

关于普通高中学生综合素质评价研究

崔允·郭 柯 政

【摘要】 普通高中学生的综合素质评价是今后高中教育和考试评价改革的一个重要热点内容。本研究首先对我国当前各地综合素质评价政策实践的现状和问题进行了考察分析,其次对其他教育发达国家和地区在学生综合素质评价方面的经验和做法进行了比较研究,最后在如上的研究基础上对如何推动和完善当前普通高中学生综合素质评价政策提出一些思考和建议:明确综合素质评价制度的功能与关系定位;对“综合素质评价”进行概念修正;用“课程学习记录”来评价学生的非学术能力表现;“课程学习记录”的具体操作建议;完善综合素质评价的监督制度。

【关键词】 普通高中 学生评价 综合素质评价

【作者简介】 崔允·郭/华东师范大学课程与教学研究所教授 (上海 200062)

柯 政/华东师范大学课程与教学研究所讲师 (上海 200062)

促进学生综合素质提高,是我国素质教育的基本内涵和要求。但当前,高中教育阶段的“应试教育”倾向还相当明显,把考试分数高低作为评价学生的唯一标准这种现象还比较突出。而构建一套综合素质评价制度对于推进素质教育,改变传统以高考成绩“一考定终身”的做法,有着十分重要的意义。但由于综合素质评价在我国还是一件新生事物,学界对综合素质评价制度的基本定位和框架进行深入和系统思考的还不多,已有的政策实践大多也还处于试误阶段。在此背景下,本研究首先对我国各地综合素质评价政策实践的现状和问题进行了考察分析,其次对教育发达国家和地区在学生综合素质评价方面的经验和做法进行了比较研究,最后对如何推动和完善当前普通高中学生综合素质评价政策提出一些建议。

一、当前我国普通高中学生综合素质评价的现状与问题

从政策实践上来说,国家提出要实施高中生综合素质评价也已有多年,各地在多年的实施和探索过程中已经形成了许多做法,取得了一些进展,但仍然存在着一系列的问题亟待解决。当前我国普通高中学生综合素质评价的主要现状和

问题可以归纳为如下几个方面。

(一)对“综合素质”的理解异中见同

什么是综合素质?这是实施综合素质评价的一个最基本的问题。根据我们对有关各地的相关政策文本和实践的考察分析,我们发现教育部以及各地对综合素质具体内容的理解相差很大,但是存在着一个很大的相同之处,那就是基本上大家都把综合素质理解成了“非学术能力表现”。

教育部在《教育部关于积极推进中小学评价与考试制度改革的通知》中,明确提出学生的综合素质主要包括道德品质、公民素养、学习能力、交流与合作、运动与健康、审美与表现等六个方面。有些省市在实施过程中直接参照了这个框架,但许多省市则对具体内容进行了调整。从总的情况来看,目前各地所界定的综合素质评价维度相差较大。如浙江省只有四个评价维度,而且只有两个“运动与健康”、“审美与艺术”评价维度处于教育部六个维度的范围之内。

但一个很有意思的现象就是,尽管大家对综合素质的具体内涵存有大的偏差,但一点却是高度的统一,那就是都把综合素质大体等同于学生所拥有的非学术性能力。如尽管教育部所提出的六大领域中包括“学习能力”这一部分,但从具体的描述中可以看出,这里的学习能力更多的是指

学习的兴趣、愿望和习惯等内容,而不是学业成就。只有北京市等极少数地方明确把学术成就作为综合素质的重要内容进行评价。

(二)综合素质评价在整个评价体系中的定位模糊

显然,综合素质评价是我国学生评价体系中的一个组成部分。除了综合素质评价之外,还有一直存在着的高考制度,以及最近同时在推进的高中学业水平考试制度。那么,综合素质评价与这些主要评价制度之间的关系如何,它在整个评价体系中的地位如何,该扮演何种角色呢?从理论上说,这也是各地实施综合素质评价政策时所首先需要回答的问题。但实际上,尽管各地也都有一些零星的探索,但总的来说,大家对此还并不清楚,还处于“试误”阶段。从政策分析的角度上来说,多采取“得过且过”的策略。

定位的不清晰,集中表现在各地对综合素质评价结果使用的规定上,尤其是在是否需要与高考做实质性“挂钩”的问题上陷入左右为难。^[1]有些省份将综合素质评价作为普通高中毕业依据的规定,如云南省规定“2011届以后的高中毕业生必须综合素质评价达到C级以上,才可拿到高中毕业证书”。而有些省份则将综合素质评价作为高校招生依据。如江苏省的2008年高考方案规定“在高校录取时,综合素质评价中考考生‘学习能力’、‘运动与健康’、‘审美与表现’单项达到A级且其他三项均为合格的,高校在同等条件下可以优先录取;以上三项均为D级的,高校可以不录取”。还有些省市干脆就模糊其辞,“评价结果应用的主要用途是为学生和家长选择适宜于学生发展的高等院校或工作岗位提供参考;为高等学校选择适合的学生及入学后开展有针对性的教育提供参考或为用人单位选择适合的人才提供参考。”

(三)综合素质评价的配套技术与制度不足

综合素质评价对相关的配套技术与制度要求很高,而现实情况是,各地在开展综合素质评价的实践中,普遍缺乏解决这些配套技术与制度问题的思路与方向,即使偶有尝试,也缺乏有说服力的论证。困扰各地实践的主要配套技术与制度主要包括如下几方面:

1. 测量、计分问题

如何对某位学生的某一项具体素质,如道德

品质进行评价,这是各地在实施高中生综合素质评价会遇到一个技术难点,也是许多人认为学校实施综合素质评价的一个最大问题。^[2]事实上这其中会涉及三个层面的问题:首先是行为样本的选择问题。比如哪些行为表现可以算是道德品质,或者审美艺术呢?

其次是评价结果的呈现形式问题。目前各地的做法主要采取评分与等级两种评价方式。对有些可以量化的内容,如“运动与健康”中的“运动能力”,采用评分与等级的方法是可行的。但对“道德品质、公民素养、交流与合作、审美与表现”使用评分与等级则并不合适。其最大的问题在于很难在等级认定的标准上取得共识。我们不能说一个学生为老师倒了五次水,另一个学生倒了三次,第一个学生就比第二个学生更加的尊敬师长。

最后是评价结果的叠加汇总问题。除了如何把行为表现转化为具体的评价结果之外,综合素质评价所涉及的测量与评分问题还包括对各项目、各维度评价结果的合成问题。就当前实践看,各省对于学生平时综合素质评价结果如何合成综合素质等级的规定很不统一,具体方法的规定都是以经验摸索为基础,显得很随意、混乱。

2. 管理的问题

综合素质评价给管理提出了很高的要求,但就目前各地的现实情况来看,管理问题还比较突出。这首先表现在如何建立一套专业、负责、有序的管理机制问题,其次是如何避免简单化、去专业化的管理措施问题。

当前的综合素质评价,各个责任主体之间的责任和权利缺乏相应的规定。针对综合素质评价,并没有一套完整的管理机制。综合素质评价具体该归属哪个部门管理,该由谁来提供专业服务,由谁来进行相应的监督和管理,没有一个清楚的框架。除了管理机制没有理顺之外,在各级教育行政部门的具体管理中,也经常出现简单化、去专业化的管理措施,与综合素质评价所倡导的评价理念相冲突。如某省规定,除“公民道德素养”外,其他五个方面被评为B的学生数在省级示范高中、市级示范高中、一般普通高中中必须分别控制在被评学生数的60%、50%、40%。又如某省规定,综合素质评为A等的,各省辖市应控制在30%以内。这种硬性的规定或者是处于某种管理

方面的考虑,但从本质上来说,这些管理措施显然并不符合综合素质评价的理念。

3. 诚信与监督机制的问题

虽然各地从文件上都明文要求建立抽测制度、监督制度、公示制度、责任追究制度等。但这些制度多停留于纸面,流于形式,并没有落实到实践操作中,造假的现象仍然比较突出。更重要的是,民众对诚信问题异常敏感,甚至容易把一些偶然出现的问题扩散为评价系统的本质与结构性问题。在这种情况下,在如何承受与化解这些方面压力上也就需要作出更多有实质性的进步,但目前来看,各地在这方面的探索与进步并不明显。

二、综合素质评价的国家和地区经验

为了借鉴其他国家和地区的有效经验,我们对美国、英国、加拿大、新加坡和我国香港等八个主要国家和地区的高中生评价体系进行了比较研究。研究发现,各国都非常重视高中学生综合素质的评价,形成了许多可资借鉴的做法和经验。

(一)对综合素质的理解和界定

各主要国家和地区在对高中学生进行评价的时候,都注重学生的综合素质。但它们对学生综合素质的理解既有相同之处,又有不同之处。

相同之处是它们都认为学生的学术能力或者学业成绩是学生综合素质的重要组成部分。在我们所搜索到的各种针对高中生的评价体系中,基本上都没有把学生的学术能力或者学业成就排除在综合素质之外,而都是把它作为一块十分重要的内容。把综合素质界定为除学生的非学术能力之外的素质,则基本没有。

如果参照我国的综合素质评价概念,国外对非学术能力表现包括哪些内容,同样是很不一致的。例如在新加坡评价体系中,非学业成就有广义和狭义之分。狭义的非学业成就是指学生在课外活动(Co-curriculum Activities, CCAs)中取得的成绩,广义的还包括学生的品性。而在加拿大安大略省的标准化学习报告单中,非学术性素质主要是社区活动。在美国,它所包含的内容则更多,任何能够说明自己在非学术方面取得成就的材料都可以作为评价内容。

(二)综合素质评价的功能定位

对高中学生进行综合素质评价的目的是什么?综合各国情况来看,大部分重要的评价都是为录用单位(主要是大学)选择候选人提供重要信息的。这其实也很好理解。任何评价都需要花费成本,而且评价所涉及的群体越大,评价难度越高,社会成本越高。所以,如果没有巨大的社会价值或者社会需求,对大量的高中生进行综合素质评价是不可能维持的。而对绝大部分高中学生或者家庭来说,毕业后能进入大学进一步深造是他们共同的需求,正是这种需求使评价产生了巨大的社会价值。换言之,如果综合素质评价对大学录取新生没有作用,那么它就不大会引起社会的关注,因而也就不会成为“重要”的评价制度。

如果说不同国家、地区的综合素质评价在功能定位上还有些区别的话,那这种区别主要就是程度的差异。有些综合素质评价成绩直接成为大学入学成绩的一部分,如韩国、新加坡等,而有些综合素质评价成绩只是大学招生的参考依据,如加拿大的高中报告单。

(三)综合素质评价方式

如果把学生的综合素质看作是学术性能力与非学术性素质的综合。那么考试是其中一个重要的评价方式。如果以可否进行考试和是否学术性素质进行交叉匹配,我们就可以简单地把综合素质划分为四类:

首先是可以考试的学术能力。

其次是很难用考试,尤其是一次性考试来测量的学术能力,如高阶的思维品质(如运用历史知识来对当下时政作出分析判断的能力)。对于这部分能力,许多国家和地区都通过用平时的课程成绩来呈现,以弥补一次性高利害考试的不足。

第三是能够用考试或者结构化较强的测验来衡量的非学术能力。在我国,我们很多人都默认非学术性素质(即我们所说的综合素质)基本上是不能通过考试来评价的。^[3]但有些国家和地区却试图把学生的非学术能力进行结构化的测试,藉此更好地纳入评价体系。如英国QCA组织的“受控评价”就是让学生置身于一个模拟真实或者类真实的环境中,让学生进行真实操作,然后对学生的行为表现进行评价,非常类似我们所熟悉的实验考试。这种评价方式可以较好地评价学生的许多非学术性能力,如家政、商业能力等。香港地

区所采用的校本评核机制的功能也与此类似。

第四类就是不能用考试来测量的非学术能力。这类素质与我国所说的综合素质最为接近,如思想道德品质、公民素养。对于那些不能用考试来测量的非学术能力,各国和地区又大致存在着两种不同的评价方式。一种只收集那些有明确证据的行为,其典型的例子有新加坡。另一种是对各种材料进行模糊的判断,其典型的有美国。

新加坡把非学业成就限定在学生课外活动中取得的成就,具体包括体育活动、团体艺术表演、童子军(Uniform Groups)、学生社团等四部分。对哪些活动可以进入评价范围,其具体的行为表现标准以及分数比重都有非常清晰的界定,如担任校学生会主席可以相当于几分。以模糊评价为主地区的典型代表是美国。各大学招生时并没有明确、系统的评价指标,而是采用了大量评语式的材料来进行定性的评价。一个完整大学入学申请材料主要包括了入学申请(application for admission)和附属材料两大部分。而综合素质评价的内容在这两部分中均得到了充分体现。它将学生的综合素质作为一个整体,采取了模糊评价、总体评价的方法,这与新加坡的方法大异其趣。

(四)综合素质评价的操作方式

大体来说,各国和地区在对学生进行综合素质评价上有两大类不同的操作方式。一类是学生收集各种能够证明自己各方面素质的材料,提交给大学进行评价。这类操作方式的前提条件是大学具有很强的自主权。

另外一类是国家规定和要求学生提供哪些资料,大学必须依据某些要求(比如哪一类材料占据多少比例,或者考虑的先后次序)对学生的综合素质进行评价。如新加坡就规定了三大类信息,这些信息同时也作为大学录取的主要考虑内容。又如韩国,学生簿是记录学生非学术能力的重要凭借,也是大学录取的重要标准,而且其权重近年来逐步提高,但国家对于学生簿的内容结构做出了比较清晰的规定。

(五)综合素质评价的监督机制

就我们所研究国家地区来看,它们主要采取了如下两套机制来尽可能地保证综合素质评价机制的公正运行。

首先是加强评分的客观性。客观性是公正的

前提,只有保证所评价内容在计分上的客观性,那么才能得到公众的认同。所以,各国、地区采用了多种办法保证评分的客观性。考试是其中一种运用较多的方法。与我们许多人的印象相反,许多国家和地区都努力通过考试成绩来表征学生的素质和能力,因为人们觉得考试拥有很高的公正性。另一个办法是只把那些可以收集到明显证据的素质和能力列入评价体系,而把那些很难具体量化的内容排除在评价体系之外。其典型的例子是新加坡,它把非学术能力就限定在四部分内容上,而且对四部分内容的计分方式都作了清晰的量化,这样的操作清晰明了,不大可能作假。当然,其弊端也明显,那就是把诸如道德品质等很难量化的指标全部剔除掉了。

其次是依赖良好的社会诚信文化。我们发现,在有些国家,由于有着比较好的诚信文化,在综合素质评价中并没有设立专门的监督机构。如在美国的大学入学申请中,并没有设置独立的机构或者机制来审查综合素质评价,而是更多地靠人们的法制意识和诚信文化来开展。但一旦查出有弄虚作假的情况,即会执行非常严厉的惩罚。社会诚信文化的另一个组成部分是公众不会对一些处于有争议范围,而又处于大学自主权范围的事件持过度的关注和焦虑,从而放大诚信问题。

三、进一步推动和完善综合素质评价政策的建议

基于上述对我国普通高中学生综合素质评价现状的分析和不同国家/地区的比较研究,本文最后对如何进一步推进和完善普通高中实施学生综合素质评价提出一些政策建议,供大家批判。

(一)明确综合素质评价制度的两个定位

我们认为,下一阶段要继续推进和完善综合素质评价政策,首先就要明晰两个定位问题:功能定位和关系定位。

1. 综合素质评价制度必须纳入高校招生的指标体系

不难理解,评价学生的综合素质是一项异常困难的事务,需要大量的社会投入。但我们如何才能把现在所需要的相关社会资源动员和组织起

来?显然,如果没有这样的运作机制,综合素质评价就无法开展起来。

在我们看来,在我国现有国情之下,设计这套运作机制的一个关键点就是,把综合素质评价纳入到高校招生,尤其是重点本科院校的招生要求当中去。只有这样,我们才能促使相关资源的拥有者,如地方政府、家长、媒体、学校、研究机构等,把相应资源分配到综合素质评价上来。也只有这样,综合素质评价才能够具有巨大的社会资源动员力量,不至于使这套制度流于形式,最终实现它的巨大社会价值。

也就是说,我们的努力方向是只有通过综合素质评价体系的构建,让评价结果成为我国的高校招生的一个影响因素,从而完善高校招生制度,最终为基础教育的发展提供良好的办学指引。我们认为,这是我们致力于推动和完善综合素质评价制度首先应该明确的一个基本定位。

2. 必须理清综合素质评价与高考和学业水平考试之间的关系

在明确了综合素质评价必须纳入高校招生的指标体系这个定位之后,我们接着就需要对综合素质评价做另外一个定位,那就是它在整个学生评价体系中的位置。我们知道,我国已经有一套非常重要的评价制度,那就是高考,它在现有高校招生指标体系中占据绝对的权重。同时,近年国家又一直在推进高中学生的学业水平考试。考虑如何推进和完善综合素质评价制度,就必须考虑它与高考、学业水平考试之间的关系定位。

显而易见,这三者共同构成了我国的学生评价体系,但它们对学生学习结果的评价重点应该是不同的。那么,高考和学业水平考试的评价重点是什么呢?由于当前高考改革的思路还并未完全明确,而且又新添了学业水平考试这项制度,所以事实上大家对这两者的评价重点以及关系还处于摸索时期。但基于对我国现有评价实践和传统的考察,以及国外评价机制的分析,我们认为未来的高考与学业水平考试很可能呈现这样的关系结构:高考主要是为了预测高中毕业生的“大学的学习潜力”的评价制度,而学业水平考试主要是为了评价高中毕业生的“学术水平和能力”。^[4]

如果这样来定位高考和学业水平考试,我们就可以清楚地看到,除了这两块内容之外,完整的

学生学习结果评价就还缺少一块重要的内容,那就是对高中毕业生非学术能力表现的评价。这应该是综合素质评价制度所要解决的主要任务。也只有这样来定位这三套评价制度,整个评价体系在结构上才是完整的,才符合素质教育理念。这是我们思考综合素质评价政策的第二个定位。

(二)对“综合素质评价”进行概念修正

承上所述,如果要理顺高考、学业水平考试和综合素质评价这三套国家基本评价制度之间的关系,使之共同构成符合我国素质教育理念的学生评价体系的话,那么就需要把综合素质评价制度定位为对学生非学术能力素养的评价上。事实上,正如本文第一部分所阐述的那样,在实践中大家实际上基本上都是把综合素质评价理解为对学生非学术能力的评价。

但把综合素质评价定位为是对学生非学术能力进行评价,会带来一个“名不副实”的问题。因为,在大家看来,顾名思义,所谓“综合素质评价”理应对学生的综合素质进行评价,而综合素质可以简单地分为学术能力和非学术能力两种。所以,只对学生的非学术能力进行评价,而把学术能力排除在综合素质评价之外,这是一件很难让人理解的事情。尽管现在看来似乎问题不大,大家也都是这样在做,但这项改革措施一旦变成高利害,引起社会关注,那么“名不顺”的问题就会成为推进这项政策的一个重大障碍。

那怎么解决这个问题呢?我们认为其中一个可能的思路就是,把“综合素质评价”在概念层级上往上提升一级,把它作为整个评价体系的一个基本理念,而不是把综合素质评价理解为某一项评价制度。这样做的好处主要是,它一方面可以保留“综合素质评价”这个政策概念,避免大的政策变动,另外一方面能够从学理上把有关概念表达清楚,而且把综合素质评价理解为整个评价体系的一个基本理念,也符合当前包括高考制度改革在内的评价改革方向,有利于社会宣传。

(三)用“课程学习记录”来评价学生的非学术能力表现

把综合素质评价这个概念理解为整个评价体系的一个基本理念之后,衍生出来的一个问题就是,用什么来涵盖原来作为一项具体评价制度的综合素质评价所包含的内容?因为,学生的非学

术能力是综合素质的重要组成部分,我们的确需要有一套相应的评价制度。

综合考虑各方面,我们认为当前可以先把学生高中三年的“课程学习记录”作为评价学生非学术性能力表现的一个评价制度,替代作为一项具体评价制度的综合素质评价。所谓“课程学习记录”是指对学生在高中三年中所学习的课程以及相应的表现或者收获做出的记录。这样做的好处至少有三方面:

首先,学校课程对学生的要求就已经基本涵盖了综合素质的要求。根据新的课程方案,普通高中除了提供学术性课程之外,还有相当大比例的课程是非学术性的,是为了促进学生非学术能力提升的。如果从必修学分比例上来说,艺术、体育与健康、综合实践活动这三个学习领域的课程时间在最低毕业必修学分中要占到三分之一以上(40/116)。所以,学生的非学术性能力素养的很大一部分内容其实已经包括在学校的课程体系中,只要把学生所学习的课程及其结果记录下来,那么就基本上可以代表了学生的非学术性能力。

其次,这有利于督促学校更好地履行应尽社会责任。作为学校,它有责任和义务向社会报告它为学生提供了什么课程,以及学生从这些课程中获得的收获。这也可以促使学校更好地实施各类课程,促进学生的全面发展。

最后,这些内容相对客观,容易评价。学生的非学术能力包含的内容很广,要对它进行全面的评价会遇到很多暂时无法克服的困难,如分类问题、测量问题、管理问题。对于这些问题其他国家和地区也没有成熟的解决经验,所以是一个世界性的难题。但其他国家和地区对于如何避开这个问题却有不少成功的经验,其中有一条就是遵循有限性原则,即把一些相对成熟、可操作的内容纳入评价体系,而放弃那些争议很大,而又难操作的内容。所以相比道德品质等内容,学生在高中三年学习了哪些课程,取得了什么成绩等内容要相对客观得多。

综上所述,我们认为如图1来表示未来学生评价制度之间的关系:

(四)“课程学习记录”的具体操作建议

由于课程学习记录报告的主要是学生的非学术性能力素养,它与学业水平考试成绩与高考成

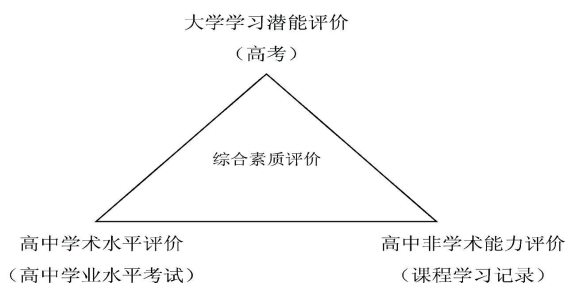


图1 综合素质评价体系

绩共同构成了学生的综合素质。基于对其他国家 and 地区对非学术性能力素养的评价机制,以及对我国国情的考虑,我们对如何进行课程学习记录作出如下建议:

首先,基于学校课程方案以及评价目的的考虑,重点记录学生在技术领域、艺术鉴赏、体育与健康、综合实践活动等四方面课程学习表现。

其次,以等级制来呈现评价结果,供高校录取时参考。如果希望更加有效地纳入高校录取条件,可以参照香港地区的做法,把每个等级换算成分数(如A等为5分),然后把5个方面的等级相加成为一个总分(如总分25分),最后规定总分在多少以上才可以毕业(如12分以上)。

第三,教育部对这四方面内容如何评分提供一个基本的评价框架,各省教育行政部门指导各地级市根据教育部的指南框架,制定符合当地的评分规则和样例,供学校参考。学校根据评分规则和样例,对学生的表现进行评价。各地级市教育行政部门每年对每所学校的评价结果进行抽样(在各等级上都抽取几份作业,包括各种证明材料),组织专家进行评核,并根据结果对学校的自评进行校正。

(五)完善综合素质评价的监督制度

尽管从学理上来说,实施综合素质评价的必要性和意义非常清楚,但在实践中,公众对于综合素质评价却是疑虑重重,他们对综合素质评价能否做到客观、公平、公正非常不确定。^[3] 公众的这种社会认知和情绪,直接威胁着这项政策的民意基础,需要决策者引起高度关注。

其实,这个问题的实质还是在于如何构建一套合理的监督机制。在我们看来,除了要继续完善常规的校长诚信承诺制度、公示制度、责任追究等制度之外,我们还可以通过如下两方面策略来加强评价过程和结果的确定性、可(下转第12页)

程的民主和公平。但是许多由各地教研室组织进行的,美其名曰“教学质量监控”的单元考和中段考,随意性和主管性就很大,应该加以规范。

教学质量监控是一把双刃剑,特别是外部的监控,做得好将有利于教学质量提升并提高教学质量,做得不好,比如说,监控的形式单一、力度过大、过于频密,都会造成危害。内部的监控也一样,教学需要多样化的方法和根据学生实际情况灵活调整的策略,但也不能变化频繁。内部的以教师为主体的监控不能过于显露和频密,要以教和学为主,将过程性的考评与监控融入到教学之中,成为教学的有机组成部分。教学质量监控的概念和相应的许多措施,是在知识经济时代,在人才竞争日益激烈的背景下发展起来,它所依托的是系统理论和管理科学。值得注意的是,系统理论和管理科学针对的是一般的自然系统、经济系统和社会系统,或甚至是抽象的系统。教育是一个特殊的系统,有着自身特定的特点和规律。比如说,许多工厂对产品的质量监控看的就是最后的产出而不问生产过程,教学就不能如此。工厂的产品质量是有完全可以测量的客观标准的,而教学质量的因素复杂,其中有许多是不可测的,甚至是多年以后才能显示出来的。不能简单地把一

些其它领域的理念和方法套用到教学系统中来。正是这样,教学质量监控到底应该怎么做?怎么和教育界一些过去行之有效的传统与方法衔接起来?如何创建新的体系和方法,都需要我们进一步的探究。

参考文献:

- [1] Lingbiao Gao & David Watkins. The relationship between school physics teachers' conceptions of teaching and student learning in China [J]. The Asia-Pacific Education Researcher, 2004(2).
- [2] Ference Marton. What does it take to learn? Some implication of an alternative view of learning [C] // Entwistle. Strategies for research and development in higher education. Amsterdam: Swets and Zeitlinger. 1976.
- [3] John Biggs. Student approaches to learning and studying, Research monograph [M]. Melbourne: Australian Council for Educational Research. 1987.
- [4] [5] [8] 高凌飏. 课程与教学质量监控[J]. 教育研究, 2004(8).
- [6] IEA. TIMSS2007 Assessment Frameworks [R]. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. 2005.
- [7] OECD. Assessing Scientific, Reading, and Mathematical Literacy: A Framework for PISA2006 [R]. OECD, Paris, 2006.

(责任校对:焦方瑞)

(上接第8页)核查性,从而降低公众对这项政策的焦虑感:

首先,努力提升评价的客观性。公众之所以会对现有的综合素质评价是否公平、公正感到担忧,其中一个很重要的原因就是对学生的非学术性表现进行评价往往依靠评价者的主观判断,而缺乏一套相对客观的标准。所以我们要坚持有所为有所不为的原则,先从课程方案明确规定的技术、艺术、体育与健康、综合实践活动这四个方面入手,而不要为了求全,把那些争议大、评价困难的内容纳入评价范围。同样,即使对如上四方面,在如何计分评价上,也要秉持简单可操作的原则,把关注点放在那些具有明确可观察的行为上。这样做可以增进评价的客观性,有利于社会的监督。

其次,建立和完善校外专家的抽样复核制度,加大纠错和惩罚的力度。对学校评价结果的抽样复核,并根据需要对各学校的评价结果进行统计校正,是增加评价信度的一个很好办法。而评价

信度的提高会直接增强民众对评价结果的确定性和安全感,从而减少他们的焦虑,缓解这项改革的社会压力。而这就需要国家充分认识到这项改革的重要性,增加相应的投入,加大专家队伍建设,制定有关操作程序,从而为扩大评价结果的校外评核范围提供保证。

参考文献:

- [1] 张远增. 学生综合素质评价实践的反思与对策[J]. 考试研究, 2008(10): 27-28.
- [2] 林珑, 王晔. 中小学综合素质评价中的问题与对策[J]. 中国考试, 2009(4): 14.
- [3] 周文叶. 论表现性评价在综合素质评价中的应用[J]. 全球教育展望, 2007(10).
- [4] 崔允漭, 杨向东. 普通高中学业水平考试国际/地区比较研究[R]. 教育部咨询报告, 2009.
- [5] 柯政. 考试评价制度改革的复杂性分析: 以综合素质评价政策为例[J]. 全球教育展望, 2010(2).

(责任校对:焦方瑞)