

# 高考科学性与受众接受度的 断裂及治理困境的突破

——以标准分制度的沿革为例

张会杰

(华东师范大学, 上海 200062)

**摘要:** 选用政策的科学性与对政策的接受度两个变量分析我国高考历史上曾经在七个省份试点和推广历时二十余年的标准分制度, 可以发现, 这项改革面临制度科学性与受众接受度断裂的治理困境。高考制度的改革必须以受众的普遍接受为基础, 否则, 即便仅仅只是技术层面的改良也将难以为继。高考治理困境的突破在于持续关注并强化受众对特定改革举措的接受程度, 逐步提升高考治理的服务性。在增进受众可接受度方面, 当年高考结束后, 关键数据信息脱敏后的适度公开及必要解释具有一定的可行性。这一提议或将面临诸多挑战。应对挑战需要理论界对关涉高考考试安全的数据信息的认知尤其是认知的误区进行澄清, 同时在顶层设计上加强信息公开的规划部署, 制定或完善相关法律或部门规章, 确立不同开放模式的实施程序, 多种途径保障数据安全, 以及组织开展基于高考数据信息的理论研究与舆情引导。

**关键词:** 高考制度; 标准分; 科学性; 接受度; 治理困境

**中图分类号:** G642.47 **文献标识码:** A **文章编号:** 1672-4038 (2019) 06-0043-08

1977年, 高考制度恢复。制度的重建为个体和社会带来新的发展契机, 但由于长期停滞, 高考的实践经验与理论积累都颇为匮乏。客观地看, 高考恢复后起初几年的考试质量并未得到充分的保障。这一问题很快得到教育主管部门的重视和回应。1985年1月, 教育部在广州召开第二届高考科研讨论会, 决定会后由广东省率先进行高考标准化改革的试验。试验采取“由点到面、逐步推进”的策略, 即部分学科、省份先试, 再逐步向其他地市推广。标准化改革试点期间进行了若干重要的探索, 在命题和

考试实施标准化方面, 试点地市开始探索以高考计分方式与分数报告为突破口, 采用将卷面原始得分转换为标准分数的标准分制度。标准分制度于1985年在广东率先试点, 先后有海南、河南、陕西、广西、山东、福建等七个省份加入进来, 辐射范围广泛, 影响巨大。然而, 从2001年开始, 试点省份相继停止使用标准分, 2007年, 最早的试点省份广东省也宣布停止使用。时至今日, 全国只有海南省还在使用这一分数制度。

实施标准分制度是我国高考发展史上一次

收稿日期: 2019-05-09

基金项目: 全国教育科学规划教育部重点课题 (DIA170359)

作者简介: 张会杰, 女, 华东师范大学教育学部考试与评价研究院副教授, 主要从事教育考试及评价制度研究。

标志性改革，被看作是恢复高考后考试技术层面的一次重大突破。测量学界普遍认为，相比于原始分数，标准分在分数评定的精确性、人才选拔的效度以及对公平的保障上具有明显的优势。这项公平与效率兼备的考试改革何以出现这种结局？有研究从制度变迁多重逻辑的视角对标准分制度发展沿革中政府部门、考生、专业群体、中学与高校的制度逻辑进行了分析，研究提出，标准分制度不能满足各主要利益相关者的诉求，政府在应对上的民意依赖使得这一整体上功能更佳的制度难以为继。<sup>[1]</sup>在本轮高考制度改革正在全面开启的时代背景下，对这项曾在七个省份试点和推广历时二十余年的改革举措的多角度思考，可为我们更好地理解当下高考改革面临的挑战提供借鉴。本文选用公共政策制定和执行中的两个核心变量即政策科学性和政策的可接受度，对标准分制度的发展沿革及其治理困境进行初步的探讨，期待通过广泛讨论推动该议题研究的深化。

## 一、政策科学性与政策的可接受度：高考政策制定和执行的核心变量

约翰·克莱顿·托马斯（John Clayton Thomas）教授在分析公共政策制定和执行时引入两个核心变量：其一是政策本身的质量；其二是政策的可接受度。<sup>[2]</sup>变量一意味着政策的科学性，即政策是否指向了客观存在着的现实问题，政策所规定的各项举措是否符合客观事物的发展规律。变量二可看作是代表不同利益的社会群体对政策的认可状态及其一致性程度，这是公共政策制定和执行不可忽视的社会资本。一项改革要成功，既取决于政策本身的科学性，亦取决于外在的社会条件和受众的接受程度。

### 1. 科学性更高：高考标准分制度兼顾公平与效率

分数合成及其报告是考试工作流程的最后一个环节。原始分是最简单的分数合成和报告方式，试卷上每一道题的原始得分直接相加即可得到该科目的原始得分。对于涉及多个科目的考试，只需把考生各科成绩的原始分数直接

相加，就生成了这位考生的原始总分。虽然算法非常简单，但原始分在多科目考试时存在内生性局限，其中最突出的问题是由于难度和区分度不同，不同科目的原始分数存在难以回避的等值、可比、可加等现实问题。针对原始分的这一内生性局限，测量学界提出了标准分的解决方案。标准分是根据教育统计与测量学的原理，按照一定规则将原始分数转化为具有相同意义、相同单位和共同参照点并能表明特定考生的考试分数在总体中排名位次的相对分数。相比于原始分，标准分在科学性上具有明显的技术优势，解决了不同科目分数之间的等值、可比、可加问题，同时更加准确地反映出每位考生在全体考生群体中的相对位次，这为考试结果的科学使用夯实了根基，对考试公平的改进也是显而易见的。

从应用的广泛性来看，标准分亦是国际重要考试项目主流的分数报告形式，如 TOEFL（非英语母语者英语能力测试）、SAT 或 ACT（服务于美国大学本科招生的标准化考试），以及经合组织的 PISA（面向 15 岁青少年的国际测评项目），等等。当前时期，有着广泛国际影响的大规模考试在成绩报告上几乎找不到使用原始分的实例。可以说，标准分制度符合教育考试的发展规律，是公认的公平与效率兼顾的制度形式。在标准分制度规划、设计与执行之初，教育主管部门对其政策科学性的认可直接推动了部分省份探索试点高考标准分制度。在标准分制度的演变进程中，海南省的做法可圈可点，该省自建省以来一直使用并在 2017 年启动的高考综合改革中继续坚持标准分制度，从 2020 年起，该省统一高考的语文、数学、外语 3 个科目成绩和考生自主选择的学业水平等级性考试的 3 个科目成绩以及高校录取总成绩，均以标准分呈现。该省教育主管部门明确表示，以标准分报告成绩是国际通用做法，既有科学依据也符合该省实际，同时还可纠正部分考生在选考等级性考试科目时“避难选易”的功利性问题，应继续坚持。由于等级性考试选考科目的群体不同，为确保标准分的转换更加科学合理，海南将根据各学科的质量因素对考生成绩先行校

准和等值处理后再进行标准分转换。<sup>[3]</sup>

2. 可接受度低：受众反对及治理风险使之难以为继

高考是大规模的国家教育考试，与千百万民众的切身利益戚戚相关。将单科原始分直接相加计算出原始总分，继而划线和录取的做法在我国各级各类升学考试中承袭已久，尽管标准分制度采用的算法使分数的评定更加精准，进而使仅根据考生高考分数排序的高校招生更加公平，但相比而言，原始分的计算更简单，也基本满足了社会各界对高考制度形式公平的独特要求，因此，原始分似乎拥有更高的信任指数和更广泛的民意基础。伴随着标准分制度的实施，来自考生、家长、高中学校和社会的负面反馈、质疑与反对的意见时而涌现。教育主管部门及其附设的考试管理机构是我国高考政策的供给主体，在高考制度框架中始终发挥着主导作用。尽管从测量学的内在原理及标准分在国际上的广泛应用状况看，相比于原始分，标准分制度更加兼顾公平与效率，但当标准分逐渐从一个专业议题演变为“社会问题”时，主管部门越来越倾向于放弃标准分以直接消解各种潜在的治理风险。在经历了二十多年试验之后，这项更加科学、更具实质公平、让社会受益更大的制度形式最终还是难以为继，绝大多数省份相继废止了标准分制度，计分方式基本上退回到原点。

3. 标准分制度沿革的治理困境：制度科学性与受众接受度之间的断裂

标准分制度通过高考计分方式的技术改良保障并增进了考试公平，是考生个体利益得到进一步保障而非牺牲个体利益的制度，却不被受众所接受，这体现了高考制度改革曾经面临的治理困境：制度科学性与受众接受度之间的断裂。标准分制度的后续走向清晰地揭示出：高考制度改革即便仅仅只是技术层面上的改良，也必须以受众的普遍接受为基础，受众的接受状况对于高考制度改革的顺利实施特别是可持续发展至关重要。值得深入探究的是，标准分制度追求公平的目标没有改变，高校招生录取也依然秉持“分数面前人人平等”总分大排序

的分数应用方式，改变的只是计分方式，受众为什么不接受这一在国际上有着广泛共识的计分方式？本文认为，受众的不接受应该是多种因素合力造成的，根源之一在于，当年高考结束后，制度供给主体对高考信息的过度保密一定程度上导致高考关键数据及必要解释的缺失，信息严重不对称为广大考生留下“暗箱操作”“不公平”的想象空间，加之基于数据的理论研究和舆情引导匮乏，由此使得考生、家长、高中学校及社会很难真正接受并完全配合这项制度。

二、不接受的根源之一：关键数据及解释的缺失影响对标准分制度的普遍接受

1. 考生对高考分数的自我“确认”难以展开  
我国有着独特的考试文化，广大学子十多年寒窗的首要目的在于争取优质高等教育的入学机会并以此为基础实现个人的高阶层发展。优质高等教育是稀缺的，要获得稀缺的入学机会，考试分数成为所有阶层子女一致追求的共同目标，是最有价值的“王牌”资本。<sup>[4]</sup>之所以称之为“王牌”资本，是因为我国高校最主流的招生渠道——统考招生总体上严格按照考生高考成绩的高低进行投档和录取。高校实际的招生是“唯分取人”，这种招生机制下，考生、家长、高中学校和教师们的核心关注自然聚焦到分数的竞争上，“分分计较、每分必究”。高考一结束，考生都会自发地依照各科目的参考答案对自己记忆中的作答情况进行成绩预估；高考分数公布后，考生还会根据估分与最终得分之间的差距判断官方评卷过程是否存在偏误的可能，对最终得分的“自我确认”过程是直接影响考生系统性信任的枢纽环节。笔者作为标准分制度的亲历考生，当年，所在省份面向考生发布的高考成绩只有各科标准分与高考总分标准分，除此之外，没有任何其他数据以及解释说明。在标准分计分制度下，考生依然习惯于预先估分，但对最终公布的标准分得分进行确认所必备的高考单科原始得分、本科目高考的原始平均得分、标准差以及单科标准分合成高考总分标准分的算法等关键考试数据、算

法参数及必要的解释说明是缺失的,对个体考生而言,对最终公布的标准分的确认难以展开。在官方阅卷偶尔存在偏误的情况下,自身考试分数“确认”环节的强制抽离势必给广大考生留下“暗箱操作”“不公平”的想象空间,猜忌心理某种程度上助长了受众对标准分计分方式和考试成绩的误读误解。

需要说明的是,国际上的一些大规模考试有的也不面向考生个体提供考生个人的原始得分等考试数据,<sup>[5]</sup>但这些标准化考试如 SAT 或 ACT,其利害程度远远不及我国的高考,SAT 或 ACT 分数仅仅只是美国大学招生的参考依据,既非最重要,更非唯一依据。<sup>[6]</sup>目前国际上最具代表性和影响力的行业标准《教育与心理测验标准》明确规定:“不能仅凭单一的测试分数去做出那些可能对学生产生重要影响的决定,还必须依据其他相关信息。”<sup>[7]</sup>在我国,高考对考生个人的终身发展具有重大影响。高考之于考生的利害程度加剧了考生应考的心理压力,也进一步强化了“分分计较、每分必究”的独特考试文化。

## 2. 基于数据的舆情引导和理论研究匮乏导致对标准分的误读误用

有观点认为标准分制度不被接受的主要原因在于标准分转化的专业壁垒及由此带来的成绩理解上的困难。显而易见,相比原始分,对标准分的准确理解确实需要受众具有一定的测评素养。但标准分的内在原理及其计算对于学习过正态分布的高中生并非难到不可理解。即便是考生和家长们起初有疑惑,考试管理机构与相关科研机构也有充裕的时间进行立足考试数据的舆情引导和理论研究。然而,各级考试管理机构在高考关键数据、算法参数,以及解释说明的适当公布、提供参与研讨平台等方面多有欠缺。与之密切关联,学术界关于高考标准分这一关键技术变革的理论研究及大众媒体舆情层面的宣传引导是匮乏的。笔者曾以“高考”+“标准分”为主题词<sup>①</sup>在中国知网的期刊数据库中检索,自 1985 年开始试点迄今三十年来,全国各类期刊刊发的高考标准分主题的论文仅检索出 54 篇,其中 CSSCI 期刊的论文

仅有 9 篇,而且多停留在一般性地阐述标准分制度的科学性上,立足本土的实证研究几近空白。标准分制度广泛应用的理论储备匮乏,基础理论研究的薄弱使大众媒体专业视角的分析、有效的舆情引导成了无源之水,更广泛的社会群体对标准分科学性的系统信任和普遍接受自然大打折扣。

考试信息不充分还直接影响公众参与考试治理的质量和考试制度改革的实施成效。在我国,考试大纲、原始分数和全省位次基本上就是考生、学校和社会能够获得的全部考试信息。长期以来,高考数据没有得到更进一步的开发应用,造成权威数据资源的浪费。在信息公开水平较低的情况下,公众的接受状况和参与能力同样处于低水平。时至今日,由于缺乏考试的效度验证和对考试结果使用的必要解释,我国基础教育领域存在的诸多顽瘴痼疾,如“唯分数论”等都源于对测试结果的误读误用。实践表明,公众普遍的不接受和低水平的参与势必对高考制度改革带来消极的影响。

## 3. 更甚于国家秘密的保密性导致高考关键数据及必要解释的缺失

关键数据的适度公开及必要的解释说明为受众的理解和接受提供权威指南,有助于强化受众的可接受度。然而,标准分二十余年的制度沿革显示,作为制度供给主体的政府及其附设的考试管理机构似乎未曾就如何有效增进受众的理解和接受状况做出更多的持续的努力。我国高考领域,出于考试安全的顾虑及对“额外”工作负担的规避,制度供给主体对于考试过程中的工作信息长期以来“不敢开放、不愿开放、不会开放”。毫无疑问,高考考试结束前,高考试题试卷等作为国家绝密级材料,必须在符合国家保密标准的设施、设备中保管和运送,所有接触试卷的人员都实行封闭式管理,未经原定密机关、单位或其上级机关批准,不得复制和摘抄。但总体而言,与国家秘密相比,如高考试卷这类所谓工作秘密信息,在保密时限、范围、事项上表现出更甚于前者的保密性。<sup>[8]</sup>高考领域定密的标准过于严格,范围和边界过于宽泛。比如,高考考试结束后如果考生对阅卷

工作有质疑，向当地教育考试院或高招办申请或走法律程序申诉复查本人答题卡等具体作答信息时，主管部门往往以“高考试卷属于国家秘密”为由拒绝考生的请求。目前全国的所有省份，当考生提出试卷复查时均不允许考生直接接触答卷。<sup>[9]</sup>复查时是否允许考生直接接触答卷还需进一步论证，但当年高考结束后，考试管理机构面向考生单单只提供一个标准分，分数转化过程中的关键数据、算法及参数、必要的解释说明是缺失的，信息不对称，考生、家长、高中学校及社会很难真正接受并完全配合这项制度。

对教育系统而言，信任是教育政策制定和实施、教育创新以及与社会和经济互动的一个重要元素，同时对于增加多层教育治理系统中利益相关者之间的合作、有效问责、专业化水平以及应对复杂性和减少信息不对称都是极为重要的。<sup>[10]</sup>高考改革的推进本应是一个逐步适应和逐步完善的过程，令人遗憾的是，高考改革相对缺乏增进受众接受配套机制建设的战略思维。制度供给主体未能在增进受众的认知和理解方面及时跟进，相关配套机制长期滞后，在这样的社会基础和国情环境下，受众对标准分制度的不接受自然日益加剧，潜在的治理风险进一步显现，政策质量更优的标准分制度在消极反馈面前显得不堪一击，最终悄然退出大多数试点省份高考历史的舞台。从制度建设的角度看，高考改革进程中对受众可接受度及其重要性的预备不足会导致更科学的制度供给无功而返，并以路径依赖方式影响后续的改革进程，使改革的探索闭锁在原初的制度结构中，考试制度改革面临停滞不前甚至被扭曲的风险，这恐怕是标准分制度被废止要付出的更大代价。

### 三、高考治理困境的突破：持续关注并强化受众对改革举措的接受程度

在我国，高考制度的研究可谓方兴未艾，但似乎忽视了一个根本性的问题：单纯以提升制度客观的科学性来推进高考制度综合改革是远远不够的。轰轰烈烈的标准分制度的兴起与

废止昭示着高考改革曾经面临的治理困境——制度科学性与受众接受度的断裂。这一困境的突破方向在于制度供给主体应持续关注并强化受众对改革举措的接受程度，提升高考治理的服务性。这固然取决于多方面的努力，其中一个值得关注的努力方向在于考试管理机构以开放的姿态，对关键数据、算法及参数信息进行脱敏处理后，向考生及专门考试研究机构适度公开，并做出必要的解释和说明。

#### 1. 更新观念，认可高考关键数据信息适度公开及做出必要解释的意义

高考结束后，相关信息不予公开，基本依据在于制度供给主体认定信息公开具有一定危害性。对危害性的顾虑首先在于考试安全，考试安全是大规模高利害考试存续发展的前提和基础，但是，考试安全并不意味着考试过程中的所有数据都是“秘密”，否则国家没有必要实行高考招生工作“阳光工程”。国际主要测试机构和协会研制的关于考试质量与公平性的标准均提倡在保障考试安全的基础上，适度公开考试信息，内容大体包含对成绩报告和测试使用的阐述等。比如《教育与心理测试标准》明确倡导在可能限度内社会各界应能得到必要的资料，以便参与辩论这些政策问题的各方都拥有充分的信息。<sup>[11]</sup>当考试分数信息公布时，考试机构有责任向受众提供适当的解释，解释中包含考试资源、原理算法、实证依据等，同时描述该测试的局限性等。当向个体报告群体统计信息时，应报告考试的信度效度证据。<sup>[12]</sup>

信息的适度公开及必要的解释说明是让测试使用者、考生、家长、教师、学校、教育政策研究者、大众媒体等心悦诚服地接受，引导如何正确地理解和应用考试，进而全面提升整个社会的测试素养，从而共同承担维护测试质量和公平协作责任的必要前提。在当今信息社会与大数据时代，让数据流动起来，使封闭的、分散的、沉睡的数据活跃起来，从而激发技术流、资源流、人才流，进而实现“用数据说话、用数据决策、用数据管理、用数据创新”。<sup>[13]</sup>已有研究建议，公开除考试大纲和原始分数之外的如成绩统计分析、试卷质量分析、效度论证

报告、教学反馈等更多信息。<sup>[14]</sup> 考虑到高考信息的特殊性,有研究建议将高考信息置于《政府信息公开条例》设定的解释框架中,重新考量并实现有选择的、有限制的不予公开。<sup>[15]</sup> 总之,数据信息是一种宝贵的社会资源,是切实推进高考治理现代化的关键载体之一。教育主管部门应更新观念,认可高考关键数据、算法及参数信息脱敏处理后适度公开及做出必要解释的意义,在可能限度内,使高考考试数据可获取、可利用,真正顺应大数据时代考试治理现代化的发展趋势。

## 2. 逐步践行,高考的关键数据、算法及参数信息脱敏后适度公开

高考结束后,当年高考的关键数据、算法及参数信息脱敏后适度公开是一个较为新颖的提法,考虑到我国的基础情况和探索条件,可以预测,现阶段这一方向的努力会遇到不少困难和挑战。应对挑战,不仅需要理论界对高考考试安全中的认知观念尤其是长期以来的认知误区进行澄清,还需要在顶层设计上加强规划部署,在保障考试安全及其他公共利益不受损害的前提下,逐步推动和促进考试数据信息公开的法治化、标准化和规范化建设。

第一,针对考试数据信息的公开及开发应用制定或完善相关法律或部门规章。加快制定出台《考试法》,或借鉴发达国家及国际成熟的关于考试数据及解释说明等方面的行业标准,使相关工作合法合规、有章可循。

第二,制定高考数据适度公开及开发应用的执行标准,确立不同开放模式的实施程序,并辅以详细的操作规范和具体指导。这需要协调好两对矛盾:一是数据开放与保守秘密之间的矛盾;二是数据开放后的社会需求与考试机构有限人力、物力之间的矛盾。与欧美国家相比,我国的高考环境更为严峻,保障考试数据安全更是高考的重中之重。建议采用多种途径保障数据安全,对信息管理主体、公布的数据类别、时间节点、公布渠道、公布范围等进行规定,可考虑对考试数据进行保密等级划分,灵活采用各种脱敏方法对数据进行处理。另外,确立考试管理机构 and 人员、考生和高考结果使

用者等各利益相关方在关键数据获取、成绩解释和应用方面的权利和责任。在此基础上设立白名单机制和申请审核发布机制。白名单机制是通过一定的审核标准,将一些信誉好、科研能力强的学术机构纳入高考数据开发应用的合作范围。申请审核发布机制按照“申请—审核批准—数据使用—成果审核—发布”步骤开展,在政府部门的监督和管理下,允许合作机构在签订保密和授权协议后,立足高考数据进行学术研究,最终的研究成果经报备和审核后发表。无论采用何种数据公开及开发模式,都应使数据使用者明确数据使用的保密义务和责任,避免数据的误用滥用。

## 3. 转变职能,组织开展基于高考数据信息的理论研究与舆情引导

当前社会背景下,信任缺失正成为影响我国国家治理的突出难题。<sup>[16]</sup> 标准分制度的变迁揭示出,高考制度改革中“科学”是必须的,但要“看得见”;“公平”也是必须的,但同样要能“感受到”。增进受众对高考改革的认知、理解、信任和接受,具有全局性和深远性意义。高考政策的制定主体与负责高考考试与招生工作的教育考试院承担着维护制度核心价值的公共责任,同时肩负着维护考试公平与效率,以及推进公众普遍接受和理性参与的双重责任。一项负责任、有能力的考试制度不应单凭公众在改革起步阶段的不接受而顺势停止更优改革举措的持续供给,关键在于积极探索增进受众可接受度的配套机制,使更高科学性的制度供给与更广泛受众的理解接受有机结合起来。

考生、家长、高中学校、教师及社交媒体等是“看得见”“感受到”进程中的关键合作伙伴,是高考治理现代化的参与力量。转变职能,提升高考治理的服务性,创设能增进受众普遍接受的配套机制是高考治理困境值得期待的突破方向,也是本轮高考制度综合改革高水平实施值得关注的保障思路。高考关键数据、算法及参数信息的适度公开亦可为学术界的理论研究打开一扇窗,为数据的深度挖掘及基于实测数据的实证研究提供必要支持。高考制度供给主体还可考虑组织开展基于高考数据的学术研

究, 以及面向更广泛社会群体的学术传播、价值引领与舆情引导。

高考制度是公众普遍关注的热点和社会利益的汇聚点。标准分制度的发展沿革为我们提供了深刻的启示, 高考改革的顺利推进在关注制度科学性的同时, 还必须关注受众的接受状况, 受众的接受程度是影响高考改革进程的先决条件之一。持续关注并通过配套机制强化受众对特定改革举措的接受程度有助于恰如其分地回应公众似是而非的朴素认知以及与公平及效率相抵触的非理性参与, 有助于促进新型信任关系的形成, 进而提升高考治理的现代化水平。

注释:

① 以“高考”+“标准分”为主题词进行检索, 意味着检索出的论文或在“篇名”、或在“摘要”、或在“关键词”中含有“标准分”和“高考”。检索时间为2019年3月21日。

参考文献:

[1] 章建石. 一项公平与效率兼备的高考改革为什么难以为继? ——标准分制度的变迁及其折射的治理困境 [J]. 北京师范大学学报(社会科学版), 2016, (1): 31—41.

[2] 约翰·克莱顿·托马斯. 公共决策中的公民参与: 公共管理者的新技能与新策略 [M]. 孙柏瑛, 等译. 北京: 中国人民大学出版社, 2005. 前言: 4.

[3] 海南省教育厅网站. 海南省高考综合改革政策解读 [EB/OL]. <http://edu.hainan.gov.cn/edu/zxjd/201803/f74abc094b6e4cd08e126c8e8e944a43.shtml>, 2019-03-21.

[4] 王金娜. 高考统考科目的“文科偏向”

与隐形教育不公平——基于场域—文化资本的视角 [J]. 教育发展研究, 2016, (20): 8—14.

[5] 程蒙蒙. 美国大学入学考试中心和培生教育测量成绩报告 [J]. 中国考试, 2008, (7): 56—61.

[6] NACAC. 2017 State of College Admission [EB/OL]. <https://www.nacacnet.org/news--publications/publications/state-of-college-admission/2019-03-21>.

[7] [11] [12] American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). Standards for educational and psychological testing [M]. Washington, DC: AERA, 2014. 198. 1. 119—121.

[8] [15] 梁艺. 工作秘密不予公开的合法性反思——从杨婷婷高考试卷“调包案”切入 [J]. 时代法学, 2015, (2): 48—55.

[9] 王俊秀, 吕昀蔓. 我的高考试卷为何成了国家机密 [N]. 中国青年报, 2011-06-24 (06).

[10] Cerna, L. Trust: What it is and Why it Matters for Governance and Education [EB/OL]. <http://dx.doi.org/10.1787/5jxswcg0t6wl-en>.

[13] 全国首届教育实证研究论坛. 大力开放教育数据的倡议书 [N]. 光明日报, 2015-10-27 (14).

[14] 程蒙蒙. 国际知名考试质量与公平性标准的比较研究与启示 [J]. 教育理论与实践, 2015, (6): 21—25.

[16] 杨中超. 教育治理: OECD 成员国的经验与启示 [J]. 外国教育研究, 2017, (2): 107—116.

(责任编辑 李威)

## The Gap between Scientificity and Acceptability of Gaokao and the Breakthrough of the Governance Dilemma --Taking the Implementation of the Standard Score System as an Example

Zhang Huijie

**Abstract:** This paper selects two variables of policy scientificity and policy acceptability to analyze the implementation of the standard score system in China. It is found that the Gaokao reform in China is confronted with the dilemma of system scientificity and the acceptability of the audience. The reform of Gaokao system must be based on the universal acceptance of the audience, or even the improvement of the technical level will be difficult to follow. In China, the breakthrough of the predicament of Gaokao is to pay attention to and strengthen the audience's acceptance of the specific reform measures. This paper discusses the possibility of the appropriate disclosure of the examination data and the explanation in improving the acceptability of the audience, as well as the challenges in the field of Gaokao in China. The study holds that it is necessary to clarify the misunderstanding of the data information concerning the security of Gaokao, furthermore, it is necessary to strengthen the planning and deployment of information disclosure on the top level and organize the theoretical research and public opinion guide based on the data information.

**Key words:** Gaokao system; Standard score system; Scientific; Acceptability; The examination governance dilemma

(上接第 8 页)

## The Prescriptive and Guiding Significance for the “Double First-class” Construction of Xi Jinping's Important Exposition on Education

Zhu Xinkai

**Abstract:** Through “Nine Persistences”, the National Education Conference summarized Xi Jinping's important exposition on education since the 18th National Congress of the Communist Party of China, providing the fundamental follow-up for us to carry out education work. “Nine Persistence” profoundly explains a series of directional, fundamental, overall and strategic issues of education, constitutes a complete and scientific ideological and theoretical system of education, is a guide for the development of education in the new era and the construction of “double first-class”, and clearly defines the leadership, fundamental tasks, strategic position, higher education direction, school-running path, value orientation, development motivation and so on, to promote the construction of “double first-class”, providing a fundamental follow for the construction of “double first-class” in the new era.

**Key words:** Xi Jinping's important exposition on education; National education conference; Double first-class; Nine persistences