、材料的改编、翻译和印刷	123
(一) 改编	123

(二) 翻译	124
(三) 印刷	124
二、考务人员	126
(一) 学校协调主任	126
(二) 测试主任	127
(三) IT 人员	128
三、考务培训	128
四、测试实施	130
(一) 测试材料的制作与运送	130
(二) 测试前一天布置机房	130
(三) 时间安排	131
(四) 实施过程	132
五、PISA 质量监控	133
(一) 任命质量监控员	133
(二) 确定质量监控校	134
(三) 质量监控员工作内容	135
六、PISA2015 国家中心值班	136
七、材料整理和数据录入	137
(一) 材料整理	137
(二) 数据录入	137
八、体会与启示	137
 第七章 PISA 数据管理	 139
一、前期准备	139
二、检查编码手册	140
三、测试 DME	143
四、导入机考数据	144
五、导入其他数据	146
六、录入数据	147
七、克隆与合并数据库	147
八、校验与提交数据	148
九、体会与启示	150

第八章 广东省 PISA2015 测试实践	152
一、广东省概况	152
二、广东省 PISA2015 实施组织机构	154
三、试测情况	154
(一) 试测的抽样	154
(二) 试测抽样中遇到的问题及解决办法	159
(三) 试测的组织与实施	160
(四) 试测结果分析	162
四、正式测试情况	165
(一) 正式测试的抽样	165
(二) 正式测试抽样中遇到的问题及解决办法	170
(三) 正式测试的组织、培训与协调	170
(四) 正式测试的实施	172
(五) 正式测试的其他相关情况	174
五、体会与启示	174
(一) 缺考问题	175
(二) PISA 研究	175
(三) 信息技术装备	175
附录 1 PISA2015 学生信息表	177
附录 2 PISA2015 正式测试抽样数据收集 FAQ	179
附录 3 PISA2015 学校协调主任工作简述	181
附录 4 PISA2015 测试主任工作简述	182
附录 5 PISA2015 正式测试常见问题解答 (FAQ)	186
附录 6 PISA2015 职业编码手册	188
一、职业大类、中类、小类、细类	188
1 管理人员	188
2 专业人员	195
3 技术员及助理专业人员	212
4 事务支持人员	227
5 服务及销售工作人员	233

6 农业、林业和渔业技术工人	240
7 工艺及有关行业工人	244
8 机械设备操作工人及组装工人	258
9 简单劳动职业人员	267
0 军人	273
二、编码思路与原则	276
三、相似职业的区别	277
四、职业编码举例	279

◆ 第一章 ◆

PISA2015 抽样

抽样是 PISA 测评的一个关键环节，抽样设计的合理性及数据信息的可靠性将最终影响测评结果的可信度及公信力。PISA 抽样综合采用多种抽样方法。从 PISA2000 至今，绝大多数参加国家或地区均采用两阶段抽样。第一阶段采用按规模大小成比例（Probability Proportional to Size，简称 PPS）系统抽样方法抽取学校。在抽取学校样本之前，由依据学校属性选定的外显分层变量将学校划分到不同分层中，再在每个外显分层中独立抽取学校。PISA 使用分层抽样是为了提高抽样设计的效率，从而确保获得可信的调查结果；确保样本涵盖了总体的各个部分，对总体具有良好代表性；与此同时，分层可满足对特定群体采用不同抽样设计的需求。第二阶段是在每所样本校内随机抽取样本学生，同一所学校内，每个符合 PISA 条件的学生被抽中的概率相同。从 PISA 以往轮次来看，仅有俄罗斯等极少数国家采用三阶段抽样，在这些国家，第一阶段则是用 PPS 方法抽取地理区域，第二、第三阶段再在抽中的区域中逐步抽取学校、学生。PISA2015 我国 4 省（市）作为整体抽取一个样本参加，使用两阶段分层抽样方法。

一、PISA2015 抽样任务与技术标准

（一）抽样任务

PISA2015 按照时间节点在试测阶段需要依次完成 ST0 ~ ST6 的抽样任务，正式测试阶段需要完成 ST7 ~ ST12 的抽样任务（Sampling Task，简称 ST），如表 1-1 所示。