

试是对学生的高中课程水平进行评价,也是英国大学入学的统一测试,负责该考试的剑桥大学考试委员会多年来通过不断增设考试科目,为学生提供更多的自我认知和发展机会,也使得大学可以更为全面地考查和评价学生。同样,美国SAT考试也是学生申请大学的依据,为了整体提升学生接受大学教育的能力和水平,组织此项考试的美国大学理事会又开发了大学先修课程考试,以推动学生更好地适应大学教育。因此,考试机构的专业化还体现在能否结合行业特点进行考试开发,就教育考试机构来说,需要根据人才培养规律和培养目标不断进行考试项目的开发。

考试是以教育学、心理学、社会学、统计学、信息技术等为基础的,围绕考试的功能与目标将各学科的研究成果综合运用于考试设计与开发也是考试机构专业化的体现。我国虽然有上千年的考试历史,但仍以传统的纸笔考试为主,以

统一时间、统一科目、统一试卷的传统方式来确保考试的公平和公正;因此,如何发挥信息技术的支撑作用,有效提升教育考试机构对大规模考试的组织能力,也是新时代发展的要求。

面对新时代提出的建设高质量教育体系要求,全方位推进中考、高考改革,完善初中、高中学业水平考试等都亟须教育考试机构在考试开发技术、考试组织方式、考试评价引领等方面发挥更加专业的作用。《总体方案》也明确要求:坚持科学有效,改进结果评价,强化过程评价,探索增值评价,健全综合评价,充分利用信息技术,提高教育评价的科学性、专业性、客观性。教育考试机构应在工作内容、队伍建设、专业服务等方面全面提升能力和水平,丰富考试评价方式,促进学生全面发展,为形成富有时代特征、彰显中国特色、体现世界水平的教育评价体系作出应有的贡献。

## 高考改革:校内成绩计入 公开考试成绩的香港经验及启示

侯杰泰

(香港中文大学,香港特别行政区 999077)

在不同国家及地区教育改革中,考试改革往往是首当其冲。中国是考试的发源地,考试改革一直都在进行中,但每一步都举步维艰,足见涉及考试改革的问题极为复杂。高考是改变普通人命运的机会,其高风险性极易引发教育及社会问题。中国大陆目前正在推进的高考综合改革,探索依据统一高考和高中学业水平考试成绩、参考综合素质评价即“两依据、一参考”进行高校招

生录取,其中综合素质评价包括的内容比较广泛,既有思想品德、身心健康、艺术素养、社会实践等方面的考核内容,也有学业水平即平时学习成绩的考量(在香港地区称为校本评核)。将校本评核成绩计入公开考试成绩内,在一些西方国家普遍采用,也曾受香港地区的推崇。本文总结香港地区采用后的经验及争论,供中国大陆在实施高考改革时参考。

收稿日期:2020-11-24

作者简介:侯杰泰(1955—),男,香港中文大学教育心理学卓敏讲座教授,美国教育研究协会院士,国际应用心理学会院士。

## 1 校本评核跨校分数可比的解决方案

“一考定终身”给学生带来严重的精神压力,为此有人提议,将学生平时的校内成绩或某些作业得分,以一定占比计入高考(各国或地区采取的大学入学公开考试,简称“高考”)成绩,或在高考成绩单中同时列出学生的校内成绩,要求大学招生时一并参考。赞成者认为,这种方法可以解决“一考定终身”,而且有些重要能力(如实验操作能力)难以在试场内考核,校内评核方式(如做实验、专题研究)能够弥补公开考试用纸笔考核方法的不足。

正是这个原因,香港地区将涉及实验科目(如化学科)的校内实验考试成绩、通识教育科专题作业成绩计入高考成绩,并从2012年开始在多个科目中采用校本评核。因为不同学校采用难度不同的试题,评分松紧度也不一致,怎样将校内成绩转为同一量尺上的分数,怎样保证各校不发“善心”地令校内成绩“人人满分”,是要处理的首要问题。为保证跨校分数可比,通常有以下2种解决方案。

第一种,专家监督调整。由考试机构(香港考试评核局)给出一些标准考题及详细的评核标准,要求各校用同一评核准则进行评分。考试机构也派出监督人员到访学校,督查教师评分松紧是否合宜。显然,这种监督的效力非常有限,教师可以在监督人员面前要求十分严格,但在其他时间,教师评分情况可能大不一样;同时,学生是否有反复、秘密地操练考题,难以监控。故此,这些校内分数究竟能否真正反映学生能力,到访的监督人员无法知晓。虽然考试机构会通过每所学校的分数分布监查教师评分松紧度,但总有一些学校钻空子,采用“拉下补上”的方法,即对一些成绩较差的学生给一个更差的校内成绩,对成绩较好的学生给一个更高的成绩,这样学校整体成绩的平均值看起来仍然正常。派专家抽查校内评分,除工作量极大之外,仍难以避免“代笔”等各类

做假行为。鉴于以上弊端,一些国家/地区在尝试后放弃了专家监督方法,而采用下述第二种方法。

第二种,用公开试分数调整校内试分数。这种解决方案是用整所学校在该科的公开试成绩作为该校的能力依据,以此调高或降低该校的校内成绩,这样就不会因为校内考试评分松紧度不一致使学生受损或获益。香港地区从2012年开始实施校本评核,主要采用这种方法调整校内考试成绩。

支持校本评核的专家认为,这种方法能拓展测评笔试无法测量的学生重要能力的范围;但实际情况是,扩大测评范围是难以实现的,因为学校明白了公开试调整方法,虽然个别学生的高考成绩受校内成绩影响,但对于整个学校来说,校内分数并不影响学校的最终高考成绩。学校只要将全部精力放在公开试即可,校本评核完全不能令学校关注高考以外科目的教学内容。此外,校本评核对每所学校学生整体成绩来说是零和游戏,提升与调低高考成绩的学生人数相近;但是,对于学生本人来说,校内成绩较佳但公开试失手的学生,在计算平时成绩后,高考成绩会有提升,平时成绩较低学生的高考成绩将被调低,同学由盟友变成敌人,不利于促进互助互爱。

## 2 校本评核的启示

西方国家与亚洲国家由于文化差异,且学生有更多升学及通过其他专业训练就业的途径,公开考试并非跨进大学的“独木桥”,因而学生和家長对分数的紧张程度较低,计算平时成绩的争论也较少。亚洲国家/地区的很多父母极度紧张孩子的考试表现,会为了测验少计几分跑到学校与教师理论。以香港地区小学五、六年级学生考核为例,各校普遍采用统一测验日、考试周、全班统一考试等笔试形式,为的是怕家長投诉不公,教育部门原本是鼓励学校用多元校本评核,但最终为了体现公平,学校又采用纸笔考试,这些经验

足可借鉴。

将平时成绩计入公开试成绩,事实上阻碍了教学改革。在“公平”的旗帜下,教师不能组织学生作小组专题报告,因为小组报告的得分不知应该归入哪个学生的成绩内,教师只能要求所有学生作同一专题、相同作业,以达至“公平”;学生遇到困难,教师也不能协助,以免偏私。此外,大家认为采用校本评核不再“一考定终身”,学生精神压力可以减低,实际上将平时成绩计入高考成绩后,学生感觉每天都是考试日读书日,投入课外活动的兴趣降低,精神压力并未减轻,反而天天处于紧张状态。教师工作量也因公开试调整大量增加。基于以上原因,香港地区原本要在很多科目推行校本评核,但实际实施时有些科目取消,多科精简执行方法或以操作性考试替代。

以上经验告诉我们,香港地区借鉴西方国家的校本评核在本地区大范围推行,未充分考虑中国强调考试成绩的文化特质,因此在实施中与预期还是有一定差距的。

### 3 结语:解决“独木桥”的问题

考试压力问题严重,源于我们很多中国人认为升大学是人生成功的唯一出路,要增建更多通往成功的桥梁,需要多方努力及配合。解决考试

问题,着力点不一定全在考试。为使人们有更多元出路,政府及社会可考虑携手创立贴近现实且具市场吸引力的专业或职业途径,如新加坡建立了比大学更美轮美奂的职业学院,政府强制所有大企业必须为毕业生提供一定比例的在职进修职位。据笔者观察,有一定比例的高中学生早已不想全日坐在教室里听课,因此增加全日制大学学位并不能解决问题;反之,让这些学生先就业,赚钱养活自己,然后依循就业方向有针对性地在职进修。为配合专业或职业发展,政府要为各行业制定专业要求,使专业/职业课程更具专门化、更有吸引力。在这方面,欧洲很多国家做得不错,值得学习和借鉴。

2000年前后,很多国家和地区都进行了课程及考试改革,近些年这些国家和地区也在重新检视教育上的重大问题,相互比较各国的改革方案,参与国际讨论(如OECD教育2030计划)。不难理解,如今世界各地皆求变心切,但面临的困难也不少,公众也常误以为存在简单的解决方案,只不过是政府或有关部门不愿意采纳。实际上,从世界各国的经验来看,改革是众多考虑的平衡,从来不易。世界上并不存在一个完美的考试系统,也没有放之四海皆准的考试供不同文化背景的国家/地区采用。

## 探索综合素质评价的科学标准

孙昌华

(美国ACT公司,美国爱荷华州爱荷华 52246)

中国于2014年开始的高考综合改革探索“两依据、一参考”的招生录取模式,其中的“一参考”指的是综合素质评价,但是从目前试点省份的招

生实践来看,与预期尚存在一定距离。综合素质评价未能真正在高校招生中发挥“一参考”的作用,主要原因在于缺乏以科研为基础的学生综合

收稿日期:2020-12-03

作者简介:孙昌华(1955—),女,美国ACT公司学习产品服务部资深研究科学家。