

指向核心素养的表现性评价

周文叶, 陈铭洲

(华东师范大学 课程与教学研究所, 上海 200062)

摘要:基于核心素养的课程改革驱动着评价体系的变革,促使我们必须超越传统的只注重“双基”的客观纸笔测验,采用并丰富“能检测学生的认知思维和推理能力以及运用知识去解决真实的、有意义的问题的能力”的表现性评价。美国新罕布什尔州指向核心素养的表现性评价试验为我们提供了一个很好的案例。首先,建立一个指向核心素养的表现性评价体系。其次,具体化核心素养,设计聚焦核心概念的表现性任务,并创建表现性任务库。再次,通过专业的程序和机制确保表现性评价实施的质量。我国基于核心素养的课程改革也亟须评价的跟进,构建更平衡更综合的评价体系、建立专业共同体和专业的评价机制、统整课程教学与表现性评价是我们应当努力探索的方向。

关键词:核心素养;表现性评价

中图分类号:G423.04 **文献标志码:**A **文章编号:**1000-0186(2017)09-0036-08

一、核心素养:驱动评价体系变革

受经济合作与发展组织(OECD)发起的“基于素养的课程或评价改革”运动的影响,全球范围内掀起了一股“素养”热。无论是八国集团(G8)、联合国教科文组织(UNESCO)、欧盟(EU)等国际组织,还是美国、澳大利亚、新加坡、日本等国都在试图勾勒新时代新型人才的形象,从而规约学校教育的方向、内容与方法。核心素养成为当前课程改革的关键。基于核心素养的课程改革意味着在“核心素养—课程标准(学科素养/跨学科素养)—单元设计—学习评价”这一连串环环相扣的链环中展开运作。^[1]而评价则是这一链环中决定核心素养是否能落地的

关键,它是实施课程与教学的指向标,若评价的不是核心素养,或是我们采用的评价评不了核心素养,那么,我们的课程和教学,以及学生的学习将不可能发生什么实质性的变化。我国当前课程改革同样面临着这样的问题,在这一连串的核心素养的课程与教学转化的链环中,必须明确什么样的评价能指引课程的开发与教学的实施。

基于核心素养的课程改革,需要构建一个“更平衡、更综合的评价体系”^[2]。这样的评价体系应当能很好地体现核心素养所描绘的学习结果,尤其是那些高阶思维、复杂的认知能力以及在新的情境中解决问题的能力等关键学习结果。而这就迫使我们必须超越传统的只注重“双基”(基本知识和基本技能)的客观纸笔测验,采用

基金项目:教育部人文社会科学项目“基于学生核心素养的学校课程发展研究”(14YJC880120)

收稿日期:2017-04-26

作者简介:周文叶,华东师范大学课程与教学研究所副教授,课程与教学系副主任,美国斯坦福大学教育学院访问学者;陈铭洲,华东师范大学课程与教学研究所硕士研究生。

并丰富“能检测学生的认知思维和推理能力以及运用知识去解决真实的、有意义的问题的能力”的表现性评价^[2]。大量的证据表明,表现性评价更适合检测高水平的、复杂的思维能力,且更有可能促进这些能力的获得;同时能支持更具诊断性的教学实践,促进课程与教学。^[2]美国在2010年颁布了各州共同核心标准(Common Core State Standards, CCSS),共同核心标准便成为引导美国基于素养的课程改革的路标。^[3]美国联邦政府随即资助了大学及职业准备评价合作伙伴(Partnership for Assessment of Readiness for College and Careers, PARCC)和智慧平衡评价联盟(Smarter Balanced Assessment Consortium, SBAC),要求两大联盟开发落实共同核心标准的评价。而能够检测学生对复杂学习目标理解程度的表现性评价成了两大联盟共同关注的焦点。^[4]

在全国性的大规模评价中加强对表现性评价的运用为指向核心素养的评价指明了方向。然而,构建各个地区、各所学校更“着地”的表现性评价体系,是核心素养落地更为直接且具体的问题。也正是为了更好地实施指向素养的课程与教学,美国部分州率先进行问责体系的改革。其中,新罕布什尔州(New Hampshire, NH)的推动力度很大且富有成效。并且,在2015年3月获得美国教育部的批准进行新的问责体系试验。新的问责体系旨在促进学生的深度学习和学校与学区强有力的组织变革。其核心就是指向核心素养的表现性评价——通过基于素养的教育路径确保学生获得关键知识与技能。

新罕布什尔州教育委员会将表现性评价与州的评价与问责体系进行统整,开发并实施全州范围的表现性评价,即共同表现性评价。由此,共同表现性评价和智慧平衡评价以及其他测验(如学校表现性评价)成为新罕布什尔州学生综合评价体系的组成要素。其中,高质量的表现性评价是新罕布什尔州指向核心素养的表现性评价系统的核心。表1是2014—2015学年参与试验的学区使用的评价体系。事实证明,2015年新罕布什尔州指向核心素养的表现性评价的试验是比较成功的,2015年12月美国通过的《每一个学生成功法》(Every Student Succeeds Act)中的

“评价和问责证明权”就是基于新罕布什尔州指向核心素养的表现性评价试验成果,并决定在七个州推广。^[5]

表1 新罕布什尔州指向核心素养的表现性评价体系^[6]

年级	英语语言艺术	数学	科学
K-2	学校表现性评价	学校表现性评价	学校表现性评价
3	智慧平衡评价和学校表现性评价	共同表现性评价和学校表现性评价	学校表现性评价
4	共同表现性评价和学校表现性评价	智慧平衡评价和学校表现性评价	学校表现性评价
5	共同表现性评价和学校表现性评价	共同表现性评价和学校表现性评价	共同表现性评价
6	智慧平衡评价和学校表现性评价	共同表现性评价和学校表现性评价	学校表现性评价
7	共同表现性评价和学校表现性评价	智慧平衡评价和学校表现性评价	共同表现性评价
8	共同表现性评价	共同表现性评价和学校表现性评价	学校表现性评价
9	智慧平衡评价和学校表现性评价	共同表现性评价和学校表现性评价	学校表现性评价
10	共同表现性评价和学校表现性评价	智慧平衡评价和学校表现性评价	共同表现性评价
11/12	智慧平衡评价(十可选择的美国学术能力评估测试)	智慧平衡评价(十可选择的美国学术能力评估测试)	学校表现性评价
	最终表现性评价		

二、指向核心素养的表现性评价设计

指向核心素养的表现性评价旨在检测高水平的、复杂的思维能力,并在这个过程中促进学生获得这些能力。实施这样的评价,首先要面临的是评价的设计问题。在新罕布什尔州指向核心素养的表现性评价系统中,共同表现性评价由参与试验地区的教师在专业人员的引领下合作开发完成,供试验地区的所有学生使用。学校表现性评价是本校教师在参考共同表现性评价的基础上,或对其进行改编或重新开发形成的,供本校学生

使用。那么，作为学校表现性评价范例的共同表现性评价是如何设计的呢？

（一）具体化核心素养

核心素养作为课程与评价的概念，是教育目的与学习结果的重要中介。要想把核心素养作为课程目标并发挥目标的功能，核心素养还需要采用专业的技术路径，在教育目的与学习结果之间设置一定的层级，并对每一层级的目标作出可理解、可传播、可实施、可评价的陈述。^[7]从整体描绘未来社会所需人才形象的核心素养到具体陈述可测可评的评价目标，这就需要一个层级地具体化核心素养。美国的“21 世纪学习结果与支持系统”彩虹图界定了学生为了面对 21 世纪挑战而需要掌握的知识和技能：学习与创新能力、生活与职业技能、信息、媒介技术素养、核心科目与 21 世纪主题。^[8]如何将这些高度概括

的、共性的 21 世纪技能落实到具体的课程教学与评价中？首先，美国各州共同核心标准，为具体化 21 世纪技能做了非常重要的工作，他们很好地将 21 世纪技能落实到了学科课程中。^[3]在此基础上，新罕布什尔州进一步将共同核心标准具体化。2014 年新罕布什尔州教育部门确立了与共同核心标准相匹配的素养^[9]，实现了从共同核心标准到新罕布什尔州素养转化的第一步。而后，新罕布什尔州又公布了素养效度评价工具（见表 2）^[10]，目的就是确保新罕布什尔州各地区学校自主开发的课程层面的素养与共同核心标准、州素养的一致性，从而使素养落实到具体的课程层面，最后试验地区的教师便可以根据具体课程所要求的素养与学生的实际情况进一步设定目标。

表 2 素养效度评价工具^①

	4	3	2	1
	←良好的素养陈述			不良的素养陈述→
内容领域的相关性 素养陈述在多大程度上与标准一致，引导学生对内容概念性的理解？	与国家、州和/或当地标准/框架一致；为了促进学习，领域可能会进行组合或集中。用一种清晰和描述性的方式表述了理解内容领域最重要的是什么。将内容与更上位的概念联系起来。			少量的证据表明与标准/框架的一致性。集中于事实性内容，没有概念的关联。
持久性概念 素养陈述在多大程度上反映了持久性概念？	包含了在内容领域间可迁移的并且能运用到真实生活情境中的技能。要求理解原理、原则和/或概念间的关系。			局限于教材/项目/资源的范围和顺序。是非常具体的内容。
认知要求 素养陈述促进了何种深度的知识？	要求深度理解内容和知识在各种情境中的应用。要求学生创造概念性的连接，表现出超出既定事实或字面解释的理解水平，通过内容的运用来支持他们的立场或观点。通过创造、分析、设计、证明、发展或制定促进复杂的联系。			常规或机械的思考或基本的回忆，并缺乏应用知识的机会。用简单的方式展示所学。回忆信息、事实、定义和术语（如背诵、陈述、识别、列表、复述、记忆）或执行简单的任务。
评价 素养陈述在多大程度上给了学生展示学习证据的机会？	用清晰和描述性的语言定义了被评价的内容。用跨学科的方式提供了多种多样展示学习证据的机会。			缺乏被评价内容的描述。限制了展示学习证据的机会。

（二）设计聚焦核心概念的任务

从上述具体化素养的各项要求中，我们已经清晰地看到，指向核心素养的评价要为学生提供展示指向素养的学习证据的机会。这里的机会，不是常规或机械的思考，不是回忆信息、事实、定义和

术语或执行简单的任务，而是要求深度理解内容和知识在各种情境中的应用，要求创造概念性的连接。因此，这就要求评价设计者研制出聚焦核心概念的表现性任务，让学生有机会展示对概念间关系的理解，并将其运用于真实生活情境的技能，而不

① 由于篇幅关系，此工具中的 2 和 3 等级的具体内容没有在本文中呈现，详见参考文献。

仅仅呈现对具体的、琐碎的事实或信息的记忆,缺乏在真实情境中应用知识的机会。

为了研制高质量的表现性任务,同时促进当地教师在研制表现性评价过程中的专业发展,新罕布什尔州最近几年与合作教育中心和斯坦福评价、学习与公平中心展开合作,在专业共同体的努力下,新罕布什尔州创建了自己的“表现性任务库”^[6]。任务库中的表现性任务与传统的纸笔测验的试题不一样,它有着自身的质量要求。^[11]第一,聚焦核心概念,即聚焦单元或课程中的大观念、核心概念或技能,且与教学一致。第二,强调学生参与,聚焦课程的重要内容,解决教室以外世界中的真实性问题,为学生提供多种方式和决策的机会。第三,注重公平,评价的条件对所有学生都是相同的,所有学生都能获得所需资源;任务在语言和设计上不含偏见,素材对于不同文化、性别、语言和群体的学生是熟悉的;任务是开放的。同时,任务设计还要遵循一些基本的原则:指导语清晰简洁,没有无关的信息;指导语中没有学生难以理解的词汇;呈现方式聚焦于期望的任务和产品;问题和提示语清楚且格式统一。当教师们合作完成一系列的评价设计之后,为了进一步提高共同表现性评价设计的质量,国家教育评价改进中心等技术伙伴会进一步对评价任务进行技术审查;而后交给教育部门的项目领导进行最后的审阅^[5],只有审阅通过,才有资格进入任务库^[12]。

在开发表现性任务的同时,教师还要开发针对任务的评分规则。同样的,评分规则要清楚地表明与所要评价的素养和标准一致。评分规则的各个维度要聚焦且相互之间不交叉,各个表现等级的表达清楚易理解。整个评分规则描述的是可观察的行为与技能。评分规则的质量直接影响着评价的质量。

三、表现性评价的实施

(一) 收集学生的表现信息

仅仅通过一个评价并不能证明学生已经达到了某种素养,学生应该有多种机会来呈现自己达到某种素养的证据。因此,要求参加指向核心素养的表现性评价试验的地区混合使用多个学校表现性评价和共同表现性评价收集学生表现信息,

同时地区可以自行决定实施表现性评价的数量以及共同表现性评价和学校表现性评价两者之间的分配比例。其中,3—8 年级每学年和高中阶段(9—12 年级)至少实施一个共同表现性评价。^[6]

虽然共同表现性评价是各地区各学校统一的,可直接从表现性任务库中获取对应年段和科目的评价任务^[13],学校表现性任务需要各学校的教师自己设计,但无论是共同表现性评价还是学校表现性评价都采取了与课程教学相统整的方式进行,且没有严格规定实施的具体时间及时间长度,只要在相应年份完成即可,教师可以根据学生的具体情况来灵活调整。如要求在 9—12 年段的英语语言艺术课上实施的共同口语交流评价,教师将其安排在学生至少完成一篇写作之后,实施时间持续 4 周。在实施评价过程中,教师可以通过一系列形成性评价,如重要事件、基准、观察、对话、学生反思以及学生最后提交的成果来了解学生对相关素养的掌握程度。共同口语交流任务中,教师通过评价学生作文的论点和证据,确认学生将在口语交流中采取的形式,连续两天观察学生在图书馆准备的表现以及学生最后要提交几分钟的口语交流视频和一定字数的作文获取信息。^[14]

(二) 确保评分的可靠性和准确性

无论表现性任务设计得多么好,如果在实施过程中,教师对学生的表现不能恰当地进行评分,或许就不能产生满意的结果。为了确保评分的可靠性和准确性,新罕布什尔州做了严谨而细致的探索。

新罕布什尔州建立了严格的评分培训机制,在正式评分前对评分者进行培训。这里的评分者指的是教师。教师在参加培训前要具备相关的知识和经验,需要对评价目标进行充分的讨论,同时进一步明确评分规则^[14]。培训以小组的形式展开,每组由 5—8 位教师组成,其中 3 位教师分别担任主持人、计时员和记录员。每位教师要在规定的时间内严格执行相应环节的任务。首先,主持人确认每位教师都有关于任务、学生表现成果样本、评分规则和评分指南等材料并介绍评分的流程和规范,教师对流程和材料不清楚的地方进行提问。接着,进入正式模拟评分阶段,教师各自独立使用评分规则对学生的作品进行分

项评分和整体评分，同时记下评分的证据。模拟评分结束后，各自展示自己所评的每个维度的分数（没有解释），记录员负责记录每个教师所评的分数。然后，教师们围绕出现不同分数的原因进行讨论，教师需要用证据给出自己评分的依据，直到小组得出一致的（至少 80%）的分项分数和整体分数。^[15]

除此之外，建立评分者之间的信度是保证评分客观性的关键一步。新罕布什尔州在 2015 年的比较性研讨会上对教师评分的严密性和一致性进行了校准。来自不同地区的教师随机进行两两分组，对近 400 份学生作品进行评分。其结果显示，除了三个作品的分数不同，其他的作品都保持了分数的一致性。而这三个作品的不同分数在预期范围内，是由样本容量和测量固有的误差引起的。^[5]

（三）评价结果的运用

1. 为学生个体提供素养发展报告单

表现性评价致力反映学生的学习进程，即通

过评价可以判断在一个特定领域中，学生在关键知识和技能的发展中所处的位置。^[2]素养报告单是汇报学生学习进程的一种方式。以参加指向核心素养的表现性评价试验的桑伯恩地区（Sanborn region）为例，桑伯恩地区的学校为学生个体提供了素养发展的报告单。表 4 是关于某学生九年级全球英语学习这门课程的相关课程素养发展的报告单。E 代表高级，指学生能够在新的任务中一贯、独立地展示出分析，总结重要内容、知识和技能的能力。P 代表精通，指学生能够在新的任务中一贯、独立地展示出应用和迁移重要内容、知识和技能的能力。BP 代表基本精通，指学生能够在熟悉的任务中展示出理解和运用重要内容、知识和技能的能力。LP 表示有限进步，指学生不能展示运用和迁移重要内容、知识和技能的能力。通过素养报告单，学生不仅能了解自己不同阶段在这门课的各个素养上的表现水平，还能根据自己薄弱的环节做出有针对性的调整。

表 4 桑伯恩地区九年级某学生素养报告单（节选）^[16]

全球英语学习				
教师：拉维依·克莉丝朵 缺席：2 天 请假：4 天				
学科素养				
能力	1/4	第一	3/4	第二学期
阅读/词汇：使用不同的策略来解读情境中不熟悉的单词	BP	BP	P	P
阅读理解：使用不同的阅读策略对文本进行关联、解释、总结、评价和推测	P	P	P	E
写作：根据不同的目的，运用标准英语写作的过程和规则形成一个书面作品	P	P	P	P
口语交流：能够有效地通过讨论和各种不同类型的媒体来进行联系，交流思想和传递信息	LP	LP	BP	P

2. 为问责提供“可比较的年度测评”信息

“年度测评”是指在学年结束时，对学生进行成就水平分类（如精通，基本精通）。那么，如何保证地区间的年度测评具有可比性呢？即如何保证学生的作品在地区 X 得到了水平三，那么与这个学生相同水平的作品在地区 Y 也被归为水平三。

指向核心素养的表现性评价的“年度测评”是由一个共同表现性评价和多个学校表现性评价组成^[5]，每个评价的实施，学生都会获得相应的评价分数，目前采用取平均值的方法，将学生获得的多个评价分数整合成一个最终的分数^[6]。通过标准设置的方式来确定划分每个成就水平的分

数，即确定在某个分数区间的学生处于哪个水平阶段。

标准设置的关键在于成就水平描述，新罕布什尔州指向核心素养的表现性评价领导小组的老师们在参考引导智慧平衡评价发展和标准设置的成就水平描述的基础上制定了英语语言艺术和数学科目年度测评的成就水平描述。接着，进入在线调查环节，使得每个学校的教师可以通过阅读成就水平描述来确定自己学生所处的成就水平。最后，通过比较学生所获得的最终分数和教师判断其所处的成就水平，确定划分每个成就水平的分数。这样，就能使最多比例的学生被划分到对应的成就水平。同时，指向核心素养的表现性评

价的“年度测评”与智慧平衡评价的相似性，暗示了参加指向核心素养的表现性评价地区和没有参加指向核心素养的表现性评价地区比较的可能性。^[5]

表5是2015年12月新罕布什尔州发布的表现性评价的结果。相关部门可以很清楚地从结果中了解到在英语语言艺术方面，每个年级学生达到甚至超过成就水平的情况。这让新罕布什尔州教育部门看到了指向核心素养的表现性评价试验带来的改变，帮助了新罕布什尔州教育部门做出继续扩大试验还是进行进一步完善评价的决定。

表5 英语语言艺术：
2015年参加试验地区的学生成就水平结果（节选）^[17]

年级	水平三	水平四	水平三和水平总和
3	30%	25%	55%
4	31%	15%	46%
5	31%	18%	50%
6	30%	11%	41%
7	38%	10%	49%
8	36%	10%	47%
11	31%	19%	50%
PACE地区	33%	16%	48%

注：水平三表示达到成就水平，水平四表示超过成就水平。

四、指向核心素养的表现性评价的启示

（一）构建更平衡、更综合的评价体系

在很大程度上，指向问责的、高利害的评价，其内容、方式和质量决定了整个评价的生态，乃至决定了整个教育生态。构建促进学生核心素养培育的评价体系，首先必须变革问责体系。新罕布什尔州正是抓住了这点，对问责体系进行了重新设计——立足素养教育、确保学生获得有意义的、促进深度学习的机会。从表1指向核心素养的表现性评价体系我们可以看到，整个评价体系的设计不仅落实了州和联邦的问责要求，在特定年级开展智慧平衡评价、美国学术能力评估测试和共同表现性评价，其中州层面的共同表现性评价就是应素养教育之需而新设的。同时，特别强调学校的主体作用，更多地给予学校自主权，学校表现性评价在整个评价体系中占据了重要位置。由此，一个更加平衡的综合的评价

体系得以建构。

我国基于核心素养的评价改革，首先面临的也必定是问责体系、高利害考试的改革。正所谓的中高考不改，其他一切都难改。新罕布什尔给我们的启示是：第一，改变原来对深度学习、关键素养关注不够的评价，进而建立高质量的表现性评价，因为它是检测学生深度理解复杂学习目标的必然。第二，改变原来以外部评价为主的评价体系，进而促进外部评价和内部评价新的平衡。第三，改变原来完全自上而下的问责模式，进而建立上下互动的合作伙伴模式。正如新罕布什尔教育部门负责人所说的：“与那些更倾向于服从取向的州问责体系相比，新罕布什尔州指向核心素养的表现性评价是一个真正的学区、教育部门和技术支持顾问之间的合作伙伴关系。”^[5]

（二）建立专业共同体和专业的评价机制

研制与实施指向核心素养的评价面临着诸多专业上的挑战，尤其是受占统治地位的传统纸笔测验的影响。新罕布什尔州通过建立内部专业共同体和借助外部专业力量来研发高质量的表现性评价。首先是每个参与试验项目的学区合作成立了高质量表现性评价研究所，团队的教师都来自这些学区。其次是内容领域专家的指导，州教育部门提供了每个内容领域的专家来指导共同表现性评价的开发，确保任务质量。再次是评价中心的技术审核，国家教育评价促进中心、全国闻名的非营利评价公司以及重要的技术合作伙伴，为共同表现性评价的技术进行审核与把关。最后还有州教育部门的项目领导的审核。由此，高质量的表现性任务库得以建立，且更为重要的是，在专业共同体中，教师获得了高水平的专业发展机会。我国当前实施基于核心素养的评价，同样面临着高质量表现性任务研制的问题。专业共同体、专业程序、表现性任务库为我们提供了可操作的路径。

同时，表现性评价的实施也是一个难题。尤其是对学生表现的判断的复杂性带来的主观性问题使表现性评价饱受诟病。新罕布什尔州一直致力于攻克这一问题，在研制评分规则、进行评分者培训和建立评分者之间的信度等方面，施行了严谨而专业的评价机制。参与指向核心素养的表现性评价试验的学校每年都会选送教师队伍参加

有关评价素养的培训,同时举行区域校正评分会议来建立评分者之间的信度和区域间评分的一致性。通过促进参与教师的专业发展,进一步辐射到学校层面,在学校进行校正评分会议。每个地区要求由学校评分的任务至少达到80%以上的评分者间信度。这些严谨而专业的评价机制,确保了表现性评价实施的质量,使得它可信、可行、可持续。专业制度和专业机制的建立与落实,在我国的基于核心素养的评价改革中将尤为迫切。

(三) 统整课程、教学与表现性评价

指向核心素养的表现性评价最重要的价值在于这个体系使每个学生的学习机会最大化。^[5] 这里的机会主要是指让学生获得有意义的知识和技能,且能进行深度理解,以使能运用到新的真实世界情境中,从而获得21世纪所需要的技能。因此,评价改革的落脚点是人的培养。这就涉及课程、教学与评价一体化的问题。也正因为如此,新罕布什尔州以评价为抓手,培养教师、学校课程领导者、学区课程领导者的课程能力,从而为整体实施基于核心素养的课程、教学与评价提供可能。无论是共同表现性评价还是学校表现性评价,新罕布什尔州都采用了统整在课程与教学中实施的做法。同样,我国实施基于核心素养的评价,需要课程、教学与评价的整体变革,将表现性评价统整在课程与教学中,共同指向培养完整的人。

同时,表现性评价本身的特质也使得它具有统整课程与教学的潜力。^[18] 其聚焦大观念的评价目标,正是课程与教学的关键目标;详细描述评分规则,让教师和学生更加清楚地知道努力的方向;让学生综合运用知识来解决真实世界中问题的评价任务,既是学习任务,又能更深层次更全面地了解学生的所知与所能。新罕布什尔州就是很好地挖掘并利用了表现性评价本身的特点,将其统整在课程与教学中进行实施,更好地促进教与学。一位桑伯恩地区的教师表示在课堂上实施表现性评价让他更加清楚学生前进的方向和他们已经走了多远。^[19] 如何利用评价更好地促进课程与教学,既是评价改革的问题,也是教师评价素养的问题。在我国,解决这个问题,更是任重而道远。

参考文献:

- [1] 钟启泉. 基于核心素养的课程发展: 挑战与课题 [J]. 全球教育展望, 2016, (1): 3-25.
- [2] Linda Darling-Hammond, Frank Adamson. Beyond basic skills: the role of performance assessment in achieving 21st century standards of learning [M]. Jossey-Bass, 2014; 23-29.
- [3] Skills P F S C. p21 Common core toolkit: A guide to aligning the common core state standards with the framework for 21st century skills [J]. Partnership for Century Skills, 2011; 48.
- [4] Bishop J M, Bristow L J, Coriell B P, et al. Assessing outside the bubble: performance assessment for common core state standards [J]. Online Submission, 2011; 127.
- [5] New Hampshire Department of Education. Moving from good to great in New Hampshire : Performance assessment of competency education (PACE) [EB/OL]. <http://education.nh.gov/assessment-systems/documents/overview.pdf>. 2016-04-08.
- [6] New Hampshire Department of Education. New hampshire performance assessment of competency education: An accountability pilot proposal to the United States Department to Education [EB/OL]. <http://education.nh.gov/assessment-systems/documents/pilot-proposal.pdf>. 2017-03-14
- [7] 崔允漭. 追问“核心素养” [J]. 全球教育展望, 2016, 45 (5): 3-10.
- [8] Partnership for 21st Century Skills. Framework for 21st century learning [EB/OL]. <http://www.p21.org/our-work/p21-framework>. 2016-12-10.
- [9] New Hampshire Department of Education. Tate model competencies [EB/OL]. http://education.nh.gov/innovations/hs_redesign/competencies.htm. 2016-07-05.
- [10] New Hampshire Department of Education. Competency validation rubric [EB/OL]. http://education.nh.gov/innovations/hs_redesign/documents/validation_rubric_for_course-level-competencies.pdf. 2016-07-05.
- [11] Performance Assessment Resource Bank. Assessment validation checklist [EB/OL]. http://performanceassessmentresourcebank.org/system/files/QPA_Tool_1_AssessmentValidationChecklist.pdf. 2017-03-15.
- [12] New Hampshire Department of Education. New hampshire performance assessment of competency

- education flexibility waiver insert [EB/OL]. <http://education.nh.gov/assessment-systems/documents/waiver-insert.pdf>. 2016-04-15.
- [13] Stanford center for Opportunity Policy in Education. New Hampshire Department of Education College and Career-Ready Accountability Options Performance Assessment of Competency Education (PACE) for Demonstrating Progress [EB/OL]. <https://edpolicy.stanford.edu/sites/default/files/events/materials/nhdoe-pace.pdf>. 2016-04-26.
- [14] Performance Assessment Resource Bank. Quality performance assessment common task teacher Directions [EB/OL]. http://cce.org/internal/documents/Tool28_QPACCommonTaskTeacherDirections.pdf. 2017-04-15.
- [15] Performance Assessment Resource Bank. Training with anchors protocol [EB/OL]. http://performanceassessmentresourcebank.org/system/files/QPA_Tool37_Training_With_Anchors_Protocol.pdf. 2017-04-15.
- [16] Sanborn Regional School District. Sanborn regional high school sample report card [EB/OL]. http://web.saul7.org/sample_srhs_rpt_card_transcript.pdf. 2016-04-08.
- [17] New Hampshire Department of Education. 2015 Performance assessment of competency education results by grade, level, and subgroups [EB/OL]. <http://education.nh.gov/assessment-systems/pace.htm>. 2016-04-08.
- [18] 周文叶. 中小学表现性评价的理论与技术 [M]. 华东师范大学出版社, 2014: 53-59.
- [19] New Hampshire Department of Education. Performance assessment of competency education reflections-Sanborn school district [EB/OL]. <http://mm.nh.gov/media/education/pace-reflections.mp4>. 2016-02-17.

(责任编辑: 刘启迪)

Core-Competence-oriented Performance Assessment

Zhou Wenye, Chen Mingzhou

(Institute of Curriculum and Instruction, East China Normal University, Shanghai, 20062, China)

Abstract: Curriculum reform based on core competencies drives the change of assessment system, prompting us to go beyond the traditional pencil and paper test that focuses on “Two Bases”, using and enriching performance assessment that can test students’ cognitive thinking, reasoning and the ability of applying knowledge to solve real and meaningful question. New Hampshire provides us a good example of performance assessment of core competencies. First, establishing a performance assessment system of core competencies. Second, specify core competencies, design the performance tasks that focus on core concepts, and create the performance task bank. Third, ensure the quality of performance assessment implement through professional procedures and mechanisms. It is urgent to promote assessment for carrying out the curriculum reform based on core competencies in China. We should strive to explore how to construct more balanced comprehensive assessment system, establish professional community and assessment mechanism, and integrate curriculum teaching and performance assessment.

Key words: core competencies; performance assessment