

■ 智能时代教师评价素养的内涵重构、价值阐释、框架构建及应用*

李毅^{1,2}, 郑鹏宇¹

1.西南大学 教育学部, 重庆 400715

2.西南大学基础教育研究中心, 重庆 400715

摘要: 智能时代的到来推动教育评价领域变革, 也对教师评价素养的内涵结构提出了新的要求。该研究首先对智能时代教师评价素养内涵进行重构, 剖析了其价值, 并构建了包含智能时代教师评价知识、智能时代教师评价技能、智能时代教师评价习惯、智能时代教师评价伦理 4 个一级维度和 12 个二级维度的结构框架。据此框架研制了测评量表, 在全国范围内展开测评, 得出研究结果: (1)智能时代教师评价素养总体达到中等水平, 评价习惯和伦理水平相对较高, 评价知识和技能水平偏低; (2)智能时代年龄和教龄更低的教师评价素养水平更高, 尤其体现在评价习惯和伦理维度; (3)智能时代东部地区的教师评价知识和技能水平显著高于中西部地区的教师, 评价习惯和评价伦理维度无显著差异; (4)智能时代城镇地区的教师评价素养总体水平及各维度水平均显著高于乡村地区的教师。基于此, 提出对策建议。

关键词: 智能时代; 评价素养; 教学评价; 智能技术

中图分类号: G434

文献标识码: A

* 本文系重庆市教育科学十四五规划重点课题“区县中小学教师数字素养提升策略与路径研究”(课题批准号: K24YB2020021)、西南大学研究生科研创新一般项目“人工智能时代高校拔尖创新人才培养模式的构建及实践”(项目编号: SWUS24015)研究成果。

2020 年, 中共中央国务院印发《深化新时代教育评价改革总体方案》(后称《方案》), 其中明确提出要“加强教师评价能力建设”。教师是教育评价实施的关键主体, 其评价水准直接关系到教育教学质量以及学生综合素质的提升。美国、英国、新西兰、澳大利亚等国都将教师评价素养相关要求列入教师专业标准中, 视其为教师专业素质的重要组成部分^[1]。随着以 DeepSeek 为代表的人工智能时代到来, 新兴技术正深刻改变着教育评价的内容和方式及其背后承载的价值观念, 为教育评价变革带来了新的机遇与挑战。这无疑也给教师的评价素养带了新的要求。但已有研究对智能时代教师评价素养的相关理论和实践研究还不足。因此, 本研究基于国内外具有代表性教师评价素养理论研究成果, 结合智能时代我国教育评价变革的特征和要求, 对智能时代教师评价素养的内涵、价值、结构及其应用进行探讨。

一、智能时代教师评价素养的内涵重构与价值阐释

人工智能与教育领域的融合重塑了教师评价活动的面貌, 对教师的评价素养提出了新的要求。为此, 迫切需要基于智能时代背景对教师评价素养的进行内涵重构与价值阐释。

(一) 教师评价素养与评价素养的关系

在教育领域中, 除教师外, 评价素养的主体还可指向学生、家长、校长、教育行政管理人员、测评专家等一系列与教育评价相关的人员。各主体承担的评价工作有所不同, 这就导致针对各主体的评价素养要求存在一定程度上的差异。理清教师评价素养与评价素养的关系, 是开展智能时代教师评价素养内涵探讨与价值阐释的前提。

在教育评价体系中, 与教师日常评价活动关联最多的是课堂教学评价和招生考试评价^[2]。相较于前者, 后者又叫外部评价。学科专家、教育测量学专家、教育评价学专家是该类评价的主要工具开发与测评方案拟定人员, 教师在该类评价中往往扮演协助者角色。而课堂教学评价是教师为了解其课堂教学情况, 掌握学生学习状况, 为促进学生有效学习与发展而采取各种方式收集、分析和不断反馈学生学习信息的活动过程^[3]。在课堂教学评价中, 教师全程主导并参与了各项评价活动。因此, 对于一线教师来说, 开展课堂教学评价才是对其评价素养的真正考验。而对于离一线教学实践相对较远的教育专家和教育行政人员来说, 课堂教学评价并非其工作的重点。因此, 教师评价素养主要侧重于实施、开展课堂教学评价所需的知识、能力、态度、理念和伦理品格。

(二) 智能时代教师评价素养的内涵

1991年, Stiggins首次提出“评价素养”这一概念, 将其界定为教师应具备的基础性教育评价知识和将这些知识运用到各种学生成就测验中的技能^[4]。此后相关的研究及政策标准大多沿用了这一理论框架, 例如 Brookhart提出的11项评价知识与技能标准^[5]以及“美国1990年标准”^[6], 只不过在具体的评价知识与技能层面存在差异。同时, 也有学者超越“知识+技能”的传统框架, 对教师评价素养的内涵进行了创新。例如, Willis从社会文化学的角度出发, 认为评价素养并不是一种单一的或固定的能力, 教师需要能够在社会的评估文化和政策背景下开展评价活动^[7]。Looney构建了超越传统意义上的“教师评价素养”的新概念——教师评价身份(Assessment Identity), 强调信念、品格、情感参与在教师评价工作中的重要意义^[8]。郑东辉也聚焦我国新时代教育评价改革的需求, 基于“为了学习的评价”这一新的评价理念, 将教师评价素养概括为教师依据促进学习的评价理念开展学生评价所表现出来的关键能力和伦理品格^[9]。近些年来, 教师评价素养内涵的研究整体呈现出三大趋势: 一是将教师评价素养视为动态化、情境化的概念, 认为其内涵应随教学情境及时代背景的变化而变化。二是力图跳出“知识+技能”的传统理论框架, 对教师在评价工作中展现出的价值理念、情感态度及伦理道德等要素予以更多关注。三是紧密结合评价理论最新进展, 将形成性评价、为了学习的评价、文化理解评价等新兴评价理论纳入其内涵中。

基于上述内容, 智能时代教师评价素养的内涵应呈现出以下新的特征: (1)以智能评价理念为引领, 紧密结合智能评价理论最新进展。随着人工智能时代的到来, 许多学者主张要发挥智能技术优势, 强化评价的过程性与发展性^[10], 促使学生在评价过程中主动反思与改进, 最终实现教学评价的个性化、全面化、精准化。这一理念应体现在智能时代教师评价素养内涵中; (2)超越“知识+技能”的传统框架, 将更多要素纳入其中。除相应的知识与技能外, 教师在智能时代开展教学评价过程中体现出的价值理念、情感态度、行为习惯, 展现出的创新素质以及对伦理道德的坚守, 都是其能否在智能时代有效开展教学评价活动所应考虑的内容; (3)紧密结合智能时代教学评价实践情境及需求。例如, 在知识与技能层面, 要在原有内容基础上融入在智能教学环境中开展评价活动所需的智能技术知识与应用技能。在伦理道德层面, 要结合智能技术带来的伦理隐患, 推动智能时代教师评价伦理标准的更新。基于以上特征, 将智能时代教师评价素养内涵概括为教师在精准化、个性化、全面化等智能评价理念的引领下, 将智能技术合理、有效地整合进教学评价过程中, 为提升教育评价功效、促进学生发展所应具备的知识、技能、情感态度、

行为习惯及伦理品格。

(三) 智能时代教师评价素养的价值

1. 推动教师评价素养标准的更新, 引领智能时代教师评价素养培育

目前世界上已有多个国家建立了专业化的教师评价素养标准, 但纵观这些标准, 鲜有将“技术”这一要素纳入其中的。智能时代的到来要求学界进一步展开智能时代教师评价素养内涵与结构的研究, 从而为更新、完善教师评价素养标准提供参考。

构建智能时代教师评价素养内涵框架, 研制相应标准, 能够为智能时代教师评价素养培育提供指引。首先是推动培育方案更加有效地回应智能时代需求。其次是推动培育课程的完善, 不仅使课程内容紧随智能时代发展潮流, 也使课程的目标、内容、评价体系更加明确, 降低智能时代教师评价素养培育的盲目性与随意性。最后体现在为智能时代教师评价素养培育项目的开展, 以及相应保障体系与认证体系的构建提供指引。

2. 提升智能时代教学评价实施成效, 促进教育教学质量提升和学生发展

提升智能时代教师评价素养, 是智能时代教学评价有效实施的必要保障。智能时代的教学评价有别于传统评价情境, 要求教师能够合理有效地利用智能技术开展一系列教学评价活动, 使智能技术充分发挥优化作用。

智能时代教师评价素养的形成与发展, 更进一步关系到教育教学质量以及学生综合素质的提升^[11]。具备智能时代评价素养的教师, 能够在教学评价过程中充分发挥智能技术的功效, 促进学生测评精准化、高效化、个性化, 为学生发展和教育教学改进提供更加有效的支持, 保障智能技术赋能教学优化及学生的发展。

3. 助推教育评价体系变革, 推动智能时代教育评价范式转型

推动教育评价体系变革, 一方面需要充分发挥教育领域的内生变革动力, 提升教师的评价素养, 若教师无法有效落实新时代教育评价理念, 无法具备评价变革所需要的创造力和创新精神, 那么教育评价系统的变革必然是收效甚微的; 另一方面需要注重智能技术发展为教育评价变革技术困境突破带来的机遇, 但要充分发挥智能技术功效, 还需要教师具备良好的智能时代评价素养^[12]。

与此同时, 智能时代教师评价素养的提升对于教育评价范式的转型也具有重要推动作用。随着智能时代的到来, 教育评价范式呈现出由依赖问卷、观察、标准化测试的数据评价, 向以多元大数据驱动下的科学决策为特征的数智评价转型的趋势^[13], 从而推动教育评价更加科学、客观。而推进教育评价范式转型同样需要教师具

备智能时代评价素养，如此才能够更加快速地适应数智评价范式的目标、理念、方法、手段。

二、智能时代教师评价素养的框架构建

首先基于智能时代教师评价素养内涵，结合国内外教师评价素养的核心要素，初拟智能时代教师评价素养的理论框架；然后采取德尔菲法，对该理论框架进行修订和完善，最终建构了完整的智能时代教师评价素养框架。

(一) 智能时代教师评价素养的框架初构

关于教师评价素养的结构，就已有研究来看，整体呈现出由一维到多维的趋势。早期主要以“知识+技能”的框架呈现。近十余年来，越来越多的研究尝试将理念、态度、伦理、品格等要素纳入其中，体现了教师评价素养向情感、文化维度的拓展（如表 1 所示）。其中，美国密歇根评价委员会在 2013 年 InTASC 教师标准中提出的“评价知识+评价技能+评价品格”的框架最具代表性^[14]。“评价品格”的涵盖范围较广，包括除知识与技能外的信念、价值观、态度、伦理道德等诸多因素。因此，本研究拟遵循此框架。但“评价品格”是复合概念，涵盖要素较多，不易测度，为此本研究参考 InTASC 的内涵，即评价品格为教师专业行动的习惯及道德承诺，将“评价品格”细分为“评价惯习”和“评价伦理”两个维度。其中“评价惯习”指向教师开展评价活动的行为倾向，它是建立在特定思维方式基础之上的。“评价伦理”即教师在评价过程中应遵循的伦理规范。

表1 国内外教师评价素养核心要素提取与分布

年份	文献来源	评价知识	评价技能	评价理念	评价态度	评价伦理	评价品格
1990	美国 1990 年标准		√				
1991	Stiggins	√	√				
2011	Brookhart	√	√				
2013	美国 InTASC	√	√				√
2016	美国密歇根评价委员会	√	√				√
2017	Looney	√	√	√	√		√
2017	耿娟娟	√	√		√		
2019	Pastore	√	√	√		√	
2019	高灿灿	√	√		√		
2021	Cecilia	√	√	√	√	√	
2022	郑东辉		√	√		√	
2024	杨卓	√	√	√			

因此，这就意味着教师不仅要具备在智能时代开展教学评价活动所需的知识与技能，还需建立起在智能时代开展教学评价应具备的行为惯习和伦理规范。基于此，本研究构建起了包含 4 个一级维度以及 11 个二级维度的智能时代教师评价素养初始框架（如表 2 所示）。

表2 智能时代教师评价素养初始框架

一级指标	二级指标
A：智能时代教师评价知识	A1：智能技术知识
	A2：教学评价知识
	A3：整合智能技术的教学评价知识
B：智能时代教师评价技能	B1：利用智能技术制定教学评价方案
	B2：利用智能技术实施教学评价活动
	B3：利用智能技术解释和应用评价结果
C：智能时代教师评价惯习	C1：智能时代教师评价信念
	C2：智能时代教师评价行为
D：智能时代教师评价伦理	D1：安全意识
	D2：以人为本
	D3：公平意识

(二) 基于德尔菲法的框架要素修正

本研究共选取了 15 名专家，开展了两轮意见咨询，选取的专家可分为两类：一类是在智能教育评价、教师评价素养与智能素养等领域有一定权威的学者；另一类是对智能技术应用于教学评价有相关实践经验和理论思考的教研员及中小学教师。调查结果表明：15 位专家的权威系数 (Cr) 均大于 0.7，表明评分结果比较可靠；两轮专家咨询后，所有指标均值由 3.67—4.40 增长到 4.53—4.93，变异系数 (CV) 由 0.07—0.20 下降到 0.05—0.11，表明专家协调程度提高。这充分表明各项指标逐步趋于更加合理，专家意见逐步趋于一致。

根据两轮专家意见反馈，对指标结构做了以下修正和补充：(1) 将 B3“利用智能技术解释和应用测评结果”改为“解释和应用智能测评结果”；(2) 将 C1 修改为“智能时代教师评价思维方式”、C3 修改为“智能时代教师评价行为倾向”，并增添维度 C2“智能时代教师评价情感态度”；(3) 将 D1 维度“安全意识”改为“智能时代教师评价隐私安全”，将 D2 维度“以人为本”改为“智能时代教师评价评价以人为主”，将 D3 维度“公平意识”改为“智能时代教师评价公平决策”，并在 D3 维度的解释中增加透明决策相关内容，即教师在选择和使用智能技术工具和资源时，应保持决策过程透明。

(三) 智能时代教师评价素养框架的正式确立

经过框架的初步构建及框架要素的修正，最终确立了智能时代教师评价素养框架（如下页图 1、表 3 所示），包含智能时代教师评价知识（简称为“评价知识”）、智能时代教师评价技能（简称为“评价技能”）、智能时代教师评价惯习（简称为“评价惯习”）、智能时代教师评价伦理四个一级维度（简称为“评价伦理”）。评价知识在整个素养结构体系中处于基础保障的地位；评价技能是教师有效开展智能教学评价的必要支撑；评价惯习包含教师在智能时代开展教学评价应具备的思维方式、情感态度及行为倾向，对于教师的评价实践起着引导作用；而评价

伦理则反映了智能时代教师评价素养的价值内核。尽管为了结构的清晰性，图中这四个要素是分立的，但在真实的实践场域中它们又在动态交互中相互影响、相互促进，最终构成完整的素养结构模型。

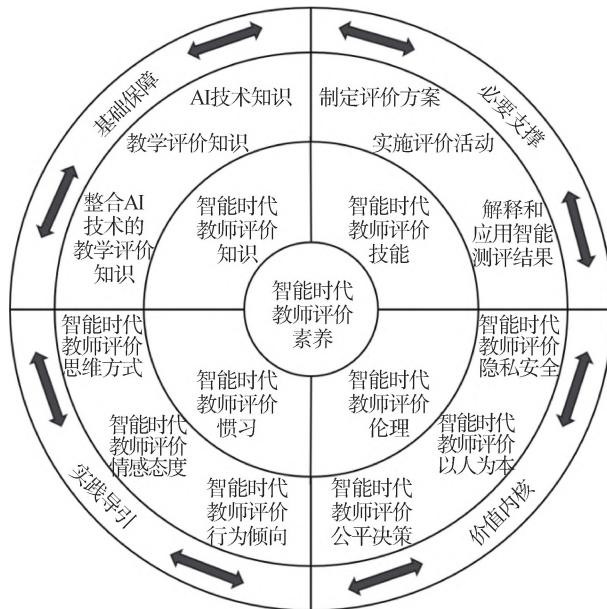


图1 智能时代教师评价素养结构图

表3 智能时代教师评价素养的正式框架

一级指标	二级指标	解释
A：智能时代教师评价知识	A1：智能技术知识	教师应具备应用于教学评价的智能技术、工具、产品的相关知识，包括概念、基本原理、使用方式等
	A2：教学评价知识	开展教学评价本身需具备的理论和实践知识，例如评价的基本概念、典型评价的方法和操作步骤
	A3：整合智能技术的教学评价知识	智能技术融入教学评价的理论及实践知识
B：智能时代教师评价技能	B1：利用智能技术制定评价方案	能够利用智能技术展开评价方案设计的辅助工作，例如能够与 DeepSeek 等生成式智能工具进行高质量交互，让其生成可供参考的评价方案
	B2：利用智能技术实施评价活动	能够合理有效地利用智能技术展开多元评价主体间的交流合作，开展多种形式、多种类型的教学测评等
	B3：解释和应用智能测评结果	能够从智能测评结果中诊断学生学习成果，并制定针对性提升策略
C：智能时代教师评价惯习	C1：智能时代教师评价思维方式	教师在智能时代开展教学评价应持的理念，包括积极的智能技术应用理念，持续反思、改进自身评价实践的理念等
	C2：智能时代教师评价情感态度	教师面对智能技术融入教学评价应持有的情感态度与心理倾向，包括积极接纳智能技术融入、克服技术焦虑及抵触
	C3：智能时代教师评价行为倾向	教师在智能时代开展教学评价过程中应呈现出的行为倾向，包括积极应用智能技术，关注学生的综合素质、个性特质、学习过程等

续表3

D：智能时代教师评价伦理	D1：智能时代教师评价隐私安全	教师在智能时代开展教学评价过程中应具有隐私风险防范意识，遵守相应的隐私伦理和数据安全规约，避免隐私数据泄露
	D2：智能时代教师评价以人为本	教师应理性、审慎地看待智能技术功效和人机关系，坚持以人为本，注重智能时代教学评价过程中学生主体性的发挥
	D3：智能时代教师评价公平决策	教师应意识到智能时代开展教学评价过程中存在的公平隐患，结合多元评价方法弥补智能评价的量化局限性，对算法偏见有所预估，并保持决策的透明性

1. 评价知识：智能时代教师评价素养的基础保障

评价知识是智能时代教师评价素养结构中的基础保障，包括开展教育评价实践需要具备的陈述性知识和程序性知识。本研究拟借鉴 TPACK 框架解读智能时代教师应具备的评价知识。TPACK 有七个要素，即技术知识(TK)、教学知识(PK)、学科知识(CK)、学科教学知识(PCK)、整合技术的教学法知识(TPK)、整合技术的学科内容知识(TCK)与整合技术的学科教学法知识(TPACK)^[15]。伴随智能技术的发展，AI-TPACK 框架应运而生，TPACK 中的技术知识演变为智能技术知识。由于本研究探讨的评价素养不指向具体学科，因此评价知识主要与 TK、PK、TPK 三项要素有着紧密关联，与之相对应，可将评价知识分为智能技术知识、教学评价知识以及整合智能技术的教学评价知识^[16]。

首先是智能技术知识，与 TPACK 中的技术知识(TK)相对应，即教学评价领域中常见的智能技术、工具、产品、资源、平台的相关知识。智能时代的教师应对这些知识具有基本的了解。这些知识既包括陈述性知识，例如学习分析技术、智能识别技术等常见智能技术的概念、原理；也包括程序性知识，即常见的智能评价工具、产品的工作原理和使用方式，例如如何向 DeepSeek 进行高质量提问以获得自己想要的回答。

其次是教学评价知识，与 TPACK 中的教学法知识(PK)相对应。教学法知识(PK)即教师所具备的一般性的、与学科无关的教学策略和教学活动方面的知识。教学评价活动是教师开展的教学活动的重要组成部分，因此教学法知识本身也包含教学评价知识。其内容既包括关于教学评价的陈述性知识，例如教学评价和测量领域的基本概念和术语，常用教学评价工具的概念原理，对于教学评价原则的理解，各类教学评价方法的发展历程、原理、功能与局限；也包括关于教学评价的程序性知识，即如何做评价的知识，例如如何确定评价目标、设计评价方案、实施评价方案和处理评价结果，表现性评价、增值性评价等各种教学评价方法的操作步骤等。

最后是整合智能技术的教学评价知识，与 TPACK 中

的整合技术的教学法知识(TPK)相对应。TPK一般意义上指将具体技术应用于教学活动时,教师对于教学和学习将发生何种改变的一种理解^[17]。这类知识是智能技术知识和教学评价知识有机结合的产物。其内容既包括理论性知识,例如智能技术与教学评价整合的价值、内涵、路径、场景及伦理规范等;也包括与具体评价情境紧密结合的实践性知识,即如何在教学评价过程中合理有效地使用智能技术,通常为教师在智能时代开展教学评价过程中通过体验和反思获得的,与实践情境高度关联。例如将智能技术应用到教学评价过程中的流程、步骤、时机和方法类知识,如何针对不同的测评方法、测评内容以及学生的特征选择合适的智能技术的知识,以及如何利用智能技术解决教育评价实践中某些具体问题的知识等。

2. 评价技能: 智能时代教师评价素养的必要支撑

评价技能是智能时代教师评价素养在实践中有效展现的支撑,是综合运用各种评价知识后表现出来的某种实践行为^[18]。杨卓将教师评价技能分为四个领域:制订评价目标、撰写评价方案、实施评价活动、解释和应用评价结果^[19]。人工智能时代,教师需能够将智能技术应用与教学评价实践相结合,由此衍生出一系列新的实践技能需求。目前来看,在教学评价领域,智能技术的应用主要集中在制定评价方案、实施评价活动以及解释和应用评价结果三个领域。因此,本研究从以上三方面解读智能时代教师应具备的评价技能。

第一是利用智能技术制定教学评价方案的技能,主要指教师能够合理有效地利用智能技术帮助自己设计评价方案。制定评价方案大体包括拟定评价目标、确定评价内容、选取评价方法、如何呈现评价结果等内容。一些智能工具可以充当评价方案设计的得力助手,这其中最有可能用到的便是以 DeepSeek 为代表的生成式人工智能。这要求教师能够对其进行高质量提问,让其按照自己的要求生成可供参考的评价方案,并为自己提供评价方案设计需要用到的资料。

第二是利用智能技术实施教学评价活动的技能。该方面技能主要指教师能够合理有效地利用智能技术和平台展开多元评价主体间的交流合作,开展多种形式、多种类型的教学测评,以及收集、整理、分析学生多方面数据信息。例如,教师可以利用物联网感知、视频监控、在线学习平台等智能工具实现多模态教育教学数据的动态抓取,还可以利用教育数据挖掘和学习分析等算法技术,对多模态教育数据进行加工和分析。

第三是解释和应用智能测评结果的技能。智能测评结果是教师在教学评价过程中应用智能技术、工具对学

生各方面发展状况进行记录、分析所得到的结果。解释和应用智能测评结果,首先要求教师能够从智能测评工具呈现出的分析结果中提取出关于学生的有效信息,例如学生现阶段的发展水平、某一时间段内的发展趋势、优势点与薄弱点、学生群体之间发展水平的差异、学生群体内部发展的平衡程度等等。其次,要求教师能够根据智能测评结果归纳原因,找准不足,制定针对性干预措施和提升策略。最后,要求教师能够将智能测评结果和提升策略以平实易懂的方式传达给学生、家长和教育行政人员。

3. 评价惯习: 智能时代教师评价素养的实践引导

评价惯习起着引导智能时代教师评价实践的作用。“惯习”这一概念最初由布迪厄所建构,指通过结构形塑机制将场域中的规则、价值和要求内化于人的思想意识中,构成一种身心图式和行为系统^[20]。“教学惯习”指教师在长期工作、生活和实践中,基于对外部教学的理解,形成一种内化结构和行为倾向系统,并对教育实践行为起着主导作用^[21]。教学惯习与特定的教学实践环境紧密结合,对于教师来说,他们在传统教学环境中形成教学惯习很难被移植到新的智能化教学环境中。因此,智能时代的教师需对其传统的教学评价惯习进行反思与解构,并通过不断吸收在智能时代开展教学评价实践经验的过程,形成与智能时代教学评价相适应的评价惯习。

就结构方面来看,Wacquant 提出了惯习的三个维度:认知、情感和实践^[22]。认知层面即惯习中的思维方式部分,情感层面即教师的情感倾向和心理态度,实践层面即教师的行为倾向。基于此观点,本研究拟从上述三个方面理解评价惯习。首先是评价思维方式,其内核是教师在智能时代开展教学评价应持的理念,主要有以下内容:其一,教师应具有积极的智能评价理念,认为在教学评价中应积极应用智能技术以发挥其优势。其二,教师愿意改变“唯分数论”,并改善以筛选、比较、结果取向为主导的教学评价理念,转向全面性、过程性、个性化的评价理念,关注学生的综合素质、学习过程以及个性特征。其三,教师应具有创新理念和专业成长意识,具备创新智能教学评价内容、流程、方式和方法的积极意识;以及持续反思、改进自身智能时代教学评价实践的意识。

其次是评价情感态度,主要指教师面对智能技术融入教学评价应持的情感态度和心理倾向。在智能技术融入教学评价的过程中,许多教师产生了一定程度的智能技术焦虑^[23]。技术焦虑使得教师更容易对智能教学评价产生负向、消极的情感态度,并持有较低的自我效能感,最终影响到教师对智能时代教学评价的正确理解,也影响到教师对智能时代评价知识与技能的有效掌握。因此,

面对智能技术融入教学评价，教师应具有积极、正向情感态度，在心理层面积极接纳智能技术的融入，克服技术焦虑、恐惧与抵触，在应用智能技术开展教学评价的过程中逐步建立起良好的自我效能感。

最后是评价行为倾向。在惯习理论中，评价行为倾向与教师的评价思维方式相匹配，大致有以下几个方面：其一，教师应在教学评价中积极应用智能技术，力求突破传统的评价方式、评价手段和评价思维，将智能技术融入到教学评价过程中。其二，教师应改善“唯分数论”和结果导向的评价实践，在智能时代教学评价过程中更多关注学生的综合素质、个性特征以及学习过程。其三，教师在评价过程中应持续学习、反思、改进和创新。具体表现为：勤于思考、勇于创新，主动探索利用智能技术解决教学评价实践问题的新方法、新路径、新手段；在评价过程中与时俱进，持续反思自身智能技术应用情况，总结正确做法，改进不当策略；不断学习智能评价新的理论成果和技术工具，并积极开展相关实践应用探索。

4. 评价伦理：智能时代教师评价素养的价值内核

作为教师开展评价活动需遵循的伦理规范及其背后所蕴含的价值导向，评价伦理是评价素养结构体系的价值内核。伴随智能技术的融入，教育评价领域产生了一些新的潜在伦理风险，例如：(1)安全风险：主要体现为数据泄露风险，导致评价相关者的隐私安全受到威胁；(2)人本风险：主要体现为技术依赖症，使评价主体将评价权力完全让渡给机器，使包括教师与学生在内的评价者的主体性受到损害；(3)公平风险：一是数据主义至上，“量化一切”的思潮泛滥。数据化教育评价看重的是评价主体可测量的外显行为，人的思想、情感、生命体验等不易被量化的容易内容被评价排除在外，导致新的评价不公问题的出现。二是算法偏见导致评价程序及结果的公正受损。算法设计可能本身带有社会偏见，算法训练数据的漏洞以及算法设计者自身可能带有偏见^{[24][25]}。

基于教师智能伦理的已有研究和智能评价的伦理风险，本研究将智能时代教师应遵循的评价伦理准则划分为以下维度：(1)隐私安全：教师在智能时代开展教学评价过程中应具备隐私风险防范意识，遵守相应的隐私伦理和数据安全规约，避免学生隐私数据的泄露；(2)以人为本：教师应理性、审慎地看待智能技术的功效、以及教学评价中的人机关系，坚持以人为本，不过分依赖智能技术，在智能时代开展教学评价的过程中注重人的主体性的发挥，赋予学生一定的评价权力^[26]；(3)公平决策：一是要认识到运用智能技术开展教学评价的量化局限性，有意识地利用多元教育评价方法弥补智能量化评价的局限，在评价过

程中关照每一位学生独特的思想特征、情感经历和人格特质；二是要对算法本身的特征和局限性有所了解，在使用某项人工智能技术或产品开展教学评价活动时，对其算法本身所包含的偏见有所预估，对于其算法生成结果中可能存在偏见与歧视的部分进行纠正；三是要在决策过程中保持透明度，即教师在教学评价过程中选择和使用智能技术工具和资源时，应向学生、家长等评价利益相关主体解释选择技术的理由以及该项技术如何辅助教学评价，包括技术如何支持学习目标，如何监控和评估这些技术的影响，从而使学生和家长能够理解这些技术如何被用于教育过程，并发挥监督技术使用的作用。

三、智能时代教师评价素养框架的应用

基于构建的智能时代教师评价素养框架，通过德尔菲法，多次修改后形成《智能时代教师评价素养调查问卷》，该问卷为李克特 5 级量表，共包含 36 道题目，四个一级维度下各 9 道题目。用问卷星平台进行发放，得到教师问卷 2320 份，其中有效问卷 2290 份，有效回收率为 98.7%。其中，男教师占比 32%，女教师占比 68%；30 岁以下教师占比 33.3%，30—45 岁教师占比 41.2%，45 岁以上教师占比 25.5%；教龄 10 年以下教师占比 41.4%，教龄 10—20 年教师占比 30%，教龄 20 年以上教师占比 28.6%；东部地区教师占比 56.1%，中西部地区教师占比 43.9%；城镇地区教师占比 62.5%，乡村地区教师占比 37.5%。运用 SPSS25.0 软件，对问卷进行信效度检验，得到 Cronbach's α 系数值为 0.984，KMO 值为 0.963，说明问卷具有良好的信效度。

(一) 研究结果

1. 智能时代教师评价素养情况的总体描述分析

结果显示，智能时代中小学教师评价素养总体得分为 3.75，达到中等水平，评价伦理得分最高，为 4.01，其次是评价惯习，得分为 3.94，评价知识和评价技能得分水平较低，分别为 3.55 和 3.43。这表明教师在评价伦理和评价惯习维度表现较好，但在评价知识和评价技能的理解和掌握方面表现不太理想。

2. 智能时代教师评价素养的组群差异比较

为进一步分析智能时代教师评价素养的表现，研究以性别、年龄、教龄、学历、任教学科、城乡和地区(中西、东)为控制因素进行差异检验。结果发现：智能时代教师评价素养在年龄、教龄、城乡、地区上存在显著差异：(1)年龄方面，30 岁以下的教师评价素养水平显著高于 30 岁以上的教师($P < 0.001$)。其中，45 岁以下的教师评价伦理水平显著高于 45 岁以上的教师($P < 0.001$)，30 岁以下的教师评价惯习水平显著高于 45 岁以上的教

师($P<0.001$)；(2)教龄方面，教龄 20 年以下的教师评价素养水平显著高于教龄 20 年以上的教师($P<0.001$)。其中，教龄 20 年以下的教师评价惯习和评价伦理水平均显著高于教龄 20 年以上的教师($P<0.001$)；(3)地区方面，东部地区的教师评价知识和技能水平显著高于中、西部地区的教师($P<0.001$)；(4)城乡方面，城镇地区的教师在总体水平以及四个维度上均显著高于乡村地区的教师($P<0.001$)；(5)在性别、任教学科、学历层面，教师评价素养不存在显著差异。

(二) 结论和讨论

1. 本研究所构建的智能时代教师评价素养框架科学合理

目前，学界基于智能化背景开展的教师评价素养研究数量较少，尚未形成成熟的理论模型。本研究通过对教师评价素养代表性理论成果的梳理，结合其研究最新发展趋势，并基于智能时代教师开展教学评价所面临的真实实践情境及需求，提出“评价知识+评价技能+评价惯习+评价伦理”的智能时代教师评价素养框架，不仅涵盖知识与技能等基础要素，还创新性地融入了评价惯习这一要素，包含智能时代教师开展教学评价的理念、态度及行为倾向；并将教师评价伦理与智能伦理的内涵相融合，凸显了智能时代教师评价伦理这一关键要素。该框架适用于学校和教育机构开展的智能时代教师评价素养培训、评估以及教育政策、教师专业发展规划的制定。同时，本研究基于该框架构建了配套的测评问卷，开展了初步的测评应用，所得出的结论如下。

2. 智能时代教师评价素养总体达到中等水平，评价惯习和伦理水平较高，评价知识和技能水平较低

首先，智能时代教师评价素养总体达到中等水平，评价惯习和伦理水平在四个维度中得分较高，这一结果与许多相关主题的研究基本一致。例如，学者林书兵针对教师数字素养的研究^[27]、学者宋乃庆针对教师智能教育素养的研究^[28]，均得出教师在意识和伦理维度得分较高的结果。造成上述结果的原因在于：近些年来我国各地区对于人工智能融入教育愈发重视，实施了一系列富有成效的举措，例如开展智能教学相关培训、建立智能教学实验基地等；再加之网络媒体对人工智能相关内容的广泛传播，使得教师对于智能评价伦理规范以及应持的价值理念有了更深入的认识。与此同时，教师评价知识和技能水平较低的结果也被许多相关主题研究所证实^[29]，造成这一结果的主要原因在于：当前有关智能技术的教师培训偏重理论讲解，相关实践项目数量欠缺、且缺乏系统性和针对性，并且当前我国智能教学评价实践场域构建还不够完善，尚有许多学校缺乏智能教学所需的人财物力资源，这使得教师在培训

中缺乏实践应用机会，难以获得智能技术知识与应用技能的有效提升^[30]。并且，智能技术发展迅速，导致相关评价知识技能更新迭代较快，这使得教师掌握新兴的智能评价知识技能存在难度。

3. 智能时代教师年龄和教龄更低的评价素养水平更高，尤其体现在评价惯习和伦理维度

已有研究指出，不同教龄的教师在智能教育素养、数智素养上存在显著差异，教龄小的教师更有优势^[31]。这与本研究的结果大致吻合。本研究更进一步发现，智能时代年龄和教龄更小的教师评价素养总体水平更高，尤其体现在评价惯习和伦理维度。究其原因在于，年轻教师多为数字原住民，对数字环境更为熟悉，也更容易对智能技术这一新鲜事物产生兴趣，有着更强的学习欲和探索欲，对智能技术应用于教育评价有着更加积极的情感态度，自然也表现出更积极的行为倾向；同时，出于对智能技术更多的了解，他们对智能技术应用于教育评价的价值理念和伦理风险也有更深刻理解。但与此同时，不同年龄和教龄阶段的教师在智能时代评价知识技能层面的水平并未呈现出显著差异。其原因可能在于，尽管年轻教师对于智能知识与技能的掌握要优于中老年教师，但其在教学评价方面的实践经验相对较少，而智能时代的教师评价知识与技能要求教师将教学评价知识技能与人工智能知识技能有效整合起来。因此，年轻教师并未在评价知识与技能维度上占据明显优势。

4. 智能时代东部地区的教师评价知识和技能水平显著高于中西部地区的教师，评价惯习和评价伦理无显著差异

以往研究较少针对教师智能素养的地区差异展开探讨，但有一些研究指出，东部地区教师的数字素养显著高于中西部地区^[32]，这与本研究结果大致相同。相较于中西部地区，东部地区经济、科技发展水平较高，人工智能基础设施较为完备，相关企业数量较多，学校具备更加便利的智能技术获取和应用条件，教师在工作和生活中接触智能技术的频率更高，接受的智能技术教学培训频次更高、内容更丰富、更具有实操性和系统性。而中西部地区学校的人工智能基础设施较为欠缺，也缺乏相关企业单位的技术支持，教师在日常工作和生活中较少会接触和使用人工智能技术、产品，所接受的教师培训也缺乏针对性和实操性。不过值得一提的是，在评价惯习和评价伦理两个维度，二者之间不存在显著差异，甚至中西部教师的得分还略高。这说明经过近些年来智能技术的发展普及和教师培训工作的开展，无论是东部教师还是中西部教师都对智能技术应用于教学评价的价值理念、伦理风险有了更深刻的理解，都对智能技术应用

于教学评价持较积极的情感态度和行为倾向，二者间的差距主要在知识技能层面。

5. 智能时代城镇教师评价素养总体水平及各维度水平显著高于乡村教师

在智能时代教师评价素养的四个维度上，城镇教师的水平均显著高于乡村教师。这一结果与先前众多研究相符合，反映了我国城乡之间的教育差距。首先，城乡之间的人工智能基础设施和智能教育资源存在显著差距，城镇教师在教学评价过程中应用智能技术、产品的便捷性要远高于乡村教师，有着更加丰富的智能评价实践机会。其次，许多乡村教师的教育理念较为滞后，对新技术的应用态度不够积极。大部分乡村教师认为在乡村中小学授课能力要求低，已有知识和专业能力已经足够应对，不需要额外学习智能技术知识^[33]，这导致乡村教师学习智能评价技术的意愿较低。最后，乡村教师在智能技术相关培训中处于劣势地位。一些新理念新技术的培训优先选拔已经具备实践条件的城市骨干教师参加^[34]，乡村教师本就缺乏智能技术应用经验，培训机会的缺乏使得其与城镇教师的智能时代评价素养水平差距进一步扩大。

(三) 对策建议

1. 构建政企校协同参与机制，打造实践导向的智能时代教师评价素养培训体系

为提升教师智能素养评价培养的实践性，应整合多方力量，增设智能时代教师评价素养实践培训项目，打造以实践为导向的针对性培训体系。在该体系中，政府应颁布激励政策并搭建协同参与平台；高校专家与中小学教师、教研员协同合作，明确主要实践问题及所需资源，设计与教师真实评价情境相匹配的实践项目；同时，联合相关企业，发挥其资金、技术、产品优势，为智能时代教师评价素养培训提供相应资源。在培训过程中，企业技术人员、高校专家和教研员应协同跟进、提供指导。例如，在以智慧课堂评价为主题的培训体系中，基于政府搭建的平台和颁布的政策，专家和教研员将针对该地区、该学校教师在智慧课堂评价实践过程中遇到的问题(如学生多模态数据采集困难)，设计相应的培训项目；随后与企业合作，购置或研发语言情感识别、眼动追踪等智能产品，为培训提供技术支持。

2. 培育中老年教师积极智能时代评价认知，构建新老教师智能评价合作小组

首先，针对中老年教师群体，应进一步提升智能时代教师评价素养培养的趣味性和实践性。学校可以在政府组织引领和政策支持的基础上，利用相关企业提供的智能产品资源，引入游戏化的智能评价培训项目，并加

强项目与中老年教师在评价实践中面临的主要问题的对接。例如，通过智能技术辅助作业批改，帮助教师减轻大班额作业批改的负担，让中老年教师能够切实感受到应用智能技术进行教学评价的乐趣和实践效用。其次，可以构建新老教师智能评价协同合作小组，将不同年龄段的教师编入同一小组，发挥教师学习共同体的效应。一方面，通过交流互动和意见共享，老教师可以在年轻教师的带动下提升智能评价的理论和实践认知，形成积极的情感态度；另一方面，充分发挥老教师在教学评价中的经验作用。例如，在人机协同评价中，新教师负责模型调配与人机交互，老教师则负责核查智能机器给出的评价结果并提出优化建议。

3. 完善区域交流帮扶机制，聚焦中西部智能时代教师评价知识技能水平提升

首先，政府可以出台相关政策，引导东部地区与中西部地区建立合作帮扶机制，形成稳定的结对区域或结对学校，加强区域之间在人工智能评价知识与技能方面的经验交流与实践合作。具体而言，东部地区可以帮助中西部地区完成本地化模型的部署，派遣优秀的中小学教师和专家技术人员，基于需求分析采集中西部地区的教学评价数据，利用 DeepSeek 大模型进行数据训练与调配，并指导中西部教师如何向大模型提出问题，协助开展评价工作。其次，在政府主导和企业协作的基础上，打造智能评价优质资源共享平台，免费向中西部地区学校开放东部地区的优质教育资源，为中西部教师学习智能评价相关知识和技能提供便捷条件。同时，政府可通过税收减免、财政补贴、宣传报道等手段，鼓励人工智能和互联网相关企业在中西部地区拓展市场，引导其在该地区设立试验区，并投放智能评价相关新产品，推动中西部地区学校智能教学设施的完善与升级。

4. 增强乡村地区资源倾斜，打造乡土化智能时代教师评价素养培育项目

首先，政府可以设立专项资金并出台优惠政策，联合企业、高校、社会组织等各方力量，进一步加强乡村地区智能教育资源的供给。当地教育部门应设立专门的乡村智能时代教师评价素养培养计划，并确保一定数量的乡村教师能够参加市级、省级的智能教学培训。其次，政府可完善城乡帮扶机制，鼓励城镇地区学校与乡村学校结对帮扶，并在高校和企业的技术支持下，打造城乡学校智能评语库和课堂行为分析模型共享系统。最后，政府与乡村学校应加强智能时代乡村教师评价意识的培养，并推出具有本地特色的智能时代教师评价素养提升项目。在乡村智能时代教师评价素养的培养过程中，应将教师培训项目与乡村教学特色的评价实践场景相结合，

让乡村教师切身体会到智能评价的优势。例如，针对复式教学、田间课堂等乡村特色教学场景，开展定制化的智能评价培训项目，指导当地教师在这些教学场景中运用智能技术采集评价数据，进而构建符合乡村实际的智能评价模型。

四、不足与展望

量化研究与质性研究各有优势，前者胜在大规模群体总体情况的把握，后者长于案例特征的深度揭示，二者通常难以兼顾。本研究重在从宏观层面把握智能时代教师评价素养的整体情况，在个案实践行为分析上不足。未来研究可以采用混合研究法，在真实的智能教学情境中，融入课堂观察、案例分析等质化研究方法开展更为全面、生动的研究。

参考文献：

- [1] Deluca C,Lapointe-mcewan D,Luhanga U.Teacher assessment literacy:A review of international standards and measures [J].Educational Assessment,Evaluation and Accountability,2016,28(3):251–272.
- [2] 辛涛,洪倩等.新时代教育评价体系的价值定位:国际趋势与中国方案 [J].国家教育行政学院学报,2024,(2):13–21.
- [3] 王爱玲.试论有效课堂评价体系的构建 [J].教育理论与实践,2014,34(34):52–55.
- [4] Stiggins R J.Assessment Literacy [J].Phi Delta Kappan,1991,72:534–539.
- [5] Brookhart S M.Educational assessment knowledge and skills for teachers [J]. Educational Measurement:Issues and Practice,2011,30(1):3–12.
- [6] AFT,NCME,NEA.Standards for teacher competence in educational assessment of students [J].Educational Measurement Issues and Practice,1990,9(4):30–32.
- [7] Willis J,Adie L,et al.Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform [J].Australian Educational Researcher,2013,40(2):241–256.
- [8] Looney A,Cumming J,van Der Kleij F & Harris K.Reconceptualising the role of teachers as assessors:teacher assessment identity [J].Assessment in Education: Principles,Policy & Practice,2017,25(5):442–467.
- [9] 郑东辉.新时期教师到底需要怎样的评价素养 [J].教育发展研究,2022,42(4):46–51.
- [10] 吴龙凯,程浩等.技术伦理视角下人机协同教育评价的运行机制与实践策略 [J].中国电化教育,2025,(1):8–16.
- [11] 谢小蓉.新时代教师效能评价的内涵重构与发展取向 [J].教师教育学报,2025,12(2):10–19.
- [12] 祝智庭,彭红超.创新发展技术赋能的智慧教育——访我国智慧教育开拓者祝智庭教授 [J].教师教育学报,2021,8(4):21–29.
- [13] 张红艳,连雅迪.数智教育评价范式转变动因、特征与实现路径研究 [J].黑龙江高教研究,2024,42(1):23–30.
- [14] Interstate New Teacher Assessment and Support Con-sortium(INTASC)Core Standard [EB/OL].<http://www.ecu.edu/cs-educ/teached/upload/INTASC StandardsIndicators.pdf>,2016–03–26.
- [15] Koehler M J,Mishra P.What happens when teachers design educational technology?The development of technological pedagogical content knowledge [J].Journal of Educational Computing Research,2005,32(2):131–152.
- [16] 闫志明,付加留等.整合人工智能技术的学科教学知识(AI-TPACK):内涵、教学实践与未来议题 [J].远程教育杂志,2020,38(5):23–34.
- [17] Cox S,Graham C R.Diagramming TPACK in practice:Using an elaborated model of the TPACK framework to analyze and depict teacher knowledge [J].TechTrends,2009,53(5):60–69.
- [18] 杨道宇.教师评价能力的内涵、结构与行为表现 [J].教育科学,2024,40(4):73–79.
- [19] 杨卓,李佳等.我国高中教师评价素养现状调查研究 [J].中国考试,2024,(5):57–68.
- [20] 宫留记.布迪厄的社会实践理论 [M].开封:河南大学出版社,2009.147–148.
- [21] 许昌良.教师教学惯习转变的理性思考及其实现 [J].天津市教科院学报,2013,(4):86–88.
- [22] Wacquant L.Homines in extremis:What fighting scholars teach us about habitus [J].Body & Society,2014,20(2):3–17.
- [23] 王天平,李珍.智能时代教师技术焦虑的形态、动因与对策 [J].电化教育研究,2022,43(10):110–115+128.
- [24] 鹿星南,高雪薇.人工智能赋能教育评价改革:发展态势、风险检视与消解对策 [J].中国教育学刊,2023,(2):48–54.
- [25] 郑永红,王辰飞等.生成式人工智能教育应用及其规制 [J].中国电化教育,2024,5(5):114–119.
- [26] 郑东辉,毛玮洁.论教师评价伦理素养的培育 [J].教师教育研究,2024,36(5):16–21.
- [27] 林书兵,姜雨晴等.中小学教师数字素养的基本现状、影响因素与提升策略——基于珠三角等地的调查与分析 [J].中国电化教育,2025,(2):84–91.
- [28] 郑智勇,宋乃庆.中小学教师智能教育素养测评指标体系构建研究 [J].中国电化教育,2023,(12):75–83.
- [29] 乔世伟,董艳等.数字化转型背景下教师数字素养:现状、影响因素与提升策略 [J].教育科学研究,2025,(4):13–19.
- [30] 彭泽平,冯橙.智能时代的教师专业发展:诉求、困境与实践进路 [J].教育学术月刊,2024,(2):98–105.
- [31] 冯剑峰,姜浩哲等.面向人机协同的教师数智素养:测评框架、现状审视与优化路径 [J].教育发展研究,2024,44(10):21–29.
- [32] 薛芳芳,张源源.农村中小学教师数字素养的调查研究 [J].中国考试,2024,(12):74–82.
- [33] 吴云鹏.乡村振兴视野下乡村教师专业发展的困境与突围 [J].华南师范大学学报(社会科学版),2021,(1):81–89+195.
- [34] 徐莉莉,高葵芬.农村小规模学校新教师专业发展的困境与对策 [J].现代中小学教育,2018,34(9):64–67.

作者简介：

李毅：教授，博士生导师，研究方向为教育数字化、教育评价。

郑鹏宇：在读硕士，研究方向为教育评价、课程与教学论。

Conceptual Reconstruction, Value Explication, Framework Development, and Application of Teachers' Assessment Literacy in the Intelligent Era

Li Yi^{1,2}, Zheng Pengyu¹

1.Faculty of Education, Southwest University, Chongqing 400715

2.Basic Education Research Center of Southwest University, Chongqing 400715

Abstract: Assessment literacy is an important part of teachers' professional quality. Arrival of the intelligent era promotes changes in the field of educational assessment, and also puts forward new requirements for the connotation structure of teachers' assessment literacy. This study first reconstructs the connotation of teachers' assessment literacy in the intelligent era, then analyzes its value, and develop a structural framework comprising four first-level dimensions named Intelligent-era teachers' assessment knowledge, Intelligent-era teachers' assessment skills, Intelligent-era teachers' assessment habitus, Intelligent-era teachers' assessment ethics and 12 second-level dimensions. Based on the established framework, this study developed an assessment instrument and conducted a nationwide assessment. The key findings reveal that:(1)Intelligent-era teachers' assessment literacy demonstrates a middle proficiency level, with relatively high performance in assessment habitus and ethics dimensions, yet comparatively lower levels in assessment knowledge and skills;(2)Younger teachers with less teaching experience exhibit superior assessment literacy, particularly evident in the dimensions of assessment habitus and ethics;(3)Teachers in eastern China show significantly higher levels in assessment knowledge and skills compared to their counterparts in central and western regions, while no significant difference in assessment habitus and ethics;(4)Teachers in urban areas have significantly higher overall levels of assessment literacy and levels in each dimension compared to those in rural areas. Building on these findings, the study further proposes evidence-based recommendations for targeted intervention strategies.

Keywords: intelligent era; assessment literacy; teaching assessment; intelligence technology

收稿日期:2025 年 3 月 5 日

责任编辑:宋灵青 张晖 张雪

(上接第 86 页)

How Can Technological Resilience Empower Rural Education Revitalization

Wu Yixiong, Long Zhou

China Digital Agriculture and Rural Research Institute, Huaihua University, Huaihua 418000, Hunan

Abstract: Technological resilience, as the steel beam supporting the revitalization of rural education, is reshaping the development form of rural education with its unique adaptability, flexibility, and sustainability. Using technological resilience as a prism, resilience theory as a foundation, and logical reasoning as a method, this four-dimensional analysis framework is based on theoretical interpretation, logical reasoning, practical difficulties, and breakthrough paths. It reflects the essence of the connotation of technological resilience empowering rural education revitalization, the analysis of generative logic, the manifestation of practical difficulties and the description of development paths. Research has found that technological resilience mainly empowers four areas: restructuring educational resources, enhancing teacher capabilities, innovating teaching models, and optimizing governance systems. The gears of these transformations require the joint drive of three driving forces: technological development, policy support, and social environment. However, the foundation of infrastructure is still shallow, the lack of passive teacher digital literacy, the outdated framework of educational governance, and the emergence of risks in technological ethics are all testing the quality of technological resilience. To this end, five paths have been constructed: strengthening technological infrastructure, cultivating teachers' digital literacy, optimizing educational governance blueprints, to open up a new era of rural education revitalization.

Keywords: technological resilience; rural education revitalization; educational technology; smart education; educational equity

收稿日期:2025 年 5 月 5 日

责任编辑:邢西深