

# The Historical Retrospect of Students' Comprehensive Quality Evaluation in Western Countries

Hualing Luo

Academy of Educational Sciences, Department of Education of Yunnan Province, Kunming, Yunnan, 650223, China

## Abstract

This paper applied literature method and historical method into the history development literatures of the comprehensive quality evaluation of students to analysis historical facts. It was an amusing finding that the thought source and practice starting point about the comprehensive quality evaluation of students can retrospect from the emergence of education phenomenon of human society to the evaluation of students which opened the road of the comprehensive quality evaluation of students.

## Keywords

the West; the comprehensive quality evaluation of students; historical traceability

## 西方国家学生综合素质评价的历史追溯

罗华玲

云南省教育厅教育科学研究院, 中国·云南 昆明 650223

## 摘要

应用文献法和历史法对西方国家学生综合素质评价的发展历史文献进行分析,发现学生综合素质评价的思想源头和实践起点从历史追溯到人类社会的教育现象产生之时,教育的产生,衍生出面向学生的测评,从学生的测评开启了学生综合素质评价之路。

## 关键词

西方国家; 学生综合素质评价; 历史追溯

## 1 引言

本研究借鉴“当下的历史”研究方法,将西方国家学生综合素质评价“回到历史”,还原其历史发展的基本脉络。关于“当下的历史”,其提出者是美国教育社会学家 Thomas Popkewitz 就历史分析框架而演绎出的基本观点,意在虑及文化理性下的知识认知皆是在社会政治、经济和历史情景的多元交互建构,力推对当下的“真”皆有社会历史根源的脉络,或盘根错节,或剥离斑驳,光怪陆离可见看似自然之物背后的运作逻辑和秩序。西方国家学生综合素质评价的历史发展脉络可划分为四个阶段<sup>[1]</sup>:萌芽启蒙时期、迂回发展时期、测量大兴与目标达成的价值判断时期、评价主体本位的相互协商和过程建构时期。

**【基金项目】**2021年度云南省教育科学规划项目立项课题基础教育专项“新时代边疆民族地区学生综合素质评价研究”(项目编号:BFJC21041)的研究成果。

**【作者简介】**罗华玲(1976-),女,中国湖南邵阳人,博士,助理研究员,从事教育学研究。

## 2 萌芽启蒙时期

西方最早的学生评价可追溯至古希腊“三哲”和古罗马时期,一个天才般的创造者,一个天才般的实践者。无论苏格拉底还是柏拉图或是亚里士多德在学生评价上都有其评价思想萌芽和评价行为的启蒙。其中,苏格拉底以其“产婆术”教学法在教学实践中,通过自然观察与学生的交流对话,判断并评价学生的思维以及语言表达的变化发展。

柏拉图则通过其所著的《理想国》阐释其学生评价思想,认为在理想国中有不同禀赋的学习者,学生评价要在不同年龄阶段区分出不同学习天赋的小铜铁人(如农民、手艺人等劳动者)、小银人(军人)以及最高层级段位的小金人(哲学家),评价标准则是围绕不同年龄阶段学习者的“天赋本质”“辩证法思维”及其“适合学问研究的最佳禀赋”等,评价功能即在不断区分“金银铜铁材质”的人以定位理想国中的基本社会分工,最终遴选出“真正的哲学王”<sup>[2]</sup>。

教育学史上首倡“和谐发展”教育思想的亚里士多德则在年龄阶段身心发展认同规律的同时,提出教育要遵循学习者的不同年龄阶段开展与其身心发展特征相适应的差异化教育,音乐教育是其和谐发展教育的核心,关注音乐

教育对学习者的心灵感受和审美体验,强调文雅教育以符合人的身心发展,服务于自由民的自由发展,关注到学习者的和谐发展,但可惜的是当时的学生评价思想以及实践启蒙带有明显的阶级性。

### 3 迂回发展时期

中世纪和文艺复兴时期至19世纪中后期是学生综合素质评价思想以及实践演进的迂回发展时期。当时主要以基督教教育为主,世俗教育为辅。教会作为西欧封建社会的精神支柱,教育的主要形式是举办教会学校,教育目的是宣扬宗教神学,培养虔诚教徒<sup>[1]</sup>。教会极为仇视动摇宗教信仰的科学文化知识,整个中古世纪在宗教蒙蔽主义的黑暗统治下,科学知识匍匐服从于宗教信仰。在教会学校之外,还有负责培养官吏的宫廷学校教育和负责培养忠君敬主维护封建统治的保卫者的骑士教育,评价标准单一,即忠实上帝、教权和服务上流社会。

中世纪大学后期,大学的出现打破了教会教育绝对控制,虽然教育仍笼罩在浓烈的宗教色彩迷雾中,但大学具有的相对独立性和自治还一定程度促进了教育普及。中世纪的大学选拔教育对象主要通过口试,对申请者的综合素质进行评价,以口试结果作为录取依据;学业结束能否如期顺利毕业,并授予相应学位同样采用口试的评价方式。

中世纪学校教育对学生综合素质(如资质、思维发展、学术思想等)以口试为主要评价方式评判学生发展的思想以及实践,显示出此时期的学生综合素质评价的思想和实践探索均处于迂回发展阶段。文艺复兴时期倡导以人性战胜神性,突破宗教神学对人性发展的绝对压抑,所以人文主义时期对学生综合素质评价侧重于学生的积极进取和努力开拓,在批驳经院主义教育的死记硬背、压制学生发展的同时,人文主义时期的学生综合素质评价关注学生学习的主动性,要求民主、大胆地追求知识、占有科学。

### 4 科学化测量大兴与目标达成的价值判断时期

社会工业化发展的到来,对人才培养提出新的要求,对知识型的劳动者需求,无论在人才数量还是人才质量上均明显需求飙升,夸美纽斯的“班级授课制”得到广泛应用,受教育学生的规模大幅度提升,此前应用于学生评价的“口试”,面对学生规模的扩大,在实践中耗时费力,引发各种批判,这时“笔试”即纸笔测验受到关注,采用“笔试”是学生评价历史上的重要变革。伴随纸笔测验——“笔试”应用于大规模的学生评价,学生评价的科学化问题引发了美国评价学研究者古巴和林肯(E.G.Guba & Y.S.Lincoln)的关注,他们不仅对教育评价领域科学化发展时期进行过划分研究,还开展了相应的批判研究,历史重现演绎出三个学生发展评价的分水岭,这一时期学生评价的主要目的和评价功能侧重于两个方面:甄别选拔优秀学生和对学生未来成就预测<sup>[4]</sup>。

#### 4.1 第一个分水岭——测验与测量时期

19世纪末至20世纪30年代是学生综合素质评价科学化发展的第一个分水岭——测验与测量时期,其标志性的教育评价事件是英国教师费希尔(George Fisher)为测评班级教学中学生的实际发展,专门编制《量表集》并在学生评价中加以实践应用。当时自然科学发展得欣欣向荣,尤其是数学统计、测量技术以及行为主义心理学的兴起<sup>[5]</sup>,激发了教育学家和心理学家对学生评价的学术探讨及实践尝试,将先进的测量统计技术以及行为主义心理学的刺激—反应—强化等行为训练策略相结合,应用到学生评价领域。

第一次世界大战后,“科学管理”的浪潮涌入教育领域,科学管理的高产效能思想之火在教育实践中炙热燃烧,导致学校成了教育的工厂,教师沦为教育生产线上的加工者,学生成了所谓的“教育产品”。这一时期的学生综合素质评价主要围绕学生的智力、人格以及能力发展等,采用测验或测试的评价方法进行评估,学生综合素质评价的组织管理权限在政府,具体实施由学校或者第三方教育评估机构负责;地方教育部门依据地方发展特色,组织涉及学生评价的培训指导;学生综合素质评价的结果受地方教育部门的监督和保障。

#### 4.2 第二个分水岭——泰勒的目标达成时期

20世纪30—50年代后期是学生综合素质评价科学化发展的第二个分水岭——泰勒的目标达成时期。这一时期美国受经济大萧条的影响,导致社会就业岗位大幅度锐减,已完成初等教育的年轻人无法顺利就业,回到高中继续接受教育成为当时无法就业的年轻人的不选之选。但当时极为棘手的是,回到高中接受教育的学生对高中学校设置的课程并不感兴趣,加之受到社会经济滞胀的影响,就业前景缥缈,学生质疑在高中学到的课程内容究竟价值意义何在?基于此教育发展背景,教育政策制定者和教育理论研究者遂下决心组建“进步教育协会”,重建美国普通高中课程体系。教育评价学家拉尔夫·泰勒(R.W.Tyler)作为负责人,围绕课程开发组织开展了为期八年的研究及实践探索,提出围绕教育目标实现的课程开发模式,即“泰勒原理(模式)”,不仅在理论层面为课程开发提供了培训指导,而且引发了学生评价的深刻变革。

“泰勒原理(模式)”关注评价与教育目标密切结合,开展学生评价就是为了促进学生实现教育预期发展目标,提出以“评价观”代替“测试观”,构建了学生综合素质评价最初的“客观描述”范式,即在评价学生的过程中,综合运用多种评价方式方法,客观描述学习者在接受教育后言行变化结果与当初明确的教育目标之间的实现程度,学生评价的实质就是通过分解确定清晰的学生行为目标,以利于评价过程中具有较强的操作性<sup>[6]</sup>。

#### 4.3 第三个分水岭——价值判断时期

20世纪50—70年代是学生综合素质评价科学化发展的

第三个分水岭——价值判断时期。美苏争霸进入冷战期间，苏联的人造卫星上天，引发了当时美国政府以及当权者的极大震撼和恐慌，美国开始深刻反思其教育发展，随即开启了轰轰烈烈的教育改革。其中颇具代表性的法案分别是《国防教育法》和《中小学教育法》，上述法案的颁布实施为美国当时的教育改革提供了相应的法律法规依据。在20世纪五六十年代，美国当时的教育改革专家一般由学科教育专家组成，他们主要继承了泰勒原理（模式）以及行为主义心理学在教育教学中对学生评价的“刺激—反应—强化”等行为训练策略，并仍大行其肆。与此同时，尤其是在20世纪70年代，越来越多质疑的声音此起彼伏，其中塔巴犀利批判泰勒原理（模式），指出泰勒原理（模式）在学生评价中实质是典型的结果评价，对于学生在整个教育教学过程的动态化行为变化发展缺乏基本的关注，并坚决认为评价需要做出相应的调整，学生评价不仅需要关注结果与目标的实现程度，同时需要关注学生在教育教学过程中的动态变化发展。艾斯纳、休伯纳等学者批判测量测试时代的学生评价方法，提出测量无法代表所有的评价理念、评价方法以及评价程序<sup>[7]</sup>。

科学化测量大兴与目标达成的价值判断时期分为三个分水岭，分别是19世纪末至20世纪30年代的学生综合素质评价科学化发展的第一个分水岭——测验与测量时期；20世纪30—50年代后期的学生综合素质评价科学化发展的第二个分水岭——泰勒的目标达成时期；20世纪50—70年代学生综合素质评价科学化发展的第三个分水岭——价值判断时期。随着社会历史背景的更迭，学生评价也在不断发生变化，但这一历史时期学生综合素质评价相关的学生评价的共同特点突出表现为前两个分水岭时期的学生评价均忽略了一些无法通过测量、测试等量化评价方法或手段能涉及的学生道德、情感、个性以及价值观等的评价<sup>[8]</sup>。

## 5 评价主体本位的相互协商和过程建构时期

从20世纪70年代末截至目前是评价主体本位的相互协商和过程建构时期，属于“第四代评价”。美国评价学研究学者古巴和林肯在总结分析麦克唐纳、艾斯纳、格林等学者对“泰勒原理（模式）”的学理批判基础上，对教育评价的发展历程以及前三代教育评价进行相对全面的反思后，提出以往的组织管理主义、科学主义等分类、测试、等级划分的系列评价思想和实践在很大程度上存在明显的弊端，诸如过分依赖科学范式的量化数据说话，为了实现所谓的科学性，忽视学生作为人的道德、情感、价值观等评价机理探究

和实践尝试以及价值主体的多元性，这与学习者实际发展需要其实是无法匹配的。

为了消弭前三代教育评价存在的弊端，推出了“第四代评价”，其基本理念强调学生综合素质评价本质是评价实践主体（评价者与评价对象以及利益相关人）之间进行相互协商和过程建构的过程，其评价实践主张主要选用以质性评价方法，辅之使用测量测试等量化评价方法，对于质性评价方法提倡使用沿用经典的“苏格拉底的自然观察+交互访谈”以及现代管理理论作为基础的“档案袋评价”，基于对学生综合素质评价的评价对象个体的个性化和多元发展特质建议采用“教育鉴赏与教育批评”等。

从20世纪70年代末截至目前的评价主体相互协商和过程建构时期，变革了过去的“自上而下”科层式的组织管理主义和科学主义的学生评价。20世纪90年代后，伴随后现代教育思想的指引下，开始采用“替代性评价”，涉及“直接性评价”“表现性评价”“真实性评价”“档案袋评价”“动态性评价”<sup>[9]</sup>。上述评价采用的评价方式方法极为丰富，诸如自然观察、实践操作、社会调研、交流合作、行动记录、实验操作、剧本表演、演讲竞赛等多种评价方法，主要采用质性评价方法，从学校到家庭到社会生活的各个方面，获取收集整理存档信息，依据多元智能理论的指导，突出评价学生综合素质的发展关键信息。

## 参考文献

- [1] 王小明.普通高中生综合素质评价机制研究[D].上海:上海师范大学,2016.
- [2] 孙培青,王承绪.中外教育比较史纲(古代卷)[M].济南:山东教育出版社,1997.
- [3] 郭守田.世界通史资料选辑(中古部分)[M].北京:商务印书馆,1981.
- [4] Guba, E G & Lincoln, Y S. Fourth Generation Evaluation[M]. Newburg Park,CA:Sage,1989.
- [5] 李树培.描述性学生评价论[M].济南:山东教育出版社,2012.
- [6] 张华.课程与教学论[M].上海:上海教育出版社,2000.
- [7] [美]威廉·派纳.理解课程[M].张华等,译.北京:教育科学社,2003.
- [8] 李树培.珍视不可测量之物——描述性学生评价研究[D].上海:华东师范大学,2008.
- [9] 田芳.体育师范生评价能力现状与发展策略研究[D].上海:华东师范大学,2017.