

# 中小學生综合素质评价研究综述

魏金宝 黄秦安 张勇

【摘要】综合素质评价的应用,起源于20世纪80年代的美国,经过四个阶段的发展,现已广泛应用于欧美各级各类学校的评价中。我国自1999年国务院发布《关于深化教育改革全面推进素质教育的决定》以来,政府多次出台政策积极推进中小學生综合素质评价改革,各地教育部门积极探索中小學生综合素质评价,但依然存在着思想理论基础不足、评价模式可操作性不强、方法技术不科学、公信力不足等问题。未来随着综合素质评价的学术研究向纵深发展,将更加注重评价工具的开发,保障机制问题将日益突出。

【关键词】教育评价;综合素质评价;教育评价改革

【中图分类号】G424.74 【文献标识码】A

【文章编号】1673-1654(2016)03-073-007

从历史源头上来看,对人的综合素质认识在人类教育史上历史悠久。古希腊教育确立语法、修辞、逻辑论辩和算术、几何、天文、音乐七门核心学科,即所谓“自由人的学科”,并确立“勇敢、智慧、正义、节制”为四大美德,这为中世纪大学的综合评价奠定了基础。在我国,刘劭(三国魏)在《人物志》中提出“八观”“五视”及“接论”的考评方法,根据不同情境中的反应,全面判断人格品质。《新唐书》卷四十五《选举志》:“择人之法有四:一曰身,体貌丰伟;二曰言,言辞辩正;三曰书,楷法遒美;四曰判,文理优长。四事皆可取,则先德行,德均以才,才均以劳。”这就是所谓的身言书判,为综合考核的标准。这些中外思想其实代表了综合评价思想的萌芽,至卢梭“自然主义教育观”培养“身心调和发达的人”的思想最终成型。

## 一、教育评价及综合素质评价概述

在人类教育产生之始,原生态的教育实施、评价、咨询同时产生,三者一直相伴发展。教育实施的成效需要作出评价,评价结果需要咨询来转化,进而促进教育实施的改进与发展。教育的这种内在原生性功能结构,决定了教育评价由测量(或计量)、评价、诊断三种基本功能构成一个完整的基础系统结构。教育评价的其他功能(如导向、预测、选拔等)则是在此基础上延伸、转化和合成出来。

相对教育实施而言,人类对教育评价的研究相对滞后,一直与人才评价及考试(含选拔官吏、职业人才等等)混杂在一起。从已有文献来看,直到十九世纪中期,教育评价与考试及人才评价等界限不清,综合素质评价与学业评价也一直没有得到明确的思

作者简介 魏金宝,中学一级教师,陕西师范大学数学与信息科学院;黄秦安,博士,教授,博士生导师,陕西师范大学数学与信息科学学院。陕西西安,710119。张勇,院长,公众教育研究院。北京海淀,100082。

本文系西安市2015年基础教育研究重大课题“基于提升教育质量的课堂教学建模研究”阶段性研究成果,项目编号2015ZB-ZD02。

想与理论上的切分。

十八世纪中叶让-雅克·卢梭(Jean-Jacques Rousseau)提出的“自然教育思想”及二十世纪初叶约翰·杜威(John Dewey)提出的“社会教育思想”(实用主义教育),为综合素质及其评价研究提供了思想基础和解释依据。

二十世纪初,由于实验心理学、教育统计学及智力理论的发展,教育评价与考试、人才评价等方得到思想与理论上的逐步划分。二十世纪初至四十年代,随着对非智力因素和人才评价的认知不断深入及“泰勒原理”的应用,才使心理学及教育学界等开始了对非学科学习素质的理论研究——综合素质理论研究开始萌芽。

从产生源头上看,人类对综合素质的认识与评价同时来源于教育评价与人才评价及对非智力因素的认识和应用(对非智力因素的认识及研究来源于对智力的认识及对其理论的反思)。综合素质评价的研究既属于教育评价并应用于教育领域,也属于人才评价并应用于人力资源领域。从趋势来看,对综合素质的研究与评价会逐步扩展和深化到人类的其他领域,其应用会越来越广泛和深入。

综合素质及其评价的研究,最初关注的是人除知识学习之外的其他能力、技能等因素,后发展至关注人的非智力因素。随着学业评价、考试、人才评价等研究与应用的深入,并受卢梭自然教育思想和杜威社会教育思想,及人才理论、建构理论、绩效理论及全面发展观、个性观等的影响,最终落实到了关注人的基础成长和成才基础的研究上。而我国对综合素质及其他评价的研究,教育部在《关于积极推进中小学评价与考试制度改革的通知》(教基[2002]26号)中最终界定为“道德品质、公民素养、学习能力、交流与合作能力、运动与健康、审美情趣”六个方面,称之为学生的基础性发展目标。

## 二、美国学生综合素质评价发展历程

学生综合素质评价的普遍应用起源于20世纪80年代的美国。其产生的背景是自20世纪60年代起,美国兴起的对标准化考试的反抗。标准化考试在美国不仅对教育决策有着重大影响,而且对雇佣

决策也产生了重大影响。由于标准化考试和学校课程的关联度较低,美国学校也陷入了为应试而忽略常规课程教学的情况。因此,很多学者对这种标准化考试滥用的现象提出批评。例如,1962年,担任多年考试顾问的Banesh Hoffman在其《测试的暴政》(The Tyranny of Testing)一书中,通过展示标准化考试本身存在的一些缺陷提出对考试的至高无上性的质疑;1984年,纽约大学教授、教育历史学者Diane Ravitch在肯定标准化考试用处的同时也指出,标准化考试所测量的内容过于狭窄,不能全面反映学生的能力,也会对学校课程的实施产生不良影响。正是在这种背景下,“表现性评价”(performance assessment)、“真实性评价”(authentic assessment)或“情境性评价”作为弥补标准化考试缺陷的新的评价方式被提出来,并被中小学广泛采用。

表现性评价,是一种要求学生通过完成特定的任务或生产一定的产品来展示其掌握的特定技能及能力的评价方式。学者认为,适用于表现性评价的项目可以有:设计和执行试验、撰写需要学生整合、应用信息的论文;需要和其他人合作的项目;展示对某种设备或技能掌握的熟练程度;建模;编制、解读以及使用地图;收藏;创作批评性文章、诗、小故事;演讲;乐器演奏;口语展示;做档案袋;体育技能等等。

表现性评价中最主要的一种形式是档案袋(portfolio)。目前全美采用这种评价方式的学校已经跨越幼儿园到大学,受评价群体从学生到行政人员,应用非常之广泛。

通过梳理美国1987年至2004年300多份关于表现性评价的专著和期刊文章,可以发现表现性评价在美国主要经历了以下几个发展阶段:

### 1. 1987-1992年,表现性评价快速发展期

表现性评价的概念一经提出,就引发了很多州的响应。如加州从1987年开始全州推行写作评估,并且在科学与历史方面开发表现性评价;康涅狄格州教育进步项目在1989年对科学、外语和商业教育实施了表现性评价;佛蒙特州在1989年开始通过立法批准基金来实施档案袋评价方法,作为对当时标准化考试成绩的一种补充。这时美国国内也有几个著名的国家层面的评价项目,如匹兹堡大学的新标

准項目、國家教育和經濟中心的项目、哈佛大學的零點計劃等等<sup>[1]</sup>。

檔案袋評價方式作為表現性評價的一種主要方式，一直是教師、學者津津樂道的话题。這個階段所呈現的特点是，對於檔案袋評價的介紹還處於入門階段，許多文章介紹檔案袋評價是什麼，如何操作<sup>[2-3]</sup>；也有很多學者對如何在教學中應用檔案袋評價法以及取得的成果加以介紹<sup>[4-8]</sup>；檔案袋評價的主要功能是用於指導教學和幫助學生自我掌握學習<sup>[9]</sup>；應用的領域主要是寫作，其他學科較少；政府主導與教師自發採用這種評價方式兼而有之。到1992年，表現性評價的研究關注點主要集中在評價任務的真實性和現實性、學生表現的測量標準、新評價打分程序以及培訓評分者如何打分、如何使用這種評價上<sup>[10]</sup>。

## 2. 1993 – 1995 年，表現性評價全面推開階段

在這個階段，檔案袋評價方式已經在全美中小學中廣泛應用，並且向下擴展到幼兒園<sup>[11]</sup>，向上擴展到大學<sup>[12]</sup>。評價對象由原來的學生擴大到教師及預備教師、行政人員。評價科目也由最早的寫作、數學發展到音樂、歷史、科學等學科中。這期間最為突出的特點有兩個，第一是將檔案袋評價和學生的畢業、教育問責、教師的僱傭、評價以及教師職業發展掛鉤。如紐約中心公園東部高中的實驗班學生，必須口頭匯報7個檔案袋任務才能准予畢業<sup>[13]</sup>。一項針對俄克拉荷馬、猶他以及新墨西哥州的偏遠和小型學校校長和教師的調查顯示，無論是普通學校還是特殊教育學校的校長和老師們都非常支持在教師聘用和評價方面採用檔案袋方式<sup>[14]</sup>。第二，開始出現了檔案袋評估的技术模型。如 Moya、S. S. 和 J. M. O'Malley 提出了 ESL 教學檔案袋評估模型。在這個模型中，評估活動包括相互關聯的六個層次：（1）明確檔案袋目的和程序的焦點；（2）計劃檔案袋內容；（3）設計分析標準和準則；（4）為指導性應用做準備；（5）確定驗證檔案袋信息準確性的程序；（6）按模型實施評價<sup>[15]</sup>。而 Paulson、F. L. P. 和 Pearl R 提出了檔案袋評估的認知模型（CMAP）<sup>[16]</sup>，該模型以“檔案袋相關者、過程性和歷時性”為相互作用的三個維度，以檔案袋發展的四個階段以及學生對檔案袋的四種反思水平為評估指標來評價檔案袋的質量。

## 3. 1996 – 2000 年，檔案袋評價法開始步入數字化和電子化時代

在此期間，檔案袋評價的應用範圍進一步擴大。體育教育也開始採用檔案袋評價方式。教師及預備教師的專業發展中大量應用這類評價方式。這個時期最突出的特點就是檔案袋評價法的數字化和電子化。早在1993年，美國就有學校開始嘗試使用數字化學生檔案，如美國肯塔基州的東部高中和新罕布什爾州的賽爾高中<sup>[17]</sup>。隨著計算機和網絡的發展，從1996年開始，電子化或數字化的檔案袋逐漸成為美國教育界熱門討論的話題。電子檔案的設計、操作步驟、存在的問題等等都有文獻論述<sup>[18-21]</sup>。數字化檔案袋因其存儲容量大、便於需要者獲取等優勢，為越來越多的學校所採用。

## 4. 2000 – 2004 年，電子檔案袋評價的深入階段

在這個階段，檔案袋評價的應用範圍繼續擴展，體育學科、諮詢師的教育以及職業晉升都開始採用這種方式。除此之外，電子檔案袋評價的研究開始有所深入。如出現了“5\*5 電子檔案袋建立的模型”以及對電子檔案進行網絡評分的操作細則<sup>[22-23]</sup>。

通過對表現性評價在美國的發展歷史的梳理，有以下發現：

第一，美國對學生能力的評估在原有的標準化測驗基礎上，又增加了表現性評價。表現性評價又以檔案袋為主要方式。將表現性評價結果納入對學生能力和素養的綜合考查範圍，其背後的理念是：學生的能力和素養由多方面構成，任何一種評價方式都有其局限性，不能僅依靠一種評價方式得出的結論來評價學生。評價應該根據學生能力和素養的不同方面選取恰當的評價方式和手段。

第二，表現性評價強調評價任務的仿真性，強調完成任務所需的能力和素養應該是成人在社會中解決類似問題所需要的能力，以此來反映學校教育是否實現了教育目標。

第三，檔案袋作為表現性評價的一種主要方式，在美國有著廣泛的應用基礎。與標準化考試提供的評價結果相比，檔案袋評價法有其不可替代的優勢，尤其表現在：促進學生自我認識能力的提升方面；良好素養的培養方面；為家長、招生單位提供更全面的資訊方面。借助信息化手段和網絡制作的電子檔案

袋,不仅具有指导学生成长的作用,存储量大和便于获取的优势,还有助于学生在需要时向各个机构,包括招生单位呈现。

第四,档案袋作为学生综合素质能力展现的一种重要手段,需要根据评价的目的精心设计档案袋内容,需要有科学的评分标准和评分量表,需要对评价结果加以解释并对应用做出指导,同时还需要对评分者进行培训,规范评分过程以保证评分信度。而这些在美国都已经具有非常成熟且丰富的可借鉴的经验。

### 三、综合素质评价在我国的学理讨论现状

1999年国务院《关于深化教育改革全面推进素质教育的决定》(以下简称《决定》)(中发[1999]9号)中提出,“高考科目设置和内容的改革应进一步突出对能力和综合素质的考查。”这是“综合素质”一词首次出现在相关政策法规中。2002年教育部《关于积极推进中小学评价与考试制度改革的通知》(以下简称《通知》)(教基[2002]26号)和《基础教育课程改革纲要(试行)》(以下简称《纲要》)(教基[2001]17号)是学生综合素质评价的直接的 policy 依据。

政策颁发之后,学者开始对综合素质评价进行了广泛而深入的讨论,其中最具代表性的探讨主要集中在以下一些方面:

#### 1. 对综合素质评价的总体看法

我国学者认为,综合素质评价要用发展的眼光看待评价对象,尊重评价对象的主体性和差异性,强调评价对象的自我评价、自我反思和自我监控,注重评价对象发展变化的过程,使评价成为促进评价对象不断发展的教育活动。评价的根本目的是为了更地促进学生的全面、均衡成长。评价方法要多样化,除考试或测验外,还要使用观察、访谈等多种科学有效、简便易行的评价方法。评价不仅要注重结果,更要注重发展和变化过程,把终结性评价与形成性评价有机地结合起来。重视学生、教师和学校在评价过程中的作用和主体地位,使评价成为教育行政部门、学校、教师、学生和家长共同积极参与的交互活动<sup>[24]</sup>。综合素质评价的出路在于:进一步加大

政策研究和制度建设的力度;营造和构建综合素质评价健康可持续发展的基础环境;将综合素质评价归置到表现性评价范畴,以“个性”与“发展”为取向,以质性评价为方法,建立与综合素质评价属性、价值与功能相适应的评价体系、机制和文化<sup>[25]</sup>。综合素质评价应遵循发展性原则、主体性原则、导向性原则、差异性原则和操作性原则。评价内容上包括模块修习记录、基本素质评价、实验操作考查、信息技术等级考试,以班级为单位成立评价小组,采用质性评价和量化评价相结合的评价方法,评价结果以写实性文字记录在普通高中学生电子档案中,作为学生能否毕业的依据和高校招生录取的重要参考<sup>[26]</sup>。

#### 2. 对综合素质评价本质的讨论

综合素质评价既是一种评价观,又是一种评价方式。作为一种评价观,它是素质教育评价体系的基本价值取向。作为一种评价方式,它与中考、高考等外部评价互动、结合,共同构成素质教育评价体系的基本内容。综合素质评价的本质是个性发展评价,也是真实性、过程性评价,同时还是内部评价<sup>[27]</sup>。综合素质评价应依据卢梭的“自然教育思想”和杜威的“社会教育思想”,关注人的基础成长和成才基础,应立足于成长规律、认知发展规律和成才规律,以“伦理、道德与品质;审美、情感与价值;健康、运动与操作;独立、交流与合作;认知、探究与应用”为目标,面向学生全面成长和个性化发展的“德、智、体、美、实践、创新”六大方面进行综合评价。

#### 3. 对综合素质评价关键的讨论

综合素质评价的关键是构建相关的配套制度和采用科学的测量评价目标、标准和技术。就配套制度而言,如申诉制度、听证制度、公开制度、问责制度、监督制度等,加强全民的法制意识和诚信文化建设,才能保障中小素质评价工作公平公正科学有序地进行<sup>[28]</sup>。完善中小学综合素质评价需要加强沟通交流,淡化选拔功能;推进教育资源分配改革,建立监管机制;积极帮助引导,实施利益补偿;加快配套制度建设,减少信息损耗等<sup>[29]</sup>。制度、机制(模式与管理)和测量评价目标、标准及技术,是决定综合素质评价实施的核心和关键。在制度、机制基本建立并科学的前提下,测量评价目标、标准和技术是决定教育评价改革尤其是综合素质评价的关键

因素<sup>[30]</sup>。

#### 4. 對綜合素質評價困境的討論

綜合素質評價的困境在於：政策文本中，“綜合素質評價”概念的內涵和功能尚處在猶豫未決狀態，尚未進入真正的教育評價領域，缺乏屬於自己的評價理論和方法<sup>[31]</sup>。綜合素質評價存在著核心概念界定不清、評價維度混亂、評價指標不合理、評價方法使用不當、監督機制不到位、使用目的不明確等問題<sup>[32]</sup>。中小學實施的綜合素質評價還存在著價值偏差導致的目標與方法錯位、權力分配造成的教育不公平、心理定勢引發的抵制行為、實際障礙減弱改革的力度等問題<sup>[33]</sup>。

### 四、國內綜合素質評價實踐的現狀與問題

#### 1. 政府多次出台政策積極推進中小學生綜合素質評價改革

1999 年在國務院的《決定》中明確使用“綜合素質”一詞，要求“高考科目設置和內容的改革應進一步突出對能力和綜合素質的考查”。2002 年，教育部頒發《通知》，指出要建立以促進學生發展為目標的評價體系，該體系的評價標準主要包括基礎性發展目標和學科學習目標兩個方面。基礎性發展目標主要包括道德品質、公民素質、學習能力、交流與合作能力、運動與健康、審美情趣。2013 年又頒發了《關於推進中小學教育質量綜合評價改革的意見》，同時給出了“中小學教育質量綜合評價指標框架（試行）”，對學生的綜合素質評價又提出了品德發展水平、學業發展水平、身心發展水平、興趣特長養成四個方面 16 個關鍵指標。較之 2002 年給出基礎性發展目標，2013 年給出的指標框架中增加了實踐能力、創新意識、潛能發展、情緒行為調控等新的評價內容。2014 年 9 月 3 日國務院《關於深化考試招生制度改革的實施意見》明確了綜合素質評價與考試招生制度實行“硬掛鉤”。上述政策的發布，表明了政府要扭轉過去單純以學生考試成績和學校升學率評價中小學教育質量這種做法的決心和意志，從中可見我國對綜合素質內涵的認識也不斷深化，對中小學生綜合素質評價的目的和定位由不明確變明確。

#### 2. 各地教育部門積極探索中小學生綜合素質

#### 評價

面對當前考試制度不能適應社會需要的狀況，國內學者、機構、政府紛紛研究和探索綜合素質評價的改革之路。國內有關課題紛紛立項，眾多學者呼籲將中小學生綜合素質評價與考試制度掛鉤。實證性研究遍地開花，如重慶的學生檔案袋評價、北京市四個維度評價、公眾教育研究院的 ICTS 綜合素質評價、南昌的綜合素質等級劃分等，甚至一些鄉鎮中學也積極參與到這次綜合素質評價的改革中。改革取得一定成效，如對綜合素質有了進一步的認識，德、智、體、美等方面達成共識；改變了包括教師在內的一些人的評價觀念，更加注重人的全面發展和個性化差異；評價指標得到一定程度的細化，一些地方開發了二級指標等。綜合素質評價表現出評價內容綜合、主體多元和方式多樣共同特點。並且有對過去的评价轉向對未來的預測，實行發展性評價，強調社會實踐和信息技術的應用。這些都表明我國的中小學生綜合素質評價正在逐漸與國際接軌，為進一步推進奠定了一定的基礎。

中國教育學會作為國內教育界最大的學術團體組織，成立“中國教育學會中小學教育質量綜合評價改革實驗區辦公室”，凝聚了諸多海內外教育評價專業組織和卓有建樹的專家，致力於為中小學綜合素質評價提供專業支持和技術服務。

#### 3. 我國中小學生綜合素質評價方面存在的主要問題

第一，從課題立項級別看，立項級別都不高，現有綜合素質評價課題研究多是地市級或國家級、省部級課題的子課題，缺少以此為主體的國家級重點重大課題。

第二，綜合素質評價的思想基礎和理論基礎不足，思想基礎和理論基礎缺乏足夠論證，評價模式可操作性不強，評價的標準和指標體系不夠完善。方法技術不科學，存在一定主觀性、隨意性和隨機性。結果不夠全面，缺少可比性，評價過程不規範。

第三，綜合素質評價的公信力不足，評價過程不透明，出現公平誠信危機、監督管理不到位。

第四，綜合素質評價是一項非常專業的工作，需要具備相應專業素養的專業人員來具體實施，但是目前具備專業能力的评价人員嚴重缺乏。

第五 综合素质评价的评价过程操作复杂,实施过程中加重了教师负担。

## 五、综合素质评价研究的未来趋势

### 1. 综合素质评价学术研究向纵深发展

中小学生综合素质评价一直以来都是教育研究人員热议的话题,随着新一轮教育质量综合评价改革的深入推进,尤其是随着教育部《关于2013年深化教育领域综合改革的意见》(教改[2013]1号)、中共中央《关于全面深化改革若干重大问题的决定》、国务院《关于深化考试招生制度改革的实施意见》(国发[2014]35号)、教育部《关于加强和改进普通高中学生综合素质评价的意见》(教基二[2014]11号)、教育部《关于全面深化课程改革落实立德树人根本任务的意见》(教基二[2014]4号)、教育部《关于深入推进教育管办评分离促进政府职能转变的若干意见》(教政法[2015]5号)等一系列政策文件的发布,必然从政策和现实需求两方面推动综合素质评价的研究向纵深发展,出现更多的具有一定专业水准的教育第三方评价机构,专门从事教育评价研究和业务开展。

### 2. 更加注重综合素质评价工具的开发

以往对综合素质评价的研究大多集中在理论探讨层面,相信随着政策和现实需求两个层面的推动,综合素质评价研究将从过去的集中在理论探讨层面转向理论探讨和工具开发并重的局面。未来研究将更加集中于进一步破解综合素质评价的理论、技术与方法难题,建立科学有效的中小学生综合素质评价标准及细化的指标体系,开发一系列有效、可靠、可信的测量及评价工具等方面。

### 3. 保障机制问题日益受重视

随着新的高校录取招生确定为“两依据一参考”模式,综合素质评价在高校录取招生中所占的比重加大,综合素质评价公信力不足、过程不透明、监督管理不到位的公平诚信危机问题将日益突出,首当其冲的问题是设计一套适用性强、易操作的实施模式以及保障机制,使中小学生综合素质评价在公开、公正、公平、规范的实施和操作中,获得科学、可信、可靠、易获取的结果,从而使中小学生综合素质评价

结果能够真正反映中小学生的基础发展目标达成状况,能够成为衡量学生是否达到毕业要求的重要依据,并且能为学校择优录取学生提供重要参考。

## 参考文献:

- [1][10] French, R. L. Portfolio Assessment and LEP Students. In Proceedings from the Second National Research Symposium on Limited English Proficient Student Issues 1991: Focus on Measurement and Evaluation, United States Department of Education: Office of Bilingual Education and Minority Language Affairs, 1992, (1): 249 - 285. Washington, DC: United States Department of Education.
- [2] Vavrus L. Put Portfolios to the Test [J]. Instructor, 1990, 100: 48 ~ 53.
- [3] Arter, J. A. & Spandel, V. Using Portfolios of Student Work in Instruction and Assessment [J]. Educational Measurement: Issues and Practice, 1992, 11, 36 - 44.
- [4] Dennie Palmer Wolf, Karen Volz Bachofer, Jean Slattery. Portfolio Assessment [M]. Corwin, 2004.
- [5] Stump Jongsma, K. Questions & Answers [J]. The Reading Teacher, 1989: 264 - 265.
- [6] Campbell, J. Laser Disk Portfolios: Total Child Assessment [J]. Educational Leadership, 1992, 49(8), 69 - 70.
- [7] Frazier, D. M. and F. L. Paulson. "How Portfolios Motivate Reluctant Writers." [J] Educational Leadership 1992, 49(8): 1992, 62 - 65.
- [8] Hettterscheidt, J., L. Pott, et al., "Using the Computer as a Reading Portfolio." [J] Educational Leadership 1992, 49(8): 73.
- [9] Hansen, J., "Literacy Portfolios: Helping Students Know Themselves." [J] Educational Leadership 1992, 49(8): p66 - 68, 1993.
- [11] Diffily, D. F., Pamela. "Sociodramatic Play: Assessment through Portfolio." EDRS Price.
- [12] Aitken, J. E., "Assessment in Specific Programs: Employment, Program, and Course Student Portfolios in Communication Studies." EDRS Price, 1994.
- [13] Darling - Hammond, L. A., Jacqueline. "Graduation by Portfolio at Central Park East Secondary School. A Series on Authentic Assessment and Accountability." EDRS Price, 1994.
- [14] Bull, K. S. A. O. "Portfolio Assessment in Teacher Evaluation: A Comparison of the Perspectives of General and Special Education Administrators and Teachers." EDRS Price, 1994.
- [15] Moya, S. S. and J. M. O'Malley. "A Portfolio Assessment Model for ESL [J]." The Journal of Educational Issues of Language Minority Students 1994: 13, 13 - 36.
- [16] Paulson, F. L. P., Pearl R. "A Guide for Judging Portfolios." EDRS Price, 1994.

- [17] [18] Niguidula, D. "Picturing Performance with Digital Portfolios [J]. " Educational Leadership ,1997, 55(3): 26-29.
- [19] Baron, C. Creating a Digital Portfolio [M]. Indianapolis: Hayden Books ,1996.
- [20] Georgi, D. C. , Judith, "Digital Portfolios: A Confluence of Portfolio Assessment and Technology [J]. " Teacher Education Quarterly , 1998, 25(1): 73-84.
- [21] Lennex, L. C. "Digital Portfolio Production Among Preservice Secondary Teachers. " EDRS Price ,1998.
- [22] Barrett, H. C. The "5-by-5" Model of Electronic Portfolio Development Overview of Technology Options 2000.
- [23] Chatel, R. G. "Portfolio Development: Some Considerations. " EDRS Price 2001.
- [24] 董奇, 赵德成. 发展性教育评价的理论与实践 [J]. 中国教育学刊, 2003, (8): 18-21, 45.
- [25] 杨九诠. 综合素质评价的困境与出路 [J]. 华东师范大学学报(教育科学版), 2013, (6): 36-41.
- [26] 黄志红. 新课程背景下普通高中学生综合素质评价的研究与构想 [J]. 课程·教材·教法, 2006, (11): 17-22.
- [27] 李雁冰. 论综合素质评价的本质 [J]. 教育发展研究, 2011, (24): 58-64.
- [28] 靳玉乐, 倪哲. 中小学综合素质评价相关制度的建立及其运用 [J]. 西南师范大学学报(自然科学版), 2014, (2): 149-153.
- [29] [33] 靳玉乐, 樊亚桥. 中小学实施综合素质评价的意义、问题及改进 [J]. 教育研究, 2012, (1): 69-74.
- [30] 姚春艳, 张勇. 新型教育评价技术破解教育评价改革难题 [J]. 考试(理论实践), 2014, (03): 8-11.
- [31] 杨九诠. 综合素质评价的困境与出路 [J]. 华东师范大学学报(教育科学版), 2013, (6): 36-41.
- [32] 洪志忠. 美国高中综合素质评价对我国的启示 [J]. 当代教育科学, 2010, (24): 17-19.

## Review of Research on the Students' Comprehensive Quality Assessment

Wei Jinbao<sup>1</sup> Huang Qin-an<sup>1</sup> Zhang Yong<sup>2</sup>

1 College of Mathematics, Shanxi Normal University, Xi'an, Shanxi, 710119

2 The Public Education Research Institute Beijing, 100082

**Abstract:** Comprehensive quality assessment originated in the United State of America in 1980s', and then it experienced four stages and is being widely applied for all kinds of schools in America and Europe. China central government has carried out many measures to promote the students' comprehensive quality assessment since the release of Resolve to Deepen Education Reform so as to Promote Quality Education by the State Council and education departments actively have been exploring the ways to make quality education assessment, but problems such as weak theoretical foundation, assessment models hard to be carried out, unscientific methodology and inefficient public incredibility do exist. With the vertical development of comprehensive quality assessment, emphasis will be laid on developing comprehensive quality assessment tools, and so problems on guarantee system will be prominent.

**Key words:** Educational Assessment, Comprehensive Quality Assessment, Educational Assessment Reform

(责任编辑: 刘清华)