

表现性评价：历史、实践及未来

赵德成

(北京师范大学 教育学部, 北京 100875)

摘要:表现性评价能较准确评价学生在真实情境中的问题解决能力及相关素质,因而在20世纪80—90年代被明确提出后即受到广泛青睐。我国在新一轮课程改革启动之后开始有关探索,取得了不少经验,但也存在诸多亟待关注的问题,主要有:较少应用非纸笔形式的表现性任务,表现性任务不真实,任务与考察点脱节,缺乏可操作的评分细则。在未来要进一步推进表现性评价的应用,要在课程标准或课标解读中加强对表现性评价的推介;加强教师培训,使教师掌握相应实践技能;加强中高考命题改革,在命题中引入表现性评价。

关键词:表现性评价;真实性评价;课程改革

中图分类号:G42 **文献标志码:**A **文章编号:**1000-0186(2013)02-0097-07

表现性评价(performance assessment)通常要求学生在某种特定的真实或模拟情境中,运用先前所获得的知识完成某项任务或解决某个问题,以考察学生知识与技能的掌握程度,或者问题解决、交流合作和批判性思考等多种复杂能力的发展状况。^[1]它强调创设真实情境,即便是模拟情境,也必须能激发学生在真实情境中相似的反应,以考察学生在现实生活中分析问题和解决问题能力。因此,有人也将它称为真实性评价(authentic assessment)。^{[2](137)}表现性评价克服了传统标准化选择题测验仅能测试低水平知识和孤立技能的弊端,能测量出学生在真实世界中的复杂成就和情意表现,替代传统标准化测验成为影响和推动教学实践的重要评价方式。因此,也有人将其称为替代性评价(alternative assessment)。^{[3](177-178)}表现性评价在学习评价中受到普

遍重视和广泛采用,是国际中小学评价与考试改革的重要趋势之一。21世纪初,伴随着我国新一轮课程改革的启动与推进,表现性评价逐步走进我国中小学教师的视野,语文、数学、英语、科学等学科的教师在实践中进行了积极而有益的探索。但必须承认的是,表现性评价在我国的应用尚处于起步阶段,不少教师对表现性评价的理解存在偏颇之处,在实践环节上也出现了一些亟待关注的问题。本文先通过历史性分析阐释表现性评价的意义和价值,接下来结合案例分析当前我国应用实践中的问题,最后对未来推进表现性评价提出建议。

一、表现性评价的历史性分析

从历史的角度分析表现性评价,可以使我們了解表现性评价的起源和发展脉络,从而深化对

收稿日期:2012-11-28;修回日期:2012-12-03

作者简介:赵德成(1971),男,内蒙古赤峰人,教育学博士,北京师范大学教育学部副教授,博士生导师,主要从事教育评价、教育人力资源管理、学校评估等领域的研究。

表现性评价的认识与理解。表现性评价兴起于 20 世纪 80—90 年代,但实际上它的起源可以追溯到 20 世纪 30—40 年代,也就是教育评价发展史中的“泰勒时期”(the Tylerian, 1930—1945 年)。^①在这一时期,泰勒(Ralph W. Tyler)第一次提出现代的教育评价概念,将教育评价定义为考察重要目标达成程度的手段,并将其视为教学方案的一部分。他先运用评价概念协助俄亥俄州立大学的教师改进课程测验,不仅采用传统选择题测验考察学生对知识信息和定义的记忆,而且基于教学目标、结合生活情境编制包括非纸笔形式在内的多样化测验题目,考察学生的理解、推理、解释和应用等复杂能力,在生物、物理、化学、哲学、数学、会计和历史等多个课程中取得了突破性进展。基于这些经验,他撰写《编制成就测验的一般技术》一文,对中小学教育工作者提出建议,指出测验编制应包括以下几个主要步骤:(1)确定教育方案的目标;(2)根据行为和 content 对每个目标清楚地加以定义;(3)确定应用目标的情境;(4)设计给出应用目标情境的途径;(5)设计取得记录的方式;(6)决定评价方式;(7)设计获取代表性样本的方法。^[4](252-255)]在这七个步骤中,第三步和第四步都在强调测验情境的创设。泰勒认为,“如果一个人习得了一些东西,那么,这些东西就被内化,人们就期望他能在任何适宜的场合运用它们。所以,如果希望确定学生是否学到了什么,就应该关心它们运用所学知识的那些情境”,“应给出适当的情境以激发他们那些在正常情境中可能激发的反应”。^[4](253)]在“八年研究”(1933—1940 年)期间,泰勒进一步发展了这套基于情境的测验编制技术,开发出行为目标评价模式(Goals-Oriented Evaluation Model),评价者要设计情境和方法激发学生的行为,然后将学生表现与既定目标作对比,以评价学生达成教学目标的程度。不难发现,泰勒时期的测验和评价已经开始强调学生在真实情境中所表现出来的复杂情意和能力,表

现性评价的思想已初见端倪。

泰勒的观点和模式一直影响着教育评价理论和技术的发展。但在 20 世纪 80 年代以前,标准化测验在整个教育评价体系中始终占据着主导地位。这与当时美国教育的时代背景有关。1957 年,苏联发射斯普特尼克号(Sputnik)人造卫星后,美国国会于 1958 年通过《国防教育法》(National Defense Education Act),授权美国联邦政府拨款对州和地方以及个人提供实质援助,更新课程内容,确保培养出质量上和数量上均适用的人才。其他各种来源的经费也纷纷投入到教育领域。伴随而来的问题是,这些金额庞大的计划如果缺乏必要的监控,很可能在某些地方造成浪费。于是,教育绩效问责受到广泛关注,教育评价成为证明项目质量的一种手段,高利害标准化测验的时代开始,人们越来越多地依赖标准化成就测验评价教育的成功。

进入 20 世纪 80 年代,以认知主义学习理论为基础的建构主义学习理论在教学领域逐渐流行起来,并进而推动了教学和评价领域的改革。建构主义很多派别一致认为,人是主动的学习者,学习不是知识由教师向学生的传递,而是学生基于自己以往的相关经验,以自己的方式建构对于事物的理解。^{[5][6]}建构主义学习理论指导下的教学十分重视情境性,主张教学应使学习在与现实情境相类似的情境中发生,以解决学生在现实生活中遇到的问题为目标,学习的内容要选择真实性任务(authentic task),不能对其做过于简单化的处理,使其远离现实的问题情境。同时,教学的过程要与现实问题解决过程相类似,所需要的工具隐含于情境当中,教师不是把提前准备好的内容教给学生,而是指导学生探究与合作,自主学习和解决问题。新的教学理论和教学实践呼唤新的教育评价。于是,在教学改革的推动下,教育者开始对高利害标准化测验进行批判性反思,指出标准化测验存在诸多问题,主要有:(1)测验内容局限于低水平的知识、孤立的内容

① 美国知名教育评价专家马多斯(George F. Madaus)和斯塔佛尔比姆(Daniel L. Stufflebeam)在对教育评价学科发展历史进行回顾和分析时,将教育评价发展分成 7 个阶段,分别是改革时期(1900 年以前)、效率及测验时期(1900—1930 年)、泰勒时期(1930—1945 年)、纯真时期(1946—1957 年)、发展时期(1958—1972 年)、专业化时期(1973—1983 年)、扩张与整合时期(1983 年以后)。泰勒时期是其中的第三个阶段。参见: Madaus & Stufflebeam. 方案评估: 历史的回顾 [M] // Stufflebeam. 评估模型. 苏锦丽,等,译. 北京: 北京大学出版社, 2007: 3—22。

与技能；(2) 测验仅测出结果，没有考虑学习者的思维与问题解决技能；(3) 测验不能测量出学习者在真实的世界中应用理解的能力。^[7]在这样的背景下，表现性评价经历长期“蛰伏”之后被明确提出。由于这种方法能准确评价学生在真实情境中的问题解决能力、批判性思维能力、交流与合作能力等重要素质，在西方国家受到教育者、研究者和决策者的青睐，被广泛应用于中小学课程评价，成为替代传统标准化测验的新兴评价方式。

二、表现性评价在当前我国教育实践中的应用

新课程呼唤与之相适应的教育评价。在我国，2001年以来，一些研究者开始发表文章（如赵德成，2002^[8]；王小明，2003^[9]）或翻译外文书籍（如 Popham，2003^[2](135-154)]；Linn & Gronlund，2003^[3](177-202)]）推介表现性评价的基本原理、程序和技术。很多教师和教研员则行动起来，在各门课程的形成性评价和终结性评价中积极探索表现性评价的应用，取得了不少有益的经验，在一定程度上推动了课程改革，特别是评价改革的深入实施。但必须承认的是，表现性评价作为“舶来品”在我国并未受到足够的重视，有关技术性要求也没有被教育者所熟知和掌握，有关应用尚处于起步阶段，实践中存在着诸多亟待关注的问题。主要问题有以下几个方面。

（一）较少应用非纸笔形式的表现性任务

表现性评价要对学生在完成表现性任务过程中的表现情况进行观察与评估。表现性任务是表现性评价的构成要件。与传统标准化测验不同，表现性任务不再局限于纸笔形式，常见的表现性任务除纸笔形式测验之外，还可以有口头表述、辩论、模拟表现、表演、实验、调查以及项目等。近些年来，随着我国新课程的深入推进，有些教师在实践中进行了大胆的探索，设计了演讲、辩论、社会调查等多样化的表现性任务，但总体来看，非纸笔形式表现性任务数量相对不足，尚不能满足教学和评价的需要。在表现性评价的发源地美国，教师不仅在形成性评价中应用非纸笔性的表现性任务，而且在终结性评价乃至高利害评价中也经常应用。在美国康涅狄格

(Connecticut) 州，高中生要获得毕业证除了参加州统一举行的阅读、写作、数学和科学考试之外，还必须通过一些课程嵌入式的表现性评价。以科学课为例，学生在课程学习期间需要完成一次实验室实验和一次科学、技术和社会调查。学生必须自己提出研究假设，实施实验或调查，分析数据和撰写报告，以证明自己具备一定的科学推理能力。在我国，教师需要进一步更新观念，在学科教学中基于评价标准设计多样化的表现性任务，将其嵌入在课程实践中，既作为教学工具，也作为评价学生达成标准程度的评价手段予以使用。

（二）表现性任务不真实

真实性是表现性任务设计的基本要求。只有确保任务的真实性，才能使我们有机会观察到学生在真实生活情境中应用知识和解决问题的能力，才能确保评价的效度。美国教育评价专家威金斯 (Grant Wiggins) 指出，一项评价任务、问题或方案是否具有真实性，需要符合以下标准：(1) 是现实的；(2) 需要判断和创新；(3) 要求学生进行探索；(4) 重复或模拟成人接受“检验”的工作场所、公民生活和个人生活等情境；(5) 评价学生能有效地使用知识、技能完成复杂任务的能力；(6) 允许有适当的机会让学生去排练、实践和查阅资料，以得到有效的反馈并不断改进。^[10](19-25)]评价者设计的表现性评价任务要源于生活，反映真实生活世界对学生学习的要求。

由于长期以来，我国传统课程与生活在很大程度上是相互脱节的，尽管新课程反复强调要将学生学习与现代社会联系起来，但不少教师所设计的表现性任务仍然在真实性上有所欠缺。以数学课程为例，数学课程标准设计、教材编写和教学实施的一个重要原则是“人人学习有用的数学”、“人人学习生活中的数学”，在数学课程评价中应用表现性评价可以考察学生利用数学解释现实世界中现象及解决现实世界中问题的能力。但有些教师和教研员编制的表现性任务实际上并没有真正联系生活，任务情境是“生编硬造”出来的，与真实生活不符。例如，某市在中考数学试卷中出了这样一个表现性评价题目：

某学校师生到距学校 30 千米的郊外去春游，

一部分同学骑自行车先走,过了1.5小时,其余的人乘汽车出发,结果他们同时到达。已知汽车的速度是自行车速度的4倍,求两种车的速度。

如果把这道题目看作是考查学生数量关系理解能力的一般性题目,本无可厚非,但从表现性评价的视角来看,这一表现性任务缺乏真实性。因为在真实生活中人们几乎不会遇到这样的数量关系,而经常实际遇到的问题是将本题中的已知与求解互换,也就是说已知去春游地点的距离和两种车的速度,求解自行车出发多久后汽车再出发才能保证两种车同时到达。将生活中不会遇到的问题转化为表现性评价,显然不能有效考查学生运用数学解释生活现象和解决生活问题的能力。同理,在任何一门课程的评价中,选择或设计的任务情境不真实,就不能有效评价学生运用相关知识和技能分析与解决实际问题的能力。

(三) 任务与考查点脱节

真实性是表现性评价的必要条件,但不是充分条件,亦即“一项任务是真实的,并不意味着该任务就是有效的”^{[10](124)}。要保证教育评价的有效性,评价者必须基于要评价的目标,设计合适的任务激发学生表现出在这个或这些目标上的反应,使评价者通过观察和评价形成有效的结论。表现性任务必须与评价目的高度相关,评价要聚焦到要评价的目标上。如果任务与考查点相脱节,评价的效度就会降低。而这种现象在当前我国的评价实践中似乎比较普遍。

以语文课程为例,《义务教育语文课程标准(2011年版)》指出,口语交际是语文教学的重要内容,口语交际的教学目标是“具有日常口语交际的基本能力,学会倾听、表达与交流,初步学会运用口头语言文明地进行人际沟通和社会交往”。某版本初中语文教材中提供了这样一道表现性评价题目,让教师用它来评价学生口语交际学习的进展。

我们每个人都有自己心底的秘密,今天晚上回家后,请你和你的父母说说你心底的秘密。你会如何说呢?请你把它写下来。

这是一个真实性任务,但通过这道题目教师能考查到的主要是学生的书面语言表达能力,而不是口语交际能力,究竟某个班级或某个学生口语交际表现是否达成了所在学段课程标准的要求

及达成程度如何,我们无从评判。在这个评价设计中,表现性任务与要考查的目标是脱节的,直接破坏了评价的效度。世界经济合作组织(OECD)在国际学生评估项目(Programme for International Student Assessment, PISA)中的经验可以借鉴。^[11]为确保表现性评价的效度,该项目不仅严格根据双向细目表编制表现性任务,而且通过试题审查(从学生及编码者的视角进行审查)、认知性访谈(让学生回答问题,在回答过程中出声思考,同时结合个人访谈和小组访谈)和试点测试(在考试情境中试测)等方法,分析表现性任务能考查到的素质是什么以及是否与预测目标匹配。

(四) 缺乏可操作的评分细则

评分细则是表现性评价方案不可或缺的重要组成部分。表现性评价属于建构反应类(constructed responses)测量,学生在评价中建构的反应或提供的答案通常是开放、复杂的,需要采用一定的表现标准(performance standards)和评分细则(rubrics)加以评判。表现性任务是根据评价目标开发的,旨在评价学生达成目标的程度,将评价目标稍加转化,再结合特定任务分析学生较差和出色表现是怎样的,就可以形成表现性评价的标准及评分细则。这些标准和细则不仅可以使学生、教师、家长等利益相关者明确优秀的含义和成功的要素,而且对于最终评价分数的信度、效度和公平来讲都十分重要。^[12]

在我国,表现性任务的开发受到了教师和教研员的广泛关注,但相对而言,评分细则的编制尚未受到足够的重视,个别地方甚至还存在有任务无细则的现象。以历史课程为例,历史课程具有综合性特点,教师要“注重历史与现实的联系,使学生逐步学会综合运用所学知识和方法对历史和社会进行全面的认识”。^[13]历史考试不仅要考查学生对历史事实的理解和记忆,还要检测学生将历史与现实联系起来,“鉴古知今”,分析问题和解决问题的能力。^[14]表现性评价可以满足这种评价需要,很多教师和教研员设计出优秀的表现性评价题目。例如,某市高中教师在历史科期末测验中曾出过这样一道题:

中央电视台拟拍一部电视连续剧《重庆谈判》,假如你是编剧,请结合所学知识,展开合

理想象，为毛泽东设计一段到达重庆机场时的讲话稿。（150 字左右）

这道题拓展了评价内容的范围，不仅考查学生历史事件记忆和理解的程度，而且考查学生用历史材料解释历史和解决问题的能力，是一道比较典型的表现性评价题目。但令人遗憾的是，当我们询问命题者和使用者如何评分时，他们却声称没有明确的评分细则，评分由教师根据经验从历史知识、语言表达能力和创造性等方面进行。在表现性评价中，如果缺乏可操作的评分标准和细则，学生和教师就不知道努力的方向，评价者也无法诊断自己在达成目标过程中的优势与不足，更无法比较学生表现水平上的差异。命题者必须在表现性任务编制的基础上开发出相应的表现标准和评分细则，以克服表现性评价的评分主观性，确保评价的评分者一致性信度和效度。表现标准既可以是绝对性标准，也可以是相对性标准；评分细则可以是整体式评分细则，也可以是分项式评分细则。评分细则不仅要重点关注评价的能力，而且要具有可操作性。

三、未来进一步推进表现性评价的建议

设计良好的表现性评价能有效地测量传统评价难以测量的复杂成就和情意表现，能有力推动新课程教学理念在实践中的落实；但如果没有精心设计，表现性评价在设计与实施中存在上述提及的某种问题，其构想效度就无法保证，评价者难以对学生表现形成准确、有效的评判。为了在未来进一步推广表现性评价，充分发挥其导向、诊断和发展作用，我们提出如下一些建议。

（一）在课程标准或课标解读中加强对表现性评价的推介

评价处于教学的中心地位，是课程的重要组成部分。泰勒于 1949 年在《课程与教学的基本原理》（*Basic Principles of Curriculum and Instruction*）一书中开宗明义地指出，开发任何课程和教学计划必须回答四个基本问题：学校应该试图达到什么教育目标？提供什么教育经验最有可能达到这些目标？怎样有效组织这些教育经验？我们如何确定这些目标正在得以实现？基于上述四个基本问题，泰勒提出了课程实践的四个步骤或环节：明确教育目标，选择教育经验，组

织教育经验，评估教育效果。^[15]四个环节依次展开，互相影响，互为依据，是一个循环往复的过程。其中，评价是十分重要的一环，它考查教育目标的达成程度，具有双重作用。它既是对前一阶段教育的终结性评估，又是一次学习需求分析，为下阶段的课程开发和调整提供依据。从这一意义上而言，课程标准不仅要對课程目标进行清晰的界定，还要对如何评价课程实施效果进行解释和说明。而鉴于表现性评价在课程实践中具有广泛的学科适用性，在当前我国深入推进新课程的进程中又具有不可替代的特殊意义和价值，各门课程标准都应对其重点推介，以切实促进课程评价实践以及学生学习方式的变革。但对 2011 年教育部颁布的义务教育课程标准进行文本分析，我们却发现，很多学科没有提及表现性评价，仅有物理、化学、英语、历史与社会、思想品德等学科在“评价建议”部分在不同程度上涉及表现性评价，并且使用的表达方式不尽相同，如活动表现评价、实践活动评价、主题活动评价、项目评价以及探究能力评价等。为弥补这一缺憾，各学科课标组有必要在课标解读中加强对表现性评价的推介，具体的建议有：（1）向教师深入阐释表现性评价应用的必要性与意义；（2）介绍表现评价应用的基本程序及注意事项；（3）提供多样化示例说明表现性评价在本学科中的应用。

（二）加强教师培训，使教师掌握相应实践技能

教师是课程实践的主体，教师是否掌握表现性评价的基本原理、程序及要求，是否能在实践中主动、有效地应用表现性评价，是影响评价改革，乃至整个课程改革向纵深发展的关键。如果教师缺乏相应的理论素养和实践能力，势必会直接影响表现性评价的信度与效度，并间接地影响以评价为基础的课程决策。在国外，无论是教师职前培养还是在职培训，都将课程评价（classroom assessment，国内也有人译作课堂评价）作为一门重要课程，教师必须掌握包括表现性评价在内的各种基本评价技能，教师既要懂教学，还要懂评价，并在实践中将教学与评价有机结合起来，使两者互为依据和相互促进。而反观我国，教师教育工作者一直未予评价技能以足够的

重视,教师没有受到系统、细致、操作性的培训,不少教师在实践中“摸着石头过河”,所以出现了前文所述的诸多问题。为避免表现性评价的被误用,确保表现性评价的构想效度和实践效果,必须分层次加强教师培训。具体建议有以下三点。(1)课程专家和教研员培训先行。有些课程专家和教研员已面向教师推荐表现性评价,也组织过有关教师培训,但从中文核心期刊相关文章来看,很多人只是在简单引介国外教材中的有关阐述,并没有结合自己在本学科领域的实践探索,本土化程度不够。所以,课程专家和教研员作为培训者必须先接受培训,并在实践中深入探索,形成本土化经验,然后在此基础上开发教师培训的案例与教材。(2)在职教师的培训要采用“在做中学”的思路,边培训,边实践,使教师切实掌握设计、分析、反思和改进表现性评价的基本技能,使表现性评价自然融入教学过程并促进教学效果。(3)改进师范生教育。师范生在校期间要加强包括表现性评价在内的各种实践技能的培养,改革师范生课程设置,将课程评价(很多学校叫“教育测量与评价”)作为必修课程,并且更新课程内容,回应时代和新课程对教师提出的新要求,在课程中加强表现性评价等新兴评价方式的教育,使师范生跟上改革形势,胜任一线教学。

(三)加强中高考命题改革,在命题中引入表现性评价

中高考是一种高利害评价,对基础教育具有强有力的导向作用,中小学不同程度地存在“中高考考什么,教师就教什么;中高考怎么考,教师怎么教”的现象。利用这种心理,我们可以加强中高考命题改革,在命题中尝试引入表现性评价。这种改革尝试不仅有利于评估中小学教学在促进学生全面发展,特别是发展实践能力和创新精神上的实效,有利于选拔综合表现优秀的学生,更重要的是,它可以对基础教育教学与评价改革发挥积极的导向作用,引导教师关注表现性评价,通过加强表现性评价改进教学效果和促进学生全面发展。目前,英国、新加坡和澳大利亚维多利亚省以及我国香港地区已将平时表现性评价成绩纳入高考总分,作为高校录取学生的重要依据;^[16]美国、日本等国在高利害考试和学习质

量监控项目命题中也非常注重基于真实情境设计表现性评价,取得良好实效。^[17]我国可以借鉴有关经验,在中考和高考命题中加大改革力度,将表现性评价理念贯穿其中,通过真实性任务考查学生的复杂能力和综合素质。如果中高考命题发生变化,中小学教育工作者就会在实践中更加重视表现性评价,在形成性评价和终结性评价中进行大胆尝试。这势必会推动表现性评价在实践中的应用,并最终促进教与学方式的深刻变革,使新课程理念落到实处。

参考文献:

- [1] 赵德成,卢慕雅.新课程与学生评价[M].北京:高等教育出版社,2004:69.
- [2] Popham W J.促进教学的课堂评价[M].国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组,译.北京:中国轻工业出版社,2003.
- [3] Linn R L, Gronlund N E.教学中的测验与评价[M].国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组,译.北京:中国轻工业出版社,2003.
- [4] 泰勒.方案评价的原则[C]//瞿葆奎.教育学文集(第16卷):教育评价.北京:人民教育出版社,1989.
- [5] 张建伟,陈琦.从认知主义到建构主义[J].北京师范大学学报(社会科学版),1996,136(4):75—82.
- [6] 陈琦,张建伟.建构主义学习观要义评析[J].华东师范大学学报(教育科学版),1998(1):61—68.
- [7] Herman J L, Aschbacher P R, Winters L. *Selecting Assessment Tasks: A Practical Guide to Alternative Assessment* [M]. Alexandria, VA: Association for Supervision and Curriculum Development, 1992: 33—43.
- [8] 赵德成.表现性测验及其在中小学课堂评价中的应用[J].语文建设,2002(11):39—41.
- [9] 王小明.表现性评价:一种高级学习的评价方法[J].全球教育展望,2003(11):47—51.
- [10] Wiggins G.教育性评价[M].国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组,译.北京:中国轻工业出版社,2005.
- [11] 张雨强,张志红. PISA2006 科学试题的设计和开发及其启示[J].外国教育研究,2011,(2):59—65.
- [12] Arter J, McTighe J.课堂教学评分规则——用表

- 现性评价准则提高学生成绩 [M]. 国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组, 译. 北京: 中国轻工业出版社, 2005: 1—16.
- [13] 中华人民共和国教育部. 义务教育历史课程标准 (2011 年版) [S]. 北京: 北京师范大学出版社, 2012.
- [14] 张静. 新课程下中学历史学科学业评价初探 [J]. 历史教学, 2004 (9): 41—46.
- [15] 泰勒. 课程与教学的基本原理 [M]. 罗康, 张阅, 译. 北京: 中国轻工业出版社, 2008.
- [16] 冯生尧. 表现性评价纳入高考制度的必要性和可行性 [J]. 全球教育展望, 2007 (9): 22—26.
- [17] Pechione P L, Kahl S. *Through A Looking Glass: Lessons Learned and Future Direction for Performance Assessment* [M]. Stanford Center for Opportunity Policy in Education, 2010: 1—41.
- (责任编辑: 苏丹兰)

Performance Assessment: History, Practice and Future

ZHAO Decheng

(Faculty of Education, Beijing Normal University, Beijing 100875, China)

Abstract: Performance assessment has become a popular approach since it was put forward in 1980, because it can measure students' problem-solving ability and related qualities in reality. In China, some teachers began to use performance assessment after the New Curriculum was implemented. There are some technical problems that need to be noticed: a) non paper-pencil performance tasks are rarely used; b) performance tasks are not authentic; c) performance tasks are not closely tied to standards; d) there is no specific rubric for assessment. To extend the use of performance assessment in the future, we provided the following suggestions: a) guidelines on the use of performance assessment in curriculum standard explanation; b) reinforcing teacher education to help teachers acquire the related skills of performance assessment; c) using performance assessment in high-stake tests.

Key words: performance assessment; authentic assessment; curriculum reform