

初中学生综合素质智能化评价体系研究

程蓓

中国教育科学研究院教育评价与督导研究所

DOI:10.12238/jief.v7i2.12831

[摘要] 在教育数字化转型浪潮下,课堂教学、校园管理与学生发展的数据采集已从单一文本向多模态延伸,但现有模型对视频、文本、生理信号的融合准确率普遍不高。因此为了提升初中学生综合素质评价水平,本文以《义务教育质量评价指南》为核心依据,遵循全面性、可操作性和动态权重三大原则,尝试构建初中学生综合素质智能化评价指标体系。

[关键词] 多模态数据; 初中学生综合素质; 评价指标

中图分类号: G625.5 **文献标识码:** A

Research on the Intelligent Evaluation System for Comprehensive Quality of Junior High School Students

Bei Cheng

Institute of Educational Evaluation and Supervision, China National Academy of Educational Sciences

[Abstract] Under the wave of digital transformation in education, data collection for classroom teaching, campus management, and student development has extended from a single text to multiple modalities. However, existing models generally have low accuracy in integrating video, text, and physiological signals. Therefore, in order to improve the comprehensive quality evaluation level of junior high school students, this article takes the "Guidelines for Quality Evaluation of Compulsory Education" as the core basis, follows the three principles of comprehensiveness, operability, and dynamic weighting, and attempts to construct an intelligent evaluation index system for the comprehensive quality of junior high school students.

[Key words] multimodal data; Comprehensive quality of junior high school students; Evaluation

引言

现阶段,我国教育评价改革已然步入深水区,传统以分数为主导的评价模式已无法满足政策要求,亟需通过智能化手段破解重结果轻过程、重学业轻素养的困境。基于此,本文从构建初中学生综合素质智能化评价体系的必要性出发,对初中学生综合素质智能化评价指标体系的构建进行了论述说明。

1 构建初中学生综合素质智能化评价体系的必要性

1.1 政策导向: 教育评价改革的顶层设计与时代使命。近年来,我国教育评价改革步入深水区,政策层面持续释放破除“五唯”痼疾的强烈信号。2020年《深化新时代教育评价改革总体方案》明确要求“完善立德树人体制机制,扭转不科学的教育评价导向”。2021年《义务教育质量评价指南》进一步细化评价维度,提出“加快建立以发展素质教育为导向的义务教育质量评价体系”,要求建立覆盖德智体美劳的全要素评价框架。与此同时,中考改革的深入推进,如北京、上海等地将综合素质评价结果按比例纳入升学总分,使得初中阶段的评价体系成为教育公平与质量的关键抓手。在此背景下,传统以分数为主导的评价模式已

无法满足政策要求,亟需通过智能化手段破解重结果轻过程、重学业轻素养的困境。

1.2 实践需求: 教育数字化转型与评价能力提升的双重驱动。教育数字化转型浪潮下,课堂教学、校园管理与学生发展数据的采集已从单一文本向多模态延伸^[1]。调查显示,当前全国中小学智慧课堂覆盖率及课堂录播系统等设备配备率均较高,日均产生大量视频数据及生理信号数据。然而,当前教育实践面临两大矛盾: 其一,海量数据与低效利用的矛盾,教师平均每周耗费大量时间填写评价档案; 其二,评价理念与技术工具的脱节,传统评价方法难以捕捉学生的隐性素养,如合作能力、情绪管理等,导致综合素质评价流于形式化。智能化评价体系的构建,既是教育数字化转型的必然要求,也是提升学校评价能力、减轻师生负担的现实需求。

1.3 问题导向: 综合素质评价的推进困境与技术瓶颈。当前初中学生综合素质评价存在三重结构性矛盾: 其一,指标体系模糊化,各地评价标准差异显著,部分指标缺乏可量化依据,导致评价结果可比性差^[2]; 其二,数据应用表层化,多模态数据多停

留于存储阶段,跨模态关联分析不足,未能有效形成“采集-分析-反馈”闭环;其三,技术伦理风险显性化,数据隐私保护机制不完善,相关利益群体担忧生物识别数据泄露,算法偏见可能加剧教育不公平^[3]。与此同时,技术层面面临多模态数据异构性高、实时分析算力不足等挑战,现有模型对视频、文本、生理信号的融合准确率普遍不高。这些问题共同指向一个核心命题,即如何通过技术创新构建科学、公平、高效的智能化评价体系,真正实现以评促育。

2 国内外关于初中学生综合素质智能化评价的研究动态

2.1 学生综合素质评价国内外研究动态。国内学生综合素质评价研究以政策驱动为核心,强调从单一学科成绩向多维度评价转变。例如,福州等地的评价体系涵盖道德品质、学习能力、社会能力等维度,注重全员、全方位、全程评价,以促进学生全面发展。近年来,国内对学生综合素质的评价呈现出注重评价方法与工具创新。研究提出多元化评价方法,包括观察法、问卷调查法、作品分析法等,强调过程性数据与结果性数据的结合。例如,通过学生成长档案袋记录学习轨迹,结合家长、教师、学生三方互评,增强评价的全面性。部分研究探索大数据与AI技术在教育评价中的应用,如多模态数据分析(行为日志、课堂互动等),但仍处于初步阶段。此外,部分研究关注实践问题与挑战,如标准与数据问题,评价指标操作性不足,部分学校仍依赖主观文字描述,缺乏统一量化标准;教师能力与参与度问题,教师培训不足,导致评价过程流于形式,部分教师仍以成绩为主要导向;结果应用局限问题,评价结果多用于升学参考,未能有效反馈至教学改进,存在为评价而评价的现象。部分地区开展了实践探索,如上海通过名额分配综合评价录取机制,将综合素质评价与招生挂钩,强调过程性参与而非竞争性差异;四川德阳引入“校园欺凌”一票否决制,推动评价与德育深度融合。

国外对学生综合评价的有关研究主要聚焦评价理念与模式、技术赋能与工具开发、伦理与公平性等方面。如美国注重个性化与多元评价,如学生自我评价问卷和关键能力框架,强调批判性思维、沟通能力等非认知素养评估。英国采用形成性评价与总结性评价相结合的方式,例如通过项目式学习(PBL)评估学生实践能力。芬兰以无标准化考试著称,通过教师观察、学生作品展示等动态评价,促进学生终身学习能力发展。国外研究较早探索技术驱动的评价工具,如澳大利亚的学生中心评价理论支持下的成长档案系统,利用数字化平台记录学生多维度表现。此外,国外学者关注评价中的算法偏见与数据隐私问题,如欧洲多国制定教育AI伦理指南,明确数据使用边界,确保评价过程的透明性。

2.2 智能化学生评价的国内外研究动态。国内对智能化学生评价方面的研究主要聚焦智能化评价体系的重构与改革、技术赋能评价方法与工具创新,以及评价场景与应用领域的扩展。国内研究聚焦于智能化教学环境下学生评价体系的重构,强调通过大数据、人工智能等技术实现从单一结果性评价向多维度过程性评价的转变。如基于智能化学习环境的学生评价体系需整

合学习行为数据、课堂互动数据等多源信息,构建动态化、个性化的评价模型。研究指出,传统评价体系在智能化环境中存在适应性不足的问题,亟需改革以支持个性化学习和精准教学。此外,国内学者还积极探索人工智能技术在教育评价中的应用,如利用自然语言处理自动批改作文,通过机器学习分析学生知识掌握程度,并生成个性化学习建议。同时,智能化教学评价系统的设计成为热点,例如基于深度学习的自适应测评系统,能够实时调整评价指标以适应不同学生的学习进度。评价场景方面,研究覆盖了从基础教育到高等教育的多阶段场景。如在基础教育中,智能化评价系统被用于课堂即时反馈;在高等教育中,则通过大数据分析学生科研能力与综合素质。部分研究还探索了智能教育平台中的学习者画像构建与动态评价机制。在此基础上,国内研究普遍关注智能化评价的潜在风险,包括数据隐私泄露、算法偏见、过度依赖技术导致教师角色弱化等问题。学者建议加强伦理规范与技术治理,建立透明、可解释的评价模型。

国外对智能化学生评价的研究主要关注了生成式人工智能的实际应用探索、伦理风险与学术诚信研究,以及技术与教育模式的深度融合。国外研究更注重生成式人工智能在教育评价中的落地应用。例如,美国在医学教育中利用ChatGPT模拟患者对话,提供个性化反馈以提升临床技能训练效果。此外,国外研究强调技术在跨学科评价中的整合,如STEM教育中的项目式学习智能评估。国外学者较早关注智能化评价的伦理问题,如学术诚信风险、算法公平性等。研究提出需建立学术道德教育框架,培养学生批判性思维以合理使用AI工具;欧洲多国推动制定AI教育伦理指南,明确数据使用边界与责任归属。同时,国外研究强调智能化评价与教学模式的协同创新。如基于生成对抗网络的虚拟实验环境,能够实时评估学生的操作能力;通过变分自编码器分析学习路径,优化课程设计。

综合分析国内外有关研究,可知国内智能化学生综合素质评价以政策驱动为主,注重评价与升学体系的衔接,但技术应用尚待深化,国外则强调评价工具的科学性与学生主体性,技术整合更成熟,但面临文化适配性挑战。二者同时存在诸如评价标准的科学性与操作性不足、定量与定性评价的平衡问题,以及教师评价能力与培训机制缺失等共性挑战。近年来呈现出AI与大数据技术将推动评价从“经验判断”向“数据驱动”转型的研究趋势,以及通过多模态数据的采集与分析,实现对学生认知、情感、能力的全方位评估。

3 初中学生综合素质智能化评价指标体系的构建

初中学生综合素质评价指标体系的构建以《义务教育质量评价指南》为核心依据,遵循全面性、可操作性和动态权重三大原则。在全面性层面,指标体系覆盖道德品质、学业水平、身心健康、艺术素养、社会实践五大维度,下设15项二级指标,例如“道德品质”维度包含“社会责任感”“诚信意识”“团队协作”等具体指标,确保评价覆盖学生发展的全领域。可操作性要求指标与多模态数据直接关联,如“团队协作”通过视频中的小组讨论频率、文本中的互评关键词,以及实践活动记录共同量化。动

态权重设计则引入熵权法,根据区域差异实时调整指标权重,避免一刀切。如针对农村学校“社会实践”资源不足的问题,系统自动降低该维度权重,同时提升“学业水平”和“身心健康”的占比,使评价更贴合实际。

表 三级评价指标及观测点

一级指标	二级指标	三级指标	观测点
道德品质	社会责任感	志愿服务参与度	志愿服务平台记录
		集体活动贡献	活动分工记录文本分析
		社会问题关注	周记或讨论中提出社会问题解决方案
	诚信意识	考试诚信记录	考场监控视频异常行为识别
		作业原创度	作业文本相似度
		日常言行一致	教师评语文本挖掘
	团队协作	小组贡献度	任务分配记录
		冲突解决能力	小组讨论音频情感分析
		倾听与反馈	语音对话关联分析
学业水平	学科知识掌握	标准化测试成绩	考试系统成绩单
		知识点迁移能力	跨学科应用题正确率分析
		课堂互动质量	课堂问答音频识别与匹配
	学科实践能力	实验操作规范性	实验视频动作分析
		问题解决效率	项目式学习任务完成时间
		跨学科应用	提交跨学科报告文本分析
身心健康	身体健康水平	新颖观点提出	开放性答案文本分析
		创客活动参与	参加校级以上创客活动
		成果转化能力	校级以上展示或竞赛获奖
	心理健康状况	体能测试达标	体育测试系统数据
		日常运动习惯	智能手环日常运动数据
		疾病预防意识	健康知识在线测评系统
艺术素养	艺术知识与鉴赏	情绪稳定性	心理测评焦虑分析
		课堂专注度	课堂专注表情与姿态分析
		社交适应性	社交平台正向互动情感分析
	艺术技能表现	规律作息	智能手环监测日均睡眠时长
		膳食均衡	食堂膳食消费数据
		卫生习惯	图像识别整洁度
	艺术创造力	艺术理论测试	标准化测试
		展览参观记录	每学期参观艺术展行为日志
		美学表达深度	课堂讨论中美学关键词分析

一级指标	二级指标	三级指标	观测点
社会实践	社会参与度	志愿服务时长	志愿服务平台记录
		社会调查深度	校级以上调查报告获奖数据
		职业体验多样性	职业体验活动日志分类统计
	实践能力	任务完成效率	社区项目按时完成率
		协作领导力	担任小组负责人次数
		资源整合能力	活动中调用外部资源
	公民意识	公德行为记录	无公共场合违纪行为
		环保行动参与	垃圾分类正确率
		法治知识掌握	在线法治测试得分

4 结语

本研究立足教育数字化转型背景,以《义务教育质量评价指南》为纲领,综合考量多模态数据融合与动态权重设计,构建了覆盖道德品质、学业水平等五大维度的初中学生综合素质智能化评价指标体系。该体系有助于破解传统评价重结果轻过程、数据利用表层化等难题,为教育评价改革提供可量化、可操作的推进路径。未来需进一步优化多模态算法,完善数据伦理规范,推动智能化评价与教学实践的深度融合,真正实现以评促育、全面发展的教育目标。

[课题信息]

中国教育科学研究院中央级公益性科研院所基本科研业务费专项资助“基于多模态数据的初中学生综合素质智能化评价体系研究(GYC2024003)”。

[参考文献]

[1]陈丽.智能技术支撑学生综合素质评价:改革与创新[J].现代教育技术,2023,33(12):5-13.

[2]王洪席.学生综合素质评价责任主体的认知分歧与路径澄清[J].教学与管理,2024,(07):67-70.

[3]郑永和,王一岩,杨淑豪.人工智能赋能教育评价:价值、挑战与路径[J].开放教育研究,2024,30(04):4-10.

作者简介:

程蓓(1987-),女,汉族,山西运城人,助理研究员,博士,中国教育科学研究院教育评价与督导研究所,研究方向:教育督导评价。