

基于结构方程模型的高中生综合素质评价实证分析

赵胜利¹, 张力芝¹, 喻小波², 钟妤玥³, 苏理云¹

¹重庆理工大学理学院, 重庆

²重庆长寿中学, 重庆

³重庆渝高中学, 重庆

Email: zhaoshengli@cqut.edu.cn

收稿日期: 2020年12月13日; 录用日期: 2021年1月7日; 发布日期: 2021年1月14日

摘要

构建了高中生综合素质评价指标体系, 建立了综合评价结构方程模型, 通过问卷调查的方式从重庆市两所中学获取了相关数据, 并依据调查数据对两所学校学生的综合素质进行了实证研究。结果表明: 分别从身心健康, 个人修养, 社会实践以及学习状态四个方面来对高中生综合素质进行量化分析是有效的, 四个方面中权重值最大的是学习状态, 权重值为0.31。考试成绩只是综合评价的一个方面, 考试总分高的综合素质得分不一定高, 且综合素质得分高的也不代表四个方面的得分都高。综合各个方面因素对高中生综合素质进行客观的评价, 这正是综合评价的魅力所在。

关键词

高中生, 综合素质, 评价指标体系, 结构方程模型

Empirical Analysis on Comprehensive Quality Evaluation of High School Students Based on Structural Equation Model

Shengli Zhao¹, Lizhi Zhang¹, Xiaobo Yu², Yuyue Zhong³, Liyun Su¹

¹School of Science, Chongqing University of Technology, Chongqing

²Chongqing Changshou Middle School, Chongqing

³Chongqing Yugao Middle School, Chongqing

Email: zhaoshengli@cqut.edu.cn

Received: Dec. 13th, 2020; accepted: Jan. 7th, 2021; published: Jan. 14th, 2021

Abstract

The evaluation index system of comprehensive quality for high school students is established, and the structural equation model of comprehensive evaluation is established through the index system. In order to study this method, the relevant data are obtained from two middle schools in Chongqing by means of questionnaire survey. Based on the survey data and model, this paper makes an empirical study on the comprehensive quality of the students in the two middle schools. The results show that it is effective to quantify the comprehensive quality of senior high school students from four aspects, namely physical and mental health, personal cultivation, social practice and learning state. Among the four aspects, learning state has the largest weight value, and the weight value is 0.31. The examination result is only one aspect of the comprehensive evaluation. High test score does not mean that the comprehensive quality score is also high. Similarly, the high score of comprehensive quality does not mean that the scores of the physical and mental health, personal cultivation, social practice and learning state are high. The charm of comprehensive evaluation lies in the objective evaluation of comprehensive quality based on various factors of high school students.

Keywords

High School Students, Comprehensive Quality, Evaluation Index System, Structural Equation Model

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,新的高考改革已经在全国范围内如火如荼地展开,与新高考制度紧密结合的高中生综合素质评价也成为了教育界的热门话题。建立系统的高中生综合素质评价指标体系,对高中生进行客观的综合评价具有重要的现实意义。

国外的中学生综合素质评价已经趋于完善,其先进教育经验也已经成为我国高考制度改革、实施高中生综合素质评价的重要参考依据。马世晔[1]介绍了国际上主要的几种大规模教育评价体系和该教育评价的系统化、规范化和多元化等主要特点,并总结出国际教育评价对我国教育的启示。孟文娉[2]分别阐述了英国、美国、新加坡以及韩国的高校招考制度中的综合素质评价,总结了他们多元化、多样化的录取标准以及对我国的启示。

国内对于高中生综合素质评价的研究大多偏向理论分析,对于实践性的应用研究很少。周先进[3]就针对高考改革问题,指出了将高中生综合素质评价纳入高考体系的可行性和不易操作性;崔允灏[4]也对我国各地的综合素质评价政策实践的现状与问题做出了考察分析,并结合与其它发达国家教育政策的对比提出自己的建议。高中生综合素质评价的关键就是评价指标的选取,有许多学者对其进行了研究。表1列举了一些研究中所建立的综合评价指标体系的部分指标。

综合上述研究,存在如下的问题:一是指标体系并没有进行科学的验证,使得各个指标间的区分度与归属感都较差;二评价的指标不统一,不同地区的教育特点不同,导致所采用的评价指标体系也会产生较大的差异,难以统一标准;三是指标权重设置方法主观性较强。因此本文将在上述研究的基础上,

构建新的综合评价指标体系，并提出一种修正综合评价指标体系的方法。

Table 1. A list of comprehensive quality assessment indicators for high school students in some studies

表 1. 一些研究中的高中生综合素质评价指标列表

作者	高中生综合素质评价指标
刘坚(2005) [5]	思想道德素质、专业素质、文化素质、身心素质、创新实践能力素质
邓志勇(2010) [6]	道德品质、科学素养、艺术素养、运动健康
张东(2014) [7]	创新能力、人文修养、学习能力、政治素质、身心健康
郑晓丹(2017) [8]	京：思想道德、学业成就、合作与交流、运动与健康、审美与表现、个性发展 沪：品德发展与公民素养、修习课程与学业成绩、身心健康与艺术素养、创新精神与实践能 浙：道德表现、学业水平、运动健康、艺术素养、创新实践
陆炳荣(2018) [9]	品德表现、学业水平、运动健康、艺术素养、创新实践
王咸伟(2019) [10]	合作交流、思想政治、学业评价、实践创新能力、审美表现、身心健康

2. 高中生综合素质评价指标体系的构建与修正

2.1. 综合评价指标体系的建立依据

教育部出台的《关于加强和改进普通高中学生综合素质评价的意见》(以下简称《意见》)中明确要求要从思想品德、学业水平、身心健康、艺术素养以及社会实践这五个方面来反映学生的全面发展情况和他们自身拥有的个性特长，考察学生的社会责任感、创新精神和实践能力。美国著名的心理学家麦克利兰曾根据人的素质表现提出“冰山模型”[11]，认为一个人的个体素质可划分为外在表现与内在表现，外在表现是能够被了解与测量的部分，容易改变，如基础知识、基本技能等，而内在表现是人体内在的、难以测量的部分，不易改变，如自我形象、社会角色等。本文以《意见》的要求以及“冰山模型”为依据，在表 1 的基础上构建了新的高中生综合素质评价指标体系。

2.2. 评价体系的初步建立

根据前文所述综合评价指标体系的建立依据，初步确定了由身心健康、个人修养、学习状态、兴趣爱好以及社会实践 5 个二级指标，一共包含 18 个三级指标的综合评价指标体系，如表 2 所示。

Table 2. Index of evaluation of comprehensive quality of high school students

表 2. 高中生综合素质评价初始体系指标

二级指标	三级指标	对应问卷问题
身心健康	睡眠状态	是否经常失眠
	身体素质	每年生病累计次数
	心理素质	产生烦躁情绪的频率
	自我调节	状态不佳时的调节能力
个人修养	思想道德	个人品德评分
	集体意识	班级凝聚力
	文明礼貌	遇见老师时的表现
	责任义务	班级活动参加次数
	人际关系	对朋友的品德评分

Continued

学习状态	学科竞赛	参加竞赛次数
	学习成绩	语数外总分等级
	校外辅导	校外辅导科目门数
	学习态度	每周请教老师次数
兴趣爱好	艺术活动	参加艺术活动次数
	文学鉴赏	欣赏文学著作时间
社会实践	志愿服务	参加志愿服务次数
	实践活动	参与实践活动积极性
	实验表现	实验课上的表现成绩

2.3. 高中生综合素质评价调查问卷

表 2 给出了问卷和指标体系中指标的对应关系。调查问卷分别在重庆长寿中学和重庆渝高中学进行发放, 这两所学校均属于重庆市普通中学, 学生层次相近, 在一定程度上能代表重庆市高中生的一般水平。总计发放了 736 份, 收回了 649 份, 剔除无效问卷 64 份, 得到有效问卷 585 份, 有效率 90.1%。作答学生分布在高二的不同班级, 其中文科班参与问卷调查的有 318 人, 占作答总人数的 54.4%; 理科班参与问卷调查的有 267 人, 占作答总人数的 45.6%。

2.4. 综合评价体系的检验与修正

首先对调查问卷进行信度分析, 运用 SPSS 软件对问卷数据进行整体的克隆巴赫系数计算, 得到的数值为 0.884, 接近 0.9, 则说明本次收集的问卷数据总体是很可信的, 然后又对每个二级指标进行克隆巴赫系数计算, 得到的具体数值如表 3 所示。

Table 3. Cronbach coefficient of initial comprehensive evaluation index system

表 3. 初始综合评价指标体系的克隆巴赫系数

二级指标	身心健康	个人修养	学习状态	兴趣爱好	社会实践
Cronbach's α 系数	0.814	0.800	0.784	0.705	0.652

其中身心健康和个人修养这两个指标的克隆巴赫系数均在 0.8 及以上, 则可以表明这两个二级指标的内在一致性良好, 但学习状态、兴趣爱好和社会实践这三个二级指标的克隆巴赫系数都小于了 0.8, 则说明这三个二级指标内的三级指标需要修正。计算该指标体系的 KMO 值为 0.868, 说明可利用主成分分析法来对初始综合评价指标体系进行探索性分析, 提取 4 个主成分, 构建新的综合评价指标体系。新的综合评价指标体系与初始指标体系基本一致, 仅有的两处不同就在于将“实验表现”归属于“学习状态”以及将“兴趣爱好”和“社会实践”合并为一个二级指标“社会实践”。此时, 4 个二级指标的克隆巴赫系数都已经大于 0.8, 如表 4 所示。这说明修正以后的综合评价指标体系的一致性良好。

Table 4. Revised Cronbach coefficient

表 4. 修正后的克隆巴赫系数

二级指标	身心健康	个人修养	学习状态	社会实践
Cronbach's α 系数	0.814	0.800	0.802	0.805

3. 高中生综合素质评价结构方程模型

结构方程模型在社会科学研究中是一种常用的融合了因素分析和路径分析的多元统计技术，可以用于处理一些不可观测的变量以及多变量、多结果的关系。

3.1. 结构方程模型

结构方程模型一般包括两个部分：测量模型和结构模型，其矩阵方程组如下所示：

$$\begin{cases} \eta = B\eta + \Gamma\xi + \zeta \\ x = \Lambda_x\xi + \delta \\ y = \Lambda_y\eta + \varepsilon \end{cases} \quad (1)$$

其中，方程组中的第一个方程为结构模型， η 、 ξ 为内生与外生潜变量向量， B 、 Γ 为 η 、 ξ 的结果系数矩阵， ζ 为结构模型的误差向量。后面两个方程为测量模型， x 、 y 为内生与外生观测变量向量， Λ_x 、 Λ_y 为回归系数矩阵， ε 、 δ 为观测模型的误差向量。

3.2. 模型识别与估计

本文结合最终构建的综合评价指标体系来建立结构方程模型。首先，根据上文所得的评价指标体系画出高中生综合素质评价模型对应的路径，如图1所示。图中均由5个潜变量(椭圆形：综合素质、身心健康、个人修养、学习状态、社会实践)、18个观测变量(长方形：身体素质、集体意识、学习态度和艺术活动等)以及22个残差变量(圆形：e1~e18, t1~t4)构成。根据路径图可知，模型中待估参数有40个，数据资料点数为171，可以进行参数估计，且自由度为131。图1中单向箭头上的数字表示为对应的回归系数。

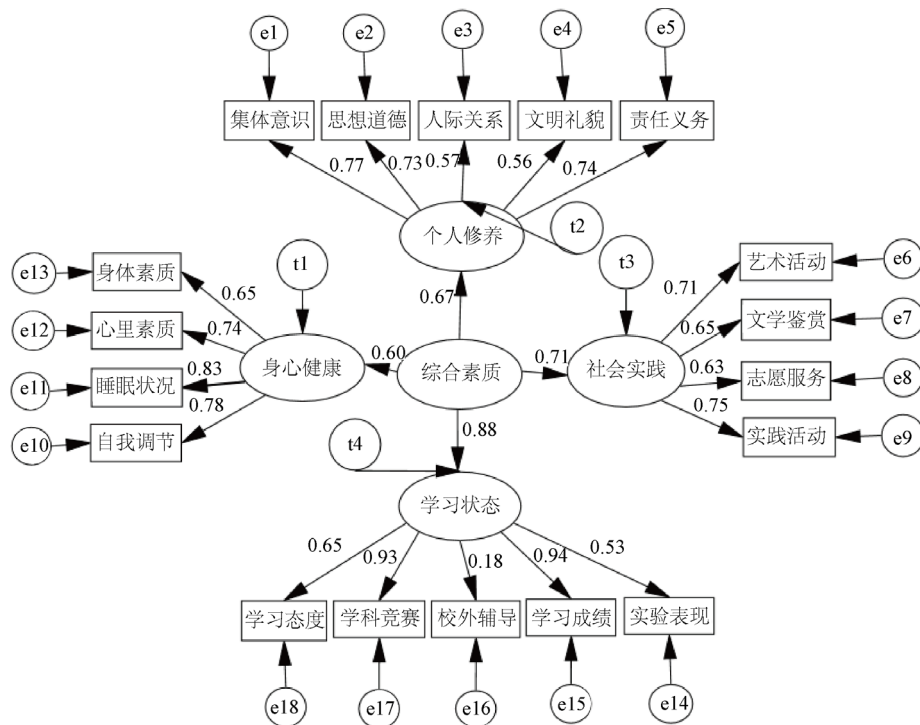


Figure 1. Initial path map of comprehensive quality evaluation for high school students
图1. 高中生综合素质评价初始路径图

3.3. 模型拟合与修正

为防止模型估计的结论产生误导性，需要对估计结果进行检验。表 5 则展示了该模型拟合指标以及对应的检验值。

Table 5. Model fitting index and test value
表 5. 模型拟合指标及检验值

拟合指标	CMIN/DF	RMSEA	IFI	TLI	CFI
理想标准值	<5	<0.08	>0.9	>0.9	>0.9
本研究中的值	2.301	0.072	0.912	0.899	0.912
检验结论	理想	理想	理想	不理想	理想
修正后	2.124	0.067	0.925	0.913	0.924

从表 5 中的模型拟合指标和对应的检验值可看出，卡方值与自由度的比值(CMIN/DF)小于了 5，属于理想结果，RMSEA (近似误差均方根)的值小于 0.08，也属于理想结果，IFI (增值拟合指数)和 CFI (比较拟合指数)均大于了 0.9，也符合理想标准范围，但 TLI (Tucker-Lewis 指数)小于 0.9，检验结果不理想，需要利用 MI 修正指数来对该模型进行修正。

通过修正得到了最终的结构方程模型，其各个变量之间的路径系数如图 2 所示，从二阶因子的分析结果来看，除了校外辅导，其它的二阶因子的载荷均在 0.5 以上，则可以表明二阶因子与一阶因子之间的关系较强，且二阶模型能够充分地表达一阶因子间的关系，从而使得这四个一阶因子可聚合为“综合素质”。最终得到了高中生综合素质评价结构方程模型，模型的具体表达式见附录。

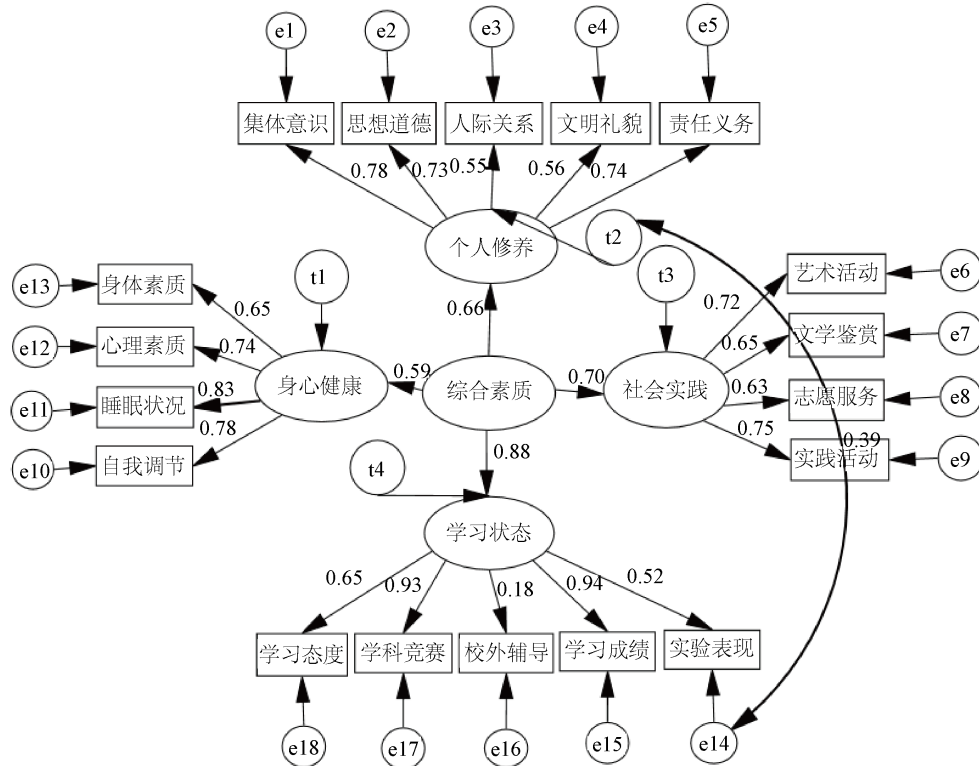


Figure 2. Final path map of comprehensive quality evaluation for high school students
图 2. 高中生综合素质评价最终路径图

3.4. 结果分析

根据图 2 的路径系数, 确定高中生综合素质评价的二级指标及所属的三级指标的路径系数和归一化处理后的权重, 结果如表 6 所示。

Table 6. Evaluation index system standardization path coefficient and weight

表 6. 评价指标体系标准化路径系数及权重

一级指标	二级指标		三级指标	
	指标及系数	权重	指标及系数	权重
综合素质(η)	身心健康 0.59 (η_1)	0.21	身体素质 0.65	0.22
			心理素质 0.74	0.25
			睡眠状态 0.83	0.27
			自我调节 0.78	0.26
	个人修养 0.66 (η_2)	0.23	集体意识 0.78	0.23
			思想道德 0.73	0.22
			人际关系 0.55	0.16
			文明礼貌 0.56	0.17
	社会实践 0.70 (η_3)	0.25	责任义务 0.74	0.22
			艺术活动 0.72	0.26
			文学鉴赏 0.65	0.24
			志愿服务 0.63	0.23
	学习状态 0.88 (η_4)	0.23	实践活动 0.75	0.27
			学习态度 0.65	0.20
			学科竞赛 0.93	0.29
			校外辅导 0.18	0.06
			学业成绩 0.94	0.29
			实验表现 0.52	0.16

表 6 给出了综合评价指标体系中各级指标之前的权重关系, 根据二级指标的权重, 可得到高中生综合素质评价的综合评价值为:

$$\eta = 0.21 \times \eta_1 + 0.23 \times \eta_2 + 0.25 \times \eta_3 + 0.31 \times \eta_4 \quad (2)$$

指标的权重反映了该指标在进行高中生综合素质评价时应占有的比重, 单从表 6 中给出的各指标所占权重结果来看, 与国家对高中生的培养方案相一致, 更注重学生学习能力与实践能力的培养。

4. 对重庆两中学的学生学习情况的评价分析

基于重庆市两中学的调查问卷数据进行实证分析。根据上文利用结构方程模型求得的各项评价指标的权重, 可结合实际调查数据计算出重庆市两中学高中生的综合素质得分。首先利用学生对问卷的回答和真实成绩, 通过加权求和的方法计算出各项二级指标的评价得分, 从而根据式(2)计算得到高中生综合素质评价的综合评价。

结合每个学生的高二总成绩以及从高一到高二的文理科年级排名变化情况(正为上升,负为下降),由于参与调研的学生较多,这里就不一一展示他们的综合得分情况,只列举出了前二十名的学生学习情况,即选取高二总分排名前二十的学生,并将其对应的综合评分进行从大到小的排序,得到表7的数据。

Table 7. Scores of the top 20 students

表 7. 学生前二十名的得分情况

学生编号	文理科	身心健康	个人修养	学习状态	社会实践	总分	排名变化	综合评分
2082	理科	4.06	4.62	3.6	3.48	591	44	3.90
2123	理科	2.48	3.4	4.36	4.5	668	0	3.78
2117	理科	3.52	4.78	3.59	2.72	590	26	3.64
2033	理科	4.32	4.18	3.34	2.45	612	1	3.52
2016	理科	4	4.39	2.98	2.69	615	21	3.45
2023	理科	3.54	4.55	3.53	2.19	636	52	3.44
2130	理科	4.51	4.78	2.2	2.72	586	58	3.41
2001	理科	4.54	4.4	3.04	1.95	615	-4	3.4
2010	理科	3.3	4.1	3.49	2.49	605	5	3.35
2067	理科	4.06	3.66	3.33	2.41	625	-1	3.33
2090	理科	3.54	3.51	3.33	2.95	596	88	3.32
1065	文科	2.8	4.11	3.94	2.19	607	0	3.31
2111	理科	2.96	4.18	3.03	2.75	588	-38	3.21
1101	文科	3.54	4.11	2.47	2.89	587.5	3	3.18
1071	文科	3.58	3.88	3.8	1.24	600.5	7	3.14
2052	理科	3.98	3.73	2.94	1.96	587.5	45	3.1
2058	理科	3.49	3.71	2.69	2.22	592	141	2.98
2054	理科	2.77	3.22	3.27	2.45	601	9	2.95
2131	理科	3	3.43	3.12	1.49	594	16	2.76
2012	理科	1.94	3.01	3.28	2.45	617	92	2.73

根据表7可看出,考试总分高的综合素质得分不一定高,且综合素质得分高的也不代表各项二级指标都高,这说明单单从考试成绩来判断一个学生的学业好坏并不科学,需要从学生的各方面来综合评价,毕竟每个人的个性与特性都不一样,存在个体差异性。

从上述表格的得分情况来看,参与调查的学生从高一升到高二的成绩排名都有很大的变化,大多数的排名均上升,也有小部分排名下降,因此,在科学评价学生的成绩时,也需要过程性评价,需要全面了解学生考试成绩的变化情况,了解学生的可塑性高低。由于综合素质评价和高考“硬挂钩”,则各个高校在录取学生的过程中,可以以自己学校所看重的点来利用加权平均的方式将综合素质与高考得分进行结合,总的分析学生的优异程度,挑选出自己所需要的人才。

5. 结论

根据国家政策的要求,在前人研究的基础上构建了高中生综合素质评价指标体系,建立了综合评价

结构方程模型。通过问卷调查的方式从重庆市两所中学获得了相关数据,并对两所学校学生的综合素质进行了实证研究。结果表明,高中生综合素质评价应该从四个方面来进行,分别是身心健康,个人修养,社会实践以及学习状态。四个方面的权重值相差不大,其中,权重值最大的是学习状态,权重值为 0.31,权重值最小的是身心健康,权重值为 0.21。综合评价结果和考试成绩之间没有必然的联系,考试总分高的综合素质得分不一定高,且综合素质得分高的不代表各项二级指标得分都高,这恰好是综合评价的目的所在,能综合各个方面去对高中生的综合素质进行综合的评价。此外,评价的结果显示,理科生的综合评分明显比文科生高,由此可看出文理分科不利于学生综合素质的培养,进而说明了新高考改革中取消文理分科的正确性。

基金项目

重庆市社会科学规划项目(2018BS56)。

参考文献

- [1] 马世晔. 从国外教育评价制度看我国基础教育评价体系的建立[J]. 中国考试, 2008(5): 13-18.
- [2] 孟文娉. 国外高考招生过程中综合素质评价的经验及启示[J]. 教育实践与研究, 2013(5): 9-10.
- [3] 周先进, 张睦楚. 高考改革: 高中生综合素质评价的“可为”与“难为”[J]. 全球教育展望, 2014, 43(7): 101-111.
- [4] 崔允灏, 柯政. 关于普通高中学生综合素质评价研究[J]. 全球教育展望, 2010, 39(9): 3-8.
- [5] 刘坚. 基于多元统计法的综合素质评价模型研究[J]. 曲阜师范大学学报, 2005, 37(3): 19-22.
- [6] 邓志勇. 构建学生综合素质评价体系的策略研究[J]. 学校管理与发展, 2010(6): 16-19.
- [7] 张东, 张恺, 等. 青少年综合素质的多层次评价[J]. 学术论坛, 2014, 8(8): 182-182.
- [8] 郑晓丹, 崔玉平. 京沪浙普通高中生综合素质评价方案的对比分析[J]. 滁州学院学报, 2017, 5(19): 103-107.
- [9] 陆炳荣, 杨大忠. 高考综合改革形式下高中生综合素质评价实践探索——以浙江省桐乡市高级中学为例[J]. 中国考试, 2018(2): 30-36.
- [10] 王咸伟, 徐晓东, 等. 信息化环境下中小学生综合素质评价指标体系构建[J]. 课程与教学, 2019(1): 67-75.
- [11] 李莉. 基于冰山模型探索高校综合评价招生模式[J]. 教育理论研究, 2016(21): 117-118.