

深化教育考试评价改革 推进中国式考试现代化

关丹丹

(教育部教育考试院, 北京 100084)

摘要: 中国式现代化道路需要教育推动, 考试作为最重要的教育评价手段之一, 需要全面深化改革, 以中国式考试现代化服务于教育评价改革。考试内容改革从考试起点出发, 目标是让考生做有意义的事情; 考试评价改革是从考试功能出发, 目标是让考生所做事情的结果得到科学、准确的阐释, 并使这些结果得到充分、有效的应用。考试评价工作通常包括评价考试、评价考生、评价教育教学3个主要方向, 从与考试相关的直接、间接数据出发, 以评价结果反馈促进考试质量的提升, 促进考生个性化发展, 促进教育教学精准改进。为推进中国式考试现代化, 深化考试评价改革的路径是: 一要强化科技赋能数字化考试评价, 二要建立有效的评价反馈与持续改进机制, 三要创新考试评价的理论与技术, 四要提升多元评价主体和社会公众的评价素养。

关键词: 考试评价; 教育评价改革; 中国式现代化; 考试现代化

【中图分类号】G405

【文献标识码】A

【文章编号】1005-8427(2023)01-0015-6

DOI: 10.19360/j.cnki.11-3303/g4.2023.01.004

党的二十大报告指出, 要以中国式现代化全面推进中华民族伟大复兴^[1]。中国式教育现代化是中国共产党领导、造福于全体人民的中国特色社会主义教育现代化, 是中国式现代化的重要组成部分, 是实现中国式现代化的基础工程和核心保障^[2]。考试作为最重要的教育评价手段之一, 是中国式教育现代化的关键组成部分, 是人才选拔的重要渠道。2020年, 中共中央、国务院发布的《深化新时代教育评价改革总体方案》(以下简称《总体方案》)引领我国教育逐步发生整体格局性变化。基于《总体方案》提出的指导思想、主要原则、改革目标, 有必要对考试的评价活动及其本质进行再认识, 通过不断深化考试评价改革,

提升高考在教育评价改革中的贡献度, 为中国式教育现代化贡献考试改革的力量。

1 概念: 对考试评价的再认知

教育考试、教育评价、考试评价、评价考试, 这些名词即使在专业领域也很容易混淆; 因此, 很多教育工作者在实践工作中往往将其混用, 并未过多注意概念之间的差异和不同导向。为了厘清本文所指考试评价的内涵, 首先需要重新审视一些基本概念。

教育考试是教育教学的一种伴生产品或手段, 是一种基于教育测量理论进行的实践活动。理论上, 学生的知识、能力或素养是人为假设存

收稿日期: 2022-11-03

基金项目: 教育部教育考试院“十四五”规划支撑专项课题“我国考试从业人员评价素养研究”(NEEA2021049)

作者简介: 关丹丹(1980—), 女, 教育部教育考试院评价处处长。

在、且无法直接观察到的变量(心理学称之为潜变量)。考试通过给考生提供一系列的任务样本,然后收集考生的作答反应,从而推测其知识水平、能力及素养的变量值^[3]。按照此定义,可以认为考试是通过设计用以获得考生特定行为样本的测量工具,按照确定的规则、过程和方法,对考生的特定心理特征进行量化的估计。例如某考试中甲的分数是80,乙的分数是75,据此我们有较大把握推断甲的水平高于乙。

与考试不同,教育评价是根据一定的教育价值观或教育目标,运用可行的科学手段,通过系统收集和分析整理信息资料,对教育活动、教育过程和教育结果等进行价值判断^[4]。教育评价具备完整的理论基础和指导思想,更注重自身规范的科学研究模式和理论自洽,利用包括但又不限于考试数据的一系列相关证据集合,综合使用多种方法对学生个体或群体作出正面或负面的结论。同样以上述2个考生为例,如果他们来自不同地域、处于不同的家庭社会经济地位、拥有不同的学习资源,则不能推断甲一定优于乙。来自西部偏远山区且学校资源环境不佳的考生,如果获得与来自东部沿海城市且教育资源充足的考生相同的分数,其付出的努力和代价是完全不同的。

因此,虽然考试和评价之间关系密切,但考试并不等同于评价。考试往往是评价的重要手段,但仅仅基于考试结果进行的评价并不一定准确。考试结果常常与数字关联,表现为客观的;评价结果往往与某种结论关联,反映评价者的价值观,因而在某种程度上是主观的。只有用考试结果作出某种决策时,考试才与评价挂钩。教育考试只是教育评价的一种工具或手段,教育评价才是背后的终极目的^[5]。很多情况下,考试和评价的应用并不限于教育内部。以高校招生为例,这类评价事实上已经成为社会治理的一种手段。在高风险的社会治理领域,概念间的微小差别甚

至可能引发严重的社会后果,因此对概念的深究绝非简单的文字游戏。

考试评价的概念鲜见于教育领域相关文献,此前并无严谨的定义。考试具有甄别、选拔、导向、评价等多元功能,这些功能依靠分数背后丰富的信息实现。本文提出考试评价概念的目的主要有2个:第一,国家考试等大规模教育考试的首要功能是服务于招生,因此必须注意考试和评价所需要的不同推论过程和注意事项,既要充分地利用考试结果,也要警惕和防止对考试分数的误用和滥用;第二,花费大量社会成本组织的考试,其结果蕴藏着对教育极其珍贵的信息,应该服务于各级各类教育评价工作,特别是在新时代教育评价改革进程中,考试评价已经成为教育改革与发展的龙头和牛鼻子。

发挥考试服务于教育评价功能的前提是,考试本身足够科学和准确;因此,考试评价首先是对考试工具的评价,然后才能用于教育教学、教育管理、学生成长和决策咨询等领域。在专业化考试机构中,考试评价工作通常包括评价考试、评价考生、评价教育教学3个主要方向。“评价不是为了证明,而是为了改进。”^[6]第三代教育评价的这个理念同样适用于考试评价。从与考试相关的直接、间接数据出发,以评价结果反馈促进考试质量的提升,促进考生个性化发展,促进教育教学精准改进,是充分发挥考试正向功能的应有之义。

2 创新:评价理念的转变

从国际视角看,在经历了“对学习的评价”(assessment of learning)、“为了学习的评价”(assessment for learning)2个阶段后,“评价即学习”(assessment as learning)的理念逐步成为当前国际教育改革与发展的新共识。“评价即学习”强调评价本身就是一种学习过程,即评学融合。例

如,美国各州普遍在推进 K-12 评价体系的创新设计,目的是使利益相关方能够更好地为每个儿童提供公平、优质的教育。表 1 为美国多家组织联合提出的美国教育评价现状及其改革愿景^[7]。由此不难看出,美国对促进评价创新的观察结论与我国“双减”政策直指的问题具有高度一致性。应试教育、考试高压、考试数据对教和学的指导不充分等,已经成为当今世界各国教育评价面临的共性问题。

《总体方案》是我国在构建富有时代特征、彰显中国特色、体现世界水平的教育体系中迈出的坚定而有力的一步。评价作为教育发展的“指挥棒”,一直是教育改革的关键问题。教育评价和评价节点不仅在中国,在全球范围看也具有深刻意义。在“评学融合”的大趋势下,重新思考评价对改进教学和学习的作用,创造更加平衡、融合的评价体系以更好地支持教和学,同时又能为政策和资源配置决策提供必要信息,这已经成为各国教育评价改革的共同目标。

国际教育评价理念的转换与更新,为我们科学开展考试评价活动设计提供了有益借鉴。我国教育评价改革政策对考试提出了探索立足“四个评价”、服务“五类主体”^[8]的新要求。无论是评价考试、评价考生、评价教育教学,都是基于评价过程、评价数据的不断反思和改进,促进考试、考生、教学的不断发展与完善。在新时代考试评价改革实践中,为了更好地实现“评价即学习”的目

标,需要重点关注以下 2 个方面:一是考试评价应以诊断和发展为主要目标,评价结果不仅要回应考试内容改革中的某些关键问题,促进考试更加科学、公平,还要回应教育发展过程中亟待解决的问题,引导学生全面健康发展;二是考试评价应从多角度、跨学科视角出发,多措并举实现对学生个体和群体的全面评价,统筹量化与质性 2 类评价方式的合理综合运用。

3 思考:深化考试评价改革,推进中国式考试现代化

深化新时代教育评价改革是提高教育治理能力和水平,加快推进中国式教育现代化的重要制度设计。立足我国大规模教育考试特点,改进结果评价,强化过程评价,探索增值评价,健全综合评价,强化考试服务“五类主体”的效能,是进一步提升教育考试服务科教兴国战略效能的重要要求。深化教育评价改革,要求作为评价工具或手段的考试也要改革,这不仅包括考试内容改革,也包括考试评价改革^[8]。从这个意义上说,考试评价改革是教育评价改革的次级概念。考试评价改革与考试内容改革既有区别又有联系:考试内容改革从考试起点出发,目标是让考生做有意义的事情;考试评价改革是从考试功能出发,目标是让考生所做事情的结果得到科学、准确的阐释,并使这些结果得到充分、有效的应用;考试评价改革既需要考试内容改革,也服务于考试内

表 1 美国多家组织总结的促进评价创新的关键主题^[7]

教育评价的共性问题	教育评价的未来愿景
囿于考试准备使课程变得窄化	既要确保在基础性技能上的公平教学,也要支持更深入、更个性化的学习
各州层面的考试数据对教学指导的可操作性不够,来自课堂的数据不受重视	平衡整合州层面的评价探索与丰富的评价方法,为学生、教师和监护人提供可操作的数据
教学时间因为准备考试和参加考试被挤占,学生考试压力大	嵌入教学的评价要为教和学提供支持,学生要参与更有意义的挑战以衡量其在学习中的位置
在评价学校时仅使用标准化考试结果,没有额外和更广泛的指标,往往使社区名声受损	评价体系的设计要适应文化需要、维持文化发展,确保每个学生都能在经历的评价中看到自己

容改革,考试内容改革的成效通过考试评价改革得到呈现和肯定。考试内容和考试评价共同构成具有中国特色的教育考试评价体系,是中国式考试现代化的重要组成部分,共同服务于教育评价改革,服务于中国式教育现代化建设。

3.1 强化科技赋能数字化考试评价

评价必然是个性化的,只有数字化才能做到个性化。大数据、人工智能、区块链等新兴技术快速融入教育领域,为推动考试评价改革创新提供了条件。深化考试评价改革,要深刻认识教育数字化战略行动在教育改革发展中的战略价值,充分重视考试评价改革中新信息技术的作用。落实国家教育数字化战略行动,必须有序推动考试评价业务的数字化、智能化升级。一是考试评价必须建立用数据说话、用数据决策、用数据管理、用数据创新的理念,实现大数据与考试评价工作的深度融合。大数据在数字考试评价中的核心价值主要体现在通过对数据的挖掘研判考试,研判教育教学问题,预测考生群体变化,从而实现考试治理科学化、教育评价服务精细化。二是数字考试评价必须具备坚实的资源保障,除考试本身产生的数据资源外,还包括与系统评价相关的非考试数据资源。数字考试评价还要实现2个转变,即从定量分析向定量与定性分析相结合的转变,从传统数据分析向依托大数据与典型调查分析相结合的转变。三是考试评价要通过信息技术形成可视化表达。在引入数据分析模型与方法进行个性化数据挖掘的基础上,能够形成三维动态展示、交互式展示,从而直观反映考试质量,以及考生个体、群体的发展情况和变化趋势,为各级各类决策者提供考试乃至教育的基本数据、过程数据、发展数据、增值数据、综合数据等,并辅助决策。总之,科技赋能是未来考试评价全“数”前进的核心引擎,也是构建教育考试新发展格局的重要力量。

3.2 建立有效的评价反馈与持续改进机制

评价结果的有效反馈与持续改进机制直接关系到评价功能的发挥;因此,深化考试评价改革既要抓好结果评价改进,关注表现特征,强调发展水平,更要抓好考试结果的反馈和使用。为贯彻落实《总体方案》,考试评价在结果反馈与使用上必须坚决落实“十个不得,一个严禁”。只有考试评价结果得到科学、合理的解释和使用,评价的诊断、改进、发展、激励功能才能得到有效发挥,教育才能从“育分”走向“育人”,学生才能全面而有个性地发展。为建立评价结果的有效反馈与持续改进机制,需要注意以下3个方面:一要增强考试分数的可解释性,充分挖掘考试及其相关数据蕴含的丰富信息。以高考为例,中国高考评价体系是新时代高考内容改革的基础工程、理论支撑和实践指南,也是考试评价的准绳和量尺。二要使用科学的评价标准构建评价指标,对评价结果进行分级、分类反馈。在立足“四个评价”、服务“五类主体”上破立结合,通过不断探索和实践,面向不同对象提供有针对性的评价报告。三要建立“评价-反馈-改进”的良性循环运行机制,迅速、准确地向不同主体反馈评价结果,并帮助各类主体正确地理解和使用评价结果,形成以持续改进为核心的评价机制。最终从关注分数转向聚焦个体和群体发展,实现学生成长、国家选才和社会公平的有机统一^[8]。

3.3 创新考试评价理论与技术

中国是考试的故乡,考试评价改革的发展有自身的特殊性,这与中国的传统文化有关。深化考试评价改革也需要推出具有中国特色的考试评价理论、方法和技术。纵观各国招生考试,大学招生制度决定考试制度,考试功能决定考试内容、形式及采用的教育测量理论和考试评价技术。由于我国大规模教育考试尤其是高考,具有超高影响力、超高利害性、超大规模、超高效率的

特点,西方教育和心理测量理论的一些基本范式并不完全适用于我国。此外,我国考试评价改革是新时代面临的新任务,是为应对新问题而提出的改革任务。时代课题是理论创新的驱动力,理论创新又为回答新时代考试评价课题开辟更加广阔的视野和有效路径。尽管学术界已有很多成熟的教育评价技术和方法,但如何更好地与我国国情结合、与当前的教育评价政策结合、与我国高考的现实需求结合,还需要在借鉴国外经验的基础上,以创新发展理念为引领,从学理研究和实践探索2个层面,探索具有中国特色的考试评价理论与方法体系,也为世界教育测量与评价理论的丰富与发展贡献中国智慧。

3.4 提升多元评价主体和社会公众的评价素养

深化考试评价改革需要实施面向多元评价主体和社会公众的评价素养提升计划。第一,着力提升考试从业人员的评价素养。根据高考综合改革需要和高考内容改革要求,全国命题人员、考务管理人员、考试研究人员将以前所未有的速度日益壮大,提升他们的评价素养既是做好各项考试的前提条件,也是落实《总体方案》关于加强国家教育考试队伍建设的必然要求。考试从业人员的评价素养至少包括5个方面:一是具有较强的政治政策意识,二是秉持科学的评价信念和积极的评价态度,三是拥有国际视野并掌握丰富的评价理论和评价过程知识,四是具备较强的评价实操、考试研究等实践技能,五是恪守安全保密、公平公正的职业操守。第二,重视教师教育评价能力建设,不断提升教师群体的评价素养。面对新高考、新课标、新课程、新教材“四改并行”的教育新局面,要为教师搭建学习发展平台,帮助教师更加深刻地理解评价并提升其评价素养,在日常教学中进一步理顺教学与评价的关系,进而带动全社会更加理性、科学地认识教育评价^[9]。第三,加大对科学的评价理念和考试技

术的宣传解读,普及评价知识,合理引导考试预期,增进社会对考试评价的共识。当前,考试与评价改革新技术新方法的推进存在一定难度,有很大一部分是源于社会公众与家长的不理解;因此,要积极宣传新方法和新技术,为考试改革创设良好的舆论环境,营造良好的社会氛围。

4 结束语

党的二十大报告指出,教育、科技、人才是全面建设社会主义现代化国家的基础性、战略性支撑^[1]。锚定我国2035年建成教育强国的目标,考试评价改革要服务教育评价改革落地落细,积极应对中国式教育现代化战略需求,以《总体方案》为指导,准确把握新形势新要求,强化考试评价的科技支撑,建立有效的评价反馈与持续改进机制,创新考试评价的理论与技术,提升多元评价主体和社会公众的评价素养。在具体实施中,既要聚焦重点又要融合推进,充分发挥考试服务于教育评价的功能,朝着实现中国式考试现代化的目标迈进。

参考文献:

- [1] 习近平:高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗:在中国共产党第二十次全国代表大会上的报告[EB/OL]. (2022-10-25) [2022-10-28]. http://www.gov.cn/xinwen/2022-10/25/content_5721685.htm.
- [2] 高书国. 以中国式教育现代化助推中华民族伟大复兴:深入学习领会《中共中央关于党的百年奋斗重大成就和历史经验的决议》[J]. 人民教育, 2021(23): 12-15.
- [3] AERA, APA, NCME. The standards for educational and psychological testing[M]. Washington, DC: American Educational Research Association, 2014.
- [4] 顾明远. 教育大辞典[M]. 上海:上海教育出版社, 1978: 767.
- [5] 于涵, 韩宁, 关丹丹. 关于新高考改革背景下考试质量监测与评价工作的思考[J]. 中国高教研究, 2018(10): 11-16.

- [6] 斯塔弗尔比姆. 方案评价的CIPP模式[M]//瞿葆奎. 教育评价. 北京: 人民教育出版社, 1989: 298-301.
- [7] Aurora Institute, The Center for Assessment, The Center for Innovation in Education, et al. Measuring forward: emerging trends in K-12 assessment innovation[R/OL]. [2022-11-02]. <https://knowledgeworks.org/resources/emerging-trends-k12-assessment-innovation/>.
- [8] 关丹丹, 韩宁, 章建石. 立足“四个评价”、服务“五类主体”进一步深化高考评价改革[J]. 中国考试, 2021(3): 1-8.
- [9] 叶丽新. 提升基础教育领域教师评价素养[N]. 光明日报, 2020-11-17(14).

Deepening the Reform of Educational Examination Evaluation and Promoting the Modernization of Chinese-style Examination

GUAN Dandan

(National Education Examinations Authority, Beijing 100084, China)

Abstract: The construction process of the Chinese path to modernization depends on education. As one of the most important means of educational assessment, the educational examination needs to be reformed in a deep and all-round manner to realize the modernization of Chinese-style examination. The reform of examination content starts from the starting point of examination with the goal to help the examinees do meaningful things; the reform of examination evaluation starts from the function of examination with the aim to make the results of what the examinees do get scientific and accurate interpretation, as well as a complete and effective application. Examination evaluation usually includes three main directions: evaluation of examination, evaluation of examinees, and evaluation of education and teaching. Based on the direct and indirect data related to the examination, efforts should be made to give full play to the positive function of examination by improving examination quality, promoting the individualized development of examinees, and the precise improvement of education and teaching through the feedback of evaluation results. The ideas of achieving examination evaluation reform in China are further clarified in this paper. First, we need to strengthen digital examination evaluation by means of the latest technology. Second, we need to establish a set of effective feedback and a continuous improvement system. Third, we need to innovate the theory and technology of examination evaluation. Fourth, we need to improve the assessment literacy of the subjects being assessed and the general public. In conclusion, various measures should be taken to deepen the reform of examination evaluation and construct the Chinese-style assessment modernization.

Keywords: the examination evaluation; the reform of educational assessment; Chinese-style modernization; modernization of examination

(责任编辑:张 丽)