

(上接第一版)

评价可能存在的误差的经验方法。上海交通大学杨惠中教授在题为《有效教学与有效测试》的报告中指出考试、教学和考试结果的使用这三者构成一个系统，要发挥系统的作用，三者必须协调，只有各方正确处理好各自的关系，做到有效测试、有效教学、有效使用，才能取得整个系统的最佳效果。英国 Cambridge Assessment 的 Nichlas Raikes 做了题为《维持不同考试划界标准的可比性》的主题报告，他以英国普通中学教育证书获得为例，分享了英国在维持不同考试划界标准的可比性方面的具体做法和经验。当天下午，在平行分会场，与会人员围绕“测评技术运用”“学习能力测评”“测评方案设计”“高校选拔标准”等议题进行了深入探讨和交流。

29日，与会人员继续聚焦测评与学习中的共同关切，激发思维，分享研究成果。

29日下午的闭幕式上，美国 ETS 专家 John Peng 首先以《题库建设及内容架构与呈现方式》为主题作大会报告，他从命题内容结构分析、呈现方法选择、设计的灵活性与可拓展性、成熟安全的主流技术选用、智慧决策等方面进行了经验分享。美国 ACT 专家 Wai Nga Woo 在题为《用于学习测评且基于人工智能的推荐及诊断引擎》的主题报告中，从美国 ACT 如何在能力模型、内容界定、配置技术、建立原则等方面通过技术监控、诊断学习、促进提高从而服务考生的个性化学习。四川省教育考试院院长刘敏作嘉宾致辞。上海市教育考试院院长郑方贤致闭幕词，他表示，以考试分数为对象的评价从以往普通的总分排序（第一阶段），到现在分维度分指标项的分析评价（第二阶段），酝酿进入利用大数据技术开展综合性分析以服务于教师教学与学生学习（第三阶段），最终推动学生的学习能力提升，这也是“考试-教育-学习”的紧密关系。相信与会人员通过国际研讨会这个学术交流平台，在探寻教育考试的本源中，进一步厘清思路，坚定探索“测评”和“教学”的融合之路。



为期两天的研讨会，200 余名与会代表围绕“测评也是学习”的会议主题进行了广泛而深入的探讨。学术氛围浓郁，研讨气氛热烈，共组织了 1 场主旨报告、5 场专题报告（测量原理、有效教学、划界标准可比性、题库建设、人工智能推荐及诊断），40 位嘉宾进行了报告交流。演讲者来自国内外，研讨案例也涵盖国内外，既有成熟理论和方案介绍，也有技术方法在不同地区、领域应用的分析研究，得到了国际知名考试测评机构、各省市考试机构、各区教育机构和高校、中小学校的积极响应。研讨会征集到了来自全国各省市专家学者的 40 余篇学术论文，9 名作者在会议现场进行了分享，为诸多焦点问题提供了实践方案。

科研速递



- ◆省教育厅日前面向省内外开展“2019 年江苏教育改革发展战略性与政策性研究课题”公开招标工作。此次课题招标设 10 个重大课题，其中综合研究课题 3 项、专项研究课题 7 项。申报截止日期至 2020 年 2 月 29 日。详见江苏省教育厅或江苏省规划办网站查询。
- ◆教育部考试中心日前公布《国家教育考试科研规划 2019 年度立项课题》，南京大学王运来《高等教育普及化背景下的自学考试转型发展研究》，江苏省教育考试院李拥军《自学考试学习资源开发建设研究》获重点资助课题，江苏省教育考试院马秀谊《高中学业水平考试选择性考试命题策略研究》获一般课题。
- ◆2018 年度“江苏省社科应用研究精品工程奖”获奖项目日前公布，江苏省教育考试院马秀谊《基于思想政治学科核心素养的考试命题与评价》获得一等奖。这是我省招考系统唯一获奖项目。

# 江苏教育考试科研月报

2019 年第 12 期（总第 60 期）

江苏省教育考试院编印

科研要闻

## 2019 考试评价国际研讨会在沪举行

2019 年 11 月 28 日—29 日，由上海市教育考试院主办的 2019 考试评价国际学研讨会在上海顺利举行。会议以“测评也是学习”为主题，围绕测评技术运用、学习能力测评、测评方案设计、高校选拔标准、在线考试探索、K-12 科学素养测评、测评素养研究等议题研讨交流最新的评价理论与实践成果，深化对考试测评的认识，为实践打下基础。我院院长袁靖宇、副院长吴仁林应邀出席。我省招考系统提交三篇学术论文。来自美国 ETS、ACT，英国 Cambridge Assessment、British Council，爱尔兰 Trinity College Dublin、澳大利亚 Victorian Tertiary Admission Centre 等国际著名考试评价机构的专家学者，以及各省市考试机构、高校和中学、科研部门的 200 余位专业人士参会。

28 日上午的开幕式上，上海市教委副主任倪闽景致大会欢迎辞，教育部考试中心主任孙海波向研讨会发来贺信。

剑桥大学考评院（Cambridge Assessment）、美国教育考试服务中心（ETS）、美国大学入学考试中心（ACT）相关负责人发来祝贺视频。

开幕式由上海市教育考试院党委书记、副院长刘玉祥主持，他表示，2017 考试评价国际研讨会和 2018 学业水平考试学术研讨会的真知灼见已经得到了各界肯定，对深化高考改革发挥了良好作用，并对全球考试评价机构对本次大会的积极响应表示感谢。上海市教育考试院院长郑方贤作大会主旨报告《测评：从选拔、评价到学习》，报告中介绍了测评的意义与作用，以及测评与学习的关系。美国教育考试服务中心（ETS）研发部首任测量创新中心主任、国际教育评估协会（IAEA）主席 Randy E.Bennett 博士作《将测量原理纳入形成性评价中》的主题报告，

他介绍了形成性评价的概念框架、实施前提及四个最基本原理，并阐释了如何将测量原理科学地纳入形成性评价，减少



(下转第四版)



【编者按】2019年11月28日—29日，由上海市教育考试院主办的2019考试评价国际学研讨会在沪顺利举行。会议围绕测评技术运用、学习能力测评、测评方案设计、高校选拔标准、在线考试探索、K-12科学素养测评、测评素养研究等议题研讨交流最新的评价理论与实践成果。我省招考系统共提交三篇论文，现将三篇论文主要观点摘编，供读者参阅。

### 高考综合改革背景下命题工作若干问题研究

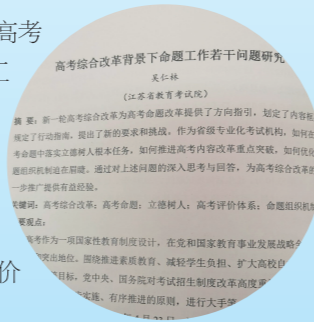
吴仁林  
(江苏省教育考试院)

摘要

新一轮高考综合改革为高考命题改革提供了方向指引，划定了内容框架，规定了行动指南，提出了新的要求和挑战。作为省级专业化考试机构，如何在高考命题中落实立德树人根本任务，如何推进高考内容改革重点突破，如何优化命题组织机制迫在眉睫。通过对上述问题的深入思考与回答，为高考综合改革的进一步推广提供有益经验。

主要观点

- ◆如何在高考命题中落实立德树人根本任务问题。立德树人是教育的根本任务，是高考育人的重要使命，也是高考命题改革的内在要求，如何将立德树人贯穿于命题工作全过程，突出命题的思想性和育人功能？一是通过彰显命题的价值导向落实立德树人。二是通过服务选才落实立德树人。三是通过引导教学落实立德树人。
- ◆如何推进高考内容改革重点突破的问题。面对当前高考综合改革内容突破的形势要求，省级高考命题如何进行有效破解和推进？一是要准确理解中国特色的高考评价体系。二是主动融入全国高考评价体系。三是要积极探索素养导向的高考评价体系。
- ◆关于优化命题组织机制的问题。高考综合改革，牵涉到命题领域多种要素的新挑战，要从优化命题实践组织机制、加强命题全流程管理入手。一是建立高考命题项目分类管理模式。二是探索命题审题评卷联动衔接机制。三是构建命题质量管理闭环控制体系。



### 面向地区和学校的增值评价报告体系框架设计

高柳萍 张 澜  
(南京大学 江苏省教育考试院)

摘要

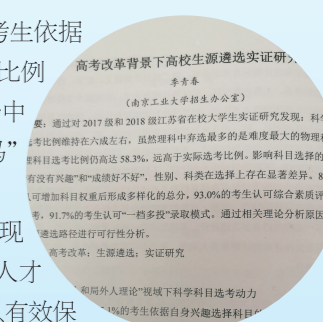
随着我国新高考改革的不断深入推进，越发凸显教育测量和评价技术在推动教育改革发展、深化评价工作方面的重要性。目前国内通用的教育评价方式是直接通过高考成绩衡量教育教学，但这种评价方式主要关注学生的高考表现，忽略了学生生源差异和学生个体的成长，基于纵向数据的发展性评价亟需建立。因此，本研究引入增值评价方法，以江苏省某市的数据为例，设计面向地区和学校的增值评价报告体系框架，更好发挥高考正向指挥棒作用。

摘要

通过对2017级和2018级江苏省在校大学生实证研究发现：科学科目总体选考比例维持在六成左右，虽然理科中弃选最多的是难度最大的物理科目，但物理科目选考比例仍高达58.3%，远高于实际选考比例。影响科目选择的因素主要是“有没有兴趣”和“成绩好不好”，性别、科类在选择上存在显著差异。85.7%的考生认可增加科目权重后形成多样化的总分，93.0%的考生认可综合素质评价作为重要参考，91.7%的考生认可“一档多投”录取模式。通过相关理论分析原因，并对新型生源遴选路径进行可行性分析。

主要观点

- ◆“局中人和局外人理论”视域下科学科目选考动力。本研究发现，若35.1%的考生依据自身兴趣选择科目的话，物理科目的选考比例可达58.3%，已远远高于实际选考比例40.9%，为何实际数据以及浙江调研数据远远小于本研究数据呢？原因在于“局中人和局外人”理论下，大部分考生放弃了自身兴趣，根据多方因素进行“田忌赛马”博弈选科的结果。
- ◆“多元智能理论”视域下综合素质评价参考动力。充分尊重考生个体差异，发现和评价考生的不同功能，即参考学生过程性的综合素质评价结果，应该是高校人才选拔的有益尝试。而综合素质评价真实性是其有益参考的前提，区块链技术可以有效保障综合素质评价信息的真实性，从而促进高考公平。
- ◆新型生源遴选路径可行性分析。研究表明，91.7%的考生认可“一档多投”招生录取方式。“一档多投”可以彻底解决高校招生自主性和公平性的矛盾。更为重要的是，可以有效打破“唯分数论”，通过参考综合素质评价，多样化的录取标准将推动高中教育进行一场深刻的变革。



主要观点

- ◆目前，我国通用的教育评价方式仍是直接通过高考成绩开展结果评价衡量教育教学，这种仅依据单一的高考成绩数据开展评价存在局限性，不仅低估了高考成绩复杂性，也窄化了普通高中发展路径。
- ◆本研究立足于江苏省某市的相关数据，尝试将增值评价法运用于大规模教育考试数据分析当中，借助国内外研发经验，来设计增值评价指标体系，初步形成面向地区和学校的增值评价报告体系框架。
- ◆增值评价报告体系框架主要精力三大设计研发阶段。一是前期数据收集。收集江苏省某市43所学校3193名学生的性别、户籍、家庭经济状况、学校类型、中考成绩、高考成绩等数据。二是增值评价模型建立。使用HLM7.0软件，构建以学生个体水平为第一层变量，学校水平为第二层变量，中考成绩为协变量，高考成绩为因变量的2水平线性回归模型，开展增值评价。三是增值评价指标体系框架设计。学习借鉴了国内外增值评价系统研发经验，结合江苏省教育实际，初步形成增值评价报告体系框架，分为地区版和学校版两类报告，框架内容遵循从宏观到微观，从整体到局部的原则来评价某地区或学校的教育教学。
- ◆地区版增值评价报告主体包括两部分。第一部分介绍报告的数据说明、增值评价模型的建立和统计指标及术语说明；第二部分介绍语文、数学、英语科目的增值评价指标体系，主要内容包括：地区整体增值水平、不同高中学校增值水平比较、不同学生群体增值水平比较、结果评价与增值评价比较。
- ◆学校版增值评价报告主体也包括两个部分。第一部分内容与地区版报告相同；第二部分介绍语文、数学、英语科目的增值评价指标体系，主要内容包括：增值水平概况、不同起点学生增值水平比较、不同学科增值水平比较、不同类型校增值水平比较。

