

学业评价标准研究：内涵、范式与策略^{*}

沈南山

(皖西学院, 安徽 六安 237012)

摘要:学业评价标准是对国家课程标准所要求的学生学习行为表现目标“测度”的具体量规,其范畴体系是一个与课程标准、内容标准和表现标准相匹配的综合评估系统。一个比较完备的学业评价标准实施方案应由评价目标、评价内容、评价范围要求和评价方法等要素组成。研制学业评价标准,需要考虑两大匹配因素:一是与国家、地方和校本课程的学习目标相一致;二是与国家基础教育质量监测目标所确立的学生发展水平相吻合。学业评价标准的建构需要制度立法,基本途径是解读国家课程标准,深入研究教材,并遵照国情,加强本土化研究,且需要根据地方和学校教育实际,制定多层面的评价标准,并在实践中不断修正。

关键词:学业评价标准;课程标准;内容标准;表现标准

中图分类号:G633.8 **文献标志码:**A **文章编号:**1000-0186(2011)11-0018-05

我国基础教育课程改革走过了十年的历程,所确立的“以学生发展为本”的评价理念,得到了社会的广泛认同,并成为主导教育实践的主流价值观念,产生了积极的改革效应。然而我们应当客观认识到:虽然我国的评价观念在发生变革,但在教育实践层面“建立促进学生全面发展的评价体系”并未得到真正落实。看看我们中小学的一些做法,背后所支撑的理念还是“应试性教育”,教育评价范式转换未能实现课程改革的预期效果。影响评价改革的因素是多方面的,调查研究表明:除了受教育制度、招生体制等一些社会发展因素制约外,教育评价机制内部“缺少规范和指导中小学日常教学与评价活动的评价标准,是一个重要原因”。^[1]新世纪课改以来,我国一直没有制定评价标准,而以“课程标准”替代“评价标准”,这样中小学在实施过程中对“课程标准”的理解与执行往往存在泛化和异化倾向,很大程度上影响了学业评价改革与发展。因此,建立学业评价标准,对规范中小学课程、教学和评价管理,引导基础教育健康持续发展具有十分重要的现实意义。本文试图从系统论和方法论的视角对学业评价标准研究的基本要素(内涵、范畴体系、建构的原则与基本途径)进行分析和

探讨,以期对学业评价改革有所裨益。

一、学业评价标准的内涵

“学业评价标准”与“学业评价”是一对关联紧密的概念。“学业评价”具有发展性和复杂性特征,主要是学业与评价发展的不同时期人们对学业和评价概念内涵的理解不同有关。本文阐释的“学业”,是指学生的学业成就(achievement),是学生经过一定时期和阶段的学习,个体的学习进展和变化程度,是衡量学生学习和发展水平的重要指标,包括学生学习成绩、思想品德、个性发展等方面的成就。而对评价(evaluation & assessment)的理解则要复杂得多,西方学者研究经历了测量、描述、判断和建构时期,形成了众多教育评价观。如,评价等同“测验”,评价是“对课程与教学目标实际达成程度的描述”(泰勒, Tyler, R. W.),评价是“为决策提供有用信息的过程”(克龙巴赫和斯塔弗尔比姆, Cronbach, L. J. & Stufflebeam, D. L.);评价是“作出价值判断的过程”(斯克瑞文, Scriven, M.),评价是通过“协商”而形成的“心理建构”(古巴和林肯, Guba, E. G. & Lincoln, Y. S.)等。^{[2][3]}关于评价的各种观点,在学术

^{*} 本文系安徽省哲学社会科学规划项目“安徽农村基础教育质量监测体系研究”(项目批准号: AHSK09-10D188)的研究成果之一。

收稿日期: 2011-01-21

作者简介: 沈南山(1964—),男,安徽六安人,皖西学院副教授,教育学博士,主要从事课程与教学论研究。

界比较权威的看法是美国“教育评价标准联合委员会”的界定：评价是对某一对象（方案、设计或内容）的价值或优点所作的系统探查。^[4]事实上，在评价的各种界定中，美国学者格朗兰德（Gronlund, N. E.）的表述十分简洁，评价=测量（量的记述）或非测量（质的记述）+价值判断，^[5]这个等量公式指出了评价的基本方法和价值取向，基本概括了评价的本质属性。

笔者认为，从现代学业评价的主要目的是为了改善学生学习、改进教学过程，促进学生全面发展，最终达到学生的个体价值和社会价值统一的评价目标来看，对学业评价的认识倾向于测量方法与价值判断的统一更符合现代教育评价观。因为新课程标准理念下，对于学业评价的定性是“以学生的发展为本”，评价方法与手段重在强调认识与实践的统一、结果和过程的统一。即不但要对学生的学业成就进行量化表征与质性判断，更重要的是判断之后怎样对学生的学习实施有效指导与行为跟进。换句话说，学业评价旨在促进学生的学业发展，“发展性”是学业评价的重点。因此，基于新课程的学业评价改革，不仅要明确目前中小学学业评价改革的目标、任务与存在的问题，更在于我们在日常的课程与教学实践中如何运用科学的学业评价体系实施有效解决。鉴于此，现实课程改革背景下，学业评价的基本涵义就是以国家课程标准所确立的教育教学目标为依据，运用质性或量性的方法、手段和途径，系统地收集一定时期或阶段内学生学习认知行为上发展变化的信息和数据，并对学生的知识、能力和情感等学业成就目标进行价值判断和实施有效指导与行为跟进的过程。

“标准”（standard & criterion），意指一定的规格水平或符合某种质量的程度，是衡量事物的依据或判断准则。一般地，“标准”是一种“测度”事物的标尺，反映事物发展质变过程中量的规定性。技术测量学意义上的“标准”，是指对所研究事物概念系统的某一或某几个方面所作的统一规定，它以科学、技术和实践经验的综合为基础，经过有关方面协商一致，由主管机构批准，以特定的形式发布，作为共同遵守的准则和依据。概括地说，学业评价标准就是学生学业成就的判断准则，是由教育部门或专门的评价机构制定的，对国家课程标准所要求的学习行为表现目标“测度”的具体量规，具体规定学生学习目标的达成程度，用以规范学校和教师的教学与评价活动的实践指南。不难看出，学业评价标准主要蕴涵两个基本要素：一是评价什么，即对学生的具体学习目标与内容加以规范；二是怎么评价，即是评价方法与尺度问题，具体规定学生学习目标的认知水平。

二、学业评价标准的范畴体系

从学业评价标准的外在属性和内在属性来分，学业

评价标准的范畴体系有匹配范畴（alignment systems）和结构范畴之分。

（一）学业评价标准的匹配范畴

学业评价标准的匹配范畴是标准的使用范围属性，是一种外在属性。从系统论的观点来看，学业评价标准并非是一个独立的评估文本，在课程、教学与评价实践的具体运用中，一个显著的特点是学业评价标准呈现一个“标准体系”，它与课程标准、内容标准（content standards）、表现标准（performance standards）构成一个综合匹配（alignment）的评估系统。^[6]研究表明，学业评价标准建基于课程标准，蕴涵内容标准和表现标准两大因素，是内容标准和表现标准的有机整合与统一，如图1所示。

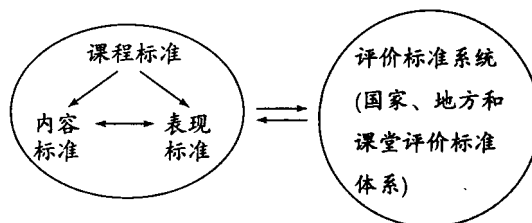


图1 学业评价标准的匹配范畴体系图

学业评价标准的匹配范畴反映了现代学业评价的复杂性，表现在诸多标准的功能作用形式的不同。但在这一系统结构中，学业评价的实践意涵就是要追求评价与标准的一致性，实现与课程标准、内容标准、表现标准是一个匹配的体系。第一，课程标准是所有标准的基石。课程标准刚性地规定内容标准和表现标准，内容标准和表现标准以课程标准为基础，不能超越课程标准。第二，内容标准与表现标准相互规约，共存于一个评价统一体中。内容标准的基本特征包含两个方面：一是陈述学生要知道的内容；二是陈述学生应该能做的事情。第二种陈述事实含有某种表现性的成分，但并不含有表现尺度（多好才算好）或表现尺度不够具体明确，它还不能为评价提供直接依据。即是说，内容标准包含表现标准的成分，表现标准以内容标准为载体。表现标准就是对内容标准的进一步分解、描述和刻画，描述的是学生对内容标准掌握的程度和质量水平，是学生要掌握知识的程度或学业成就水平的判别标准。表现标准不能离开内容标准而存在，否则失去表现标准的意义。第三，评价标准则是内容标准和表现标准的有机整合，表现形式反映学生学习内容目标和认知能力表现目标的二维性，具体规定某一知识能力要素在何种认知水平上进行评价。

（二）学业评价标准的结构范畴

学业评价标准的结构范畴反映标准的组成结构特征，是标准的内在属性。学业评价标准作为一个时期指导教学和评价活动的评估文本，其具体构成内容要素及结构

范畴的确定不仅和一个国家的课程标准有很大关联,而且与一个国家(地区)的教育实情、评价目标、评价对象、开发者的主客观理解水平等因素有很大关联,存在着多维度、多层面的学业评价标准模式。美国加利福尼亚大学评价检测研究中心在2007年的一份报告中指出,开发评估文本应当基于学生的学习(learning)、认知(cognitive)及其表现水平(expert performance)三个维度,这是专家和专业团体开发评估文本最趋于一致的强烈建议之一。^[6]这实际上指出了评估文本至少含有学习内容、认知目标和表现尺度三个基本要素。也有学者基于标准的可操作性和实用性方法的考量,认为学业评价标准主要包括表现标准、评价方法、评价样例和学生作品及评议等要素。^[1]不难看出,学业评价标准研究受诸多因素制约,不存在统一的标准模式,主要是由于研究问题的视角和着眼点不同,结构划分存在一些差异,但开发学业评价标准的一个基本原则是基于国家课程标准以及能衡量出课程标准所期望的学生学习目标的评价。^[7]

学业评价标准根植于课程标准统摄下的内容标准和表现标准,其宏观结构主要包括内容标准和表现标准,并作为一个整体功能发挥评价系统的作用。这样说未免过于笼统,需要选择一个问题视阈进一步整合与分解。从方法论的角度来看,一个比较完备的学业评价标准,应当是完整地描述了内容标准和表现标准实施细则,其评价方案应是明确的,所以从这个角度去探讨学业评价标准,其微观结构可设置为评价目标、评价内容、评价范围要求和评价方法等要素,且各要素相互联系与制约,共存于评价系统的统一体中。

1. 评价目标。评价目标是学业评价标准的“导航系统”,规定评价活动预期实现的目的和达到的要求,体现国家教育方针政策和课程目标,引导学生学业评价的方向。评价标准要使评价目标具体化,规范学生学习行为的测量目标,因而评价目标具有明确性、具体性和可测性特点。可测的学习目标依据课程目标的三维性来衡量:一是知识目标,通常在课程标准中,以分段目标或分类目标等形式列出的基础教育阶段学生具体的学习结果或学习目标,反映学生需要掌握的知识内容要求;二是能力目标,描述的是学生学习行为的表现性水平,反映学生学习的认知能力水平要求;三是情感目标,反映在整个学习过程中表现出来的与他人合作的态度、表达与交流的意识和实践探索的精神等层次上的学习目标要求。评价目标体现现代学业评价的核心价值理念,它包含的不仅有社会权威的价值,更有学生个体发展的价值;不仅蕴涵整体性的国家课程标准总体目标设定,而且包括各学科三维目标的细化,具体到分级的可测学习目标,如一级、二级、三级学习目标评价指标体系等。

2. 评价内容。评价内容是学业评价标准的核心内容

和中心轴线,是评价实施的主要载体。评价内容整体呈现学生在一定时期和阶段内应完成的学习任务,具有系统性、科学性和逻辑性。实施评价的内容基于课程标准,至少覆盖课程标准知识领域的核心概念(big ideas)和重要原理法则(key principles)。^[6]这是保证课程标准中所有的关键概念、原理、命题、公式法则和思想方法等在评价标准中得到全面的体现。

3. 评价范围要求。评价范围要求实质与评价目标是统一的,这里区分的目的是凸显“范围要求”的明确性,通常反映评价标准中表现标准的具体细则,是评价标准的重要组成部分。评价范围是学业评价标准的量和质的限度,反映学生学习目标表现水平的上限与下限,是实施评价活动的必要条件。因为只有当评价标准规定的内容领域具体、认知范围清晰时,评价的目标才有可能是清晰的,评价才能做到“有的放矢”。评价范围实质是对于特定的评价对象及确定的评价内容,学生的认知能力水平在什么限度范围内进行评价,它具体限定学生学习目标的达成程度。一般学业评价标准要规定每个年级每个学科的具体内容,并且对每个年级每个学科学习内容与目标的表现水平给予详尽的描述。

4. 评价方法。现代学业评价的方法具有多元化、多样化特征,主要实行量性评价和质性评价、形成性评价和终结性评价相结合的方法。例如,纸笔测验、表现评定、档案袋评价、案例分析评价、网络平台评价、日常观察等。评价方法要提供相应的评价作品案例,使教师、学生、家长通过作品表现描述、表现水平的划分以及作品范例说明等能够容易实现对学生表现任务的评分和解释。

三、学业评价标准研制的基本原则与途径

(一) 学业评价标准研制的基本原则

学业评价标准建基于内容标准和表现标准,研制的主要任务之一是制定中小学各年级内容标准、表现标准细则。

1. 内容标准研制的基本原则

内容标准是评价标准的一个基本维度,反映课程标准所规定的课程领域的具体学习目标。内容标准是学习和教学系统的核心,规定着课程与教学的方向。我国中小学各科课程标准规定了课程性质、目标、基本理念、设计思路、内容标准以及实施建议等,其中“内容标准”,是以学段、内容领域或模块(专题)等形式划分与设置的,规定一个时期或学段内具体学科课程领域学生学习和目标要求。但我国课程标准规定的内容标准仍是一个较大的范围标准,弹性很大且不含有校本课程内容,更由于教材的多样化和差异性,各年级的内容标准界限比较模糊,评价缺乏统一的尺度,操作比较困

难。因此,学业评价标准建构的一个重要任务是根据课程标准、教材和具体教育实际制定各年级学科的内容标准。一般内容标准研制应遵循如下基本原则。

- 内容标准描述的是学生的学习目标,要与国家、地方和校本课程的学习目标相一致,这是制定内容标准的必要条件。

- 内容标准在分解过程中要反映课程标准所规定的学生学习目标的核心概念和基本思想方法,整体上呈现九年一贯系统,且各年级自成体系,知识系统科学,逻辑连贯,符合学生的认知特点。

- 内容标准学习目标明确可测,具体适用,使学校、教师和学生容易理解与执行。

- 内容标准能够反映本地区的教育实情,并根据新教学信息适时加以修订。

2. 表现标准研制的基本原则

表现标准是评价标准的另一个重要维度,是学生学习行为表现目标的具体规范。表现标准在我国中小学教育评价者的应用意识中还相当薄弱和生疏。不仅因为“表现标准”是一个“舶来”词汇,在我国使用频率较低,而且“表现标准”是一个比较复杂的概念。关于表现标准的研究,其称谓、格式和具体内容十分不一致。表现标准体系一般由表现水平、表现描述、范例和分数线等要素组成。^[5]表现标准开发应遵循的基本原则。

- 表现标准研制需要与国家基础教育质量监测目标所确立的学生发展水平相吻合,实施国家基础教育质量监测制度,基于表现标准的评价数据及其分析是评价学生、教师和学校的重要依据。

- 表现标准的表现水平是明确的,能够清晰地区分每一级表现水平的等级。如,要了解“掌握”的表现水平,读了表现标准描述之后,能够清楚地知道该表现水平学生的学习目标是什么,相应的学生作品类别和样例是什么样的。

- 表现标准的表现描述是准确的,能够清晰地刻画每一级表现水平上所期望的学生表现任务,即能够清晰地刻画每一级表现水平的主要概念和基本技能的学习目标,并描述学生不同发展阶段,反映各种不同类型学生学业成就的多种类型的作品和样例。

- 表现标准符合地方和学校的实际教育水平。

(二) 学业评价标准建构的基本途径

1. 政府的制度立法

为了确保学业评价标准的有效执行和功能发挥,学业评价标准应当纳入国家和地方的教育教学法规的范畴之内,需要在课程标准的框架内,制定学业评价标准的有关制度条款,对学业评价标准的实施细则进行更为具体的规范。

建构学业评价标准,需要广泛的社会支持,特别是

各级政府的重视与领导。学业评价标准研究的组织机构通常是政府引领下的教育行政与科研部门,如,美国联邦政府于1994年颁布了《美国2000年教育目标法》,着手建立国家教育标准和提高委员会(National Education Standards and Improvement Council),该机构的重要举措就是和非政府机构的研究部门合作,开发评价课程标准的标准。其研究组织一般是由美国的一些著名的、具有一定权威的非营利性研究机构,如美国教师联合会(American Federation of Teacher)、成就公司(Achieve Inc.)等。我国学业评价标准建设也需要一个强有力的社会支持体系,表现在政府立法管理、财政支持与社会参与上,其制度体系的建立需要政府行为决策,并体现地方教育意志。

2. 加强本土化研究

关于学业评价标准,近年来我国学者也多有研究,重点探讨了“怎样将课程标准化为评价标准”的理论方法、“表现标准”相关研究成果的评介以及研制评价标准的构想上等。^{[1][8]}这些研究成果对我国学业评价标准研究走上科学化道路具有积极意义。但应当看到,我国学业评价标准研究起步较晚,历史研究和实验研究薄弱,基本还停留在“技术理论化”阶段,“拿来”成分较多,还要进一步针对我国教育实情,加强学业评价标准研发的基础和实验研究。这就是说,我国学业评价标准研究在认识方法上还存在引进与吸收、继承和发展的问題。第一,学业评价标准研究首先在认识上要突破“外来移植”的思维局限,应当基于我国教育国情和民族文化的生长点。国情不同、文化有差异,话语体系意境有别,这就要求我们在借鉴国外的相关研究成果的同时,必须加强本土化研究,要剥离西方教育劣质文化及其所根植的历史土壤成分,使学业评价标准能够真正有效地服务于我国基础教育质量监测,切实为中小学课程与教学评价提供评价指南。第二,学业评价标准研究在方法上应当是在专业化水准上建构的基础教育阶段学生学习质量评估标准,需要研究人员具有坚实的教学理论专业知识,才能克服经验行为的依赖性,防止研究发生偏向。学业评价标准作为一个教育评价项目,其计划、制定、执行和颁布,必须满足一个国家(地区)教育目标的现实需要和长远发展,至少应当满足两大匹配因素:(1)与国家、地方和校本课程的学习目标相一致;(2)与国家基础教育质量监测目标所确立的学生发展水平相吻合。只有这样,学业评价标准才有生根的土壤。

为此,学业评价标准的本土化研究,需要对国家课程标准和本地区使用教材进行深度解读与分析。因为一方面,课程标准是“教材编写、教学、评估和考试命题的依据,是国家管理和评价课程的基础”,也是指导学业评价的纲领性文件和建立学业评价标准的原则依据;另

一方面,教材(包括校本课程)多样化问题是学业评价标准建构的难点,如何分析与研究教材,分离中小学各年级教学内容并确定其学习认知范围,建立年级内容标准和相应的表现标准是学业评价标准建构的基本而又重要的工作。此外,学业评价标准如何体现课程标准的素质教育理念,如何设计真实性评价方法,发挥学业评价的情意功能,突出学生在学习评价活动中的主体地位等也是必须突破的难点。

3. 建立多层面的评价标准体系

从世界众多国家的评价实践来看,无论教育体制是中央集权型或地方分权型国家,学业评价标准既有国家指导方针的统一性,又有地方教育特征的差异性。我国是一个偌大的教育国度,社会经济、文化发展不平衡,地域差异性较大,基础教育存在明显的地区差异、文化差异、学生差异以及教学差异等,建立“大一统”的学业评价标准既是不可能的,也是不现实的。应当基于不同地区不同经济文化背景下的教育实际,形成国家、地方和学校等多层面的评价标准体系,在归属关系上,国家评价标准(抑或课程标准)是指导和统摄性的,地方和学校层面的评价标准应该融入国家评价标准,且在实践中不断修正。

参考文献:

- [1] 徐岩,丁朝蓬. 建立学业评价标准 促进课程教学改革 [J]. 课程·教材·教法, 2009 (12): 3—14.

- [2] 陈玉琨. 教育评价学 [M]. 北京: 人民教育出版社, 1999: 14—17, 8.
- [3] 张华. 课程与教学论 [M]. 上海: 上海教育出版社, 2000: 381—391.
- [4] Joint Committee on Standards for Educational Evaluation. *Standards for Evaluations of Educational Programs, Projects and Materials*, 1981.
- [5] Hansche. *Meeting the Requirements of Title I: Handbook for the Development of Performances Standards* [Z]. Washington, D C: Council of Chief States School Offiences, 1998: 11—31.
- [6] David Niemi, Julia Vallone, Jia Wang, Noelle Griffin. *Recommendations for Building A Valid Benchmark Assessment System: Interim Report to the Jackson Public Schools* [R]. CRESST Report 723. [DB/OL]. <http://www.cse.ucla.edu/products/reports/R723.pdf>. 2007.
- [7] Robert Rothman, Jean B Slattery and Jennifer L Vranek. *Benchmarking and Alignment of Standard and Testing* [R]. CSE Technical Report 566. [DB/OL]. <http://www.cse.ucla.edu>.
- [8] 崔允漭, 夏雪梅. 试论基于课程标准的学生学业成就评价 [J]. 课程·教材·教法, 2007 (1): 3—18.

(责任编辑: 刘启迪)

Research on the Achievement Assessment Standards: Intension, Model and Strategy

SHEN Nan-shan

(West Anhui University, Luan Anhui 237012, China)

Abstract: Achievement assessment is the concrete measure for students' learning behavior required by national curriculum standards, which is the comprehensive assessment system matched with curriculum standard, content standard and performance standard. A complete assessment standard project should be constructed by assessment goal, assessment content, assessment scope and assessment method etc. We need to consider the two matched factors when making out the assessment standards. Firstly, to be corresponded to the state, local and school-based curriculum learning goals; secondly, to be consistent with the performance standards, content of education and the quality monitoring. The construction of achievement assessment standard need legislation, and the basic path is to interpret the national curriculum standard and research materials thoroughly, to strengthen the research of localization according to the national situation and to multi-layered achievement assessment standard based on the different regions and school education practice and to be revised in practice.

Key words: achievement assessment standards; curriculum standards; content standards; performance standards