

技术赋能表现性评价:框架设计与实践路径

张尧¹, 左明章²

(1.华中师范大学 人工智能教育学部, 湖北 武汉 430079;

2.华中师范大学 信息化与基础教育均衡发展省部共建协同创新中心, 湖北 武汉 430079)

[摘要] 技术为教育评价改革提供了条件,也为表现性评价的创新发展带来前所未有的机遇。据此,文章探讨了表现性评价的价值取向,分析表现性评价在实践中存在的表现数据收集尚不全面、评价过程主观性较强且耗时耗力、评价结果的应用与反馈效果欠佳等问题,并在此基础上构建了技术赋能表现性评价的实践框架,包括:坚持以育人目标为指引,确保评价设计与实施的一致性;提高数据采集质量,保障评价的客观性;推动人机协同支持的数据分析,提供发展性结果与反馈;促使学生开展反馈调节,引领个体自我发展。在实践框架的基础上,进一步提出了技术赋能表现性评价的四大实践路径,即构建技术赋能表现性评价支持环境,支撑评价工作的开展;完善技术标准体系框架,促进评价的规范与成熟;优化评价过程与技术工具,推动评价的可持续发展;建立技术赋能表现性评价运行机制,确保评价的有效运行。

[关键词] 技术赋能;表现性评价;课程教学;发展性评价;评价改革

[中图分类号] G434

[文献标志码] A

[作者简介] 张尧(1995—),男,河南郑州人。博士研究生,主要从事学生评价、人工智能教育应用研究。E-mail: yaochang@mails.cnu.edu.cn。

一、引言

随着我国教育教学改革的发展与新时代教育评价改革的深入推进,表现性评价(Performance Assessment)作为一种超越传统标准化测验的评价方式,受到广泛关注。表现性评价为当前我国学生评价改革提供了有效途径,已在各级教育的学生评价体系中得到应用,旨在衡量学生的学业表现与技能掌握情况,评价学生的高阶思维和素养等,并进一步支持国家义务教育质量监控。然而,表现性评价仍存在数据采集范围有限、主观评价偏差和反馈针对性不强等关键问题^[1]。随着科学技术的快速发展,尤其是人工智能技术与教育领域的深度融合为上述问题提供了解决方案。中共中央、国务院印发的《深化新时代教育评价改革总体方案》、教育部印发的《基础教育课程教学改

革深化行动方案》均强调要充分利用信息技术,提高教育评价的科学性、专业性、有效性。现有部分研究阐释了智能技术可在一定程度上改善评价过程与方法、创新评价内容与模式、提升评价的数字化水平与评价质量等^[2-3]。因此,本研究基于表现性评价的价值取向及其在实践中存在的主要问题,尝试以数据采集、评价过程与评价反馈为抓手,构建并阐述技术赋能表现性评价的实践框架,探讨其具体的实践路径,以期推动智能时代表现性评价的创新发展。

二、表现性评价的价值取向与实践问题

(一)表现性评价的价值取向

在二十世纪标准化测验盛行之际,以建构主义学习理论、多元智能理论为代表的教学理论与实践催生了对新型教育评价的需求^[4]。随着教育教学改革的推

基金项目:2022年度教育部人文社会科学重点研究基地重大项目“指向核心素养的课堂学习研究”(项目编号:22JJD880026);2021年度信息化与基础教育均衡发展省部共建协同创新中心研究项目“面向智能时代的农村中小学表现性评价体系研究”(项目编号:xtkjr2021-004)

进,人们开始质疑并反思标准化测验的局限性,如测验内容关注低阶的知识与技能、忽略学生差异、与实际学习生活脱节以及难以全面评价学生能力等。因此,表现性评价作为应对标准化测验局限性的一种方式而再度兴起,已在许多国家、地区和学校中作为大规模外部管理评价项目或课程教学评价方式得到应用。在概念定义方面,布兰德(Bland)等认为,表现性评价通过创设一定的任务情境,不仅能够引发、记录和评价学生在任务情境中的过程表现与结果,而且还能以超越实际分数的方式为学生提供有价值的学习体验、兴趣与动机^[9]。波帕姆(Popham)认为,表现性评价是通过让学生学会完成某一特定任务来测量学生的状态,该过程既要求学生生成建构反应,评价者也要观察学生的建构过程,并对其行为表现进行评判^[6]。本研究认为,表现性评价是指评价学习者在任务情境中的过程表现与结果。它注重学习者的行为表现、知识的理解程度^[7]、高阶思维能力^[8]和核心素养^[9]等评价目标,强调通过学习者完成任务的表现,衡量其在评价目标方面的真实水平。

表现性评价的回归与兴起反映了教育对多元、多维的高阶思维能力与素养评价的需求,体现了评价从重视选拔性评价向强调发展性评价的转变,是培养适应未来社会需求人才的必然选择。表现性评价强调评价要回归育人本位,着力于实现立德树人的根本任务。它倡导从单一的知识考核转向能力与素养评价,遵循人才成长规律。表现性评价将高阶思维能力、素养等有意识地贯穿于评价设计与实施的全过程,发挥评价的“指挥棒”作用,进一步促进教学过程、教学内容等发生深刻变化。表现性评价注重多元参与的评价体系,尊重学校、教师、学生、家长等多元主体的地位,能够促进评价的多元化和全面化,并激发了各方参与评价的积极性。表现性评价从单一的结果导向转变为过程和发展导向,视学生为完整的个体,聚焦于学生的现实需求和未来发展,重视学生的主体性及其与真实生活的关联,强调学生的自我发展意识与自我调节能力。

(二)表现性评价在实践中存在的问题

尽管一线教师和教育研究者在过程性评价和总结性评价中积极探索表现性评价的应用,积累了一定的研究成果,推动了学生评价改革,但表现性评价的研究与实践仍面临诸多亟须解决的问题,这阻碍了表现性评价的应用与发展,也影响了社会对其评价结果的认同^[10]。

一是学习表现数据采集不全面。当前,课程教学中实施的表现性评价在技术应用方面较为有限,且多

依靠相对传统的表现性任务,如纸笔形式的表现性任务,其倾向于强调学生最终提交的个人作品,而难以实时、持续地采集学生的过程表现数据等,从而限制了数据采集的广度与深度,不利于对高阶思维能力和素养等评价目标开展客观、科学、全面的评价。

二是评价过程主观性较强且耗时耗力。当前表现性评价主要通过人工评分方式对采集到的学生个人作品等进行评定,该过程耗时耗力,且依赖于评价者自身的判断力以及评价规则的可操作性。评定过程会受评价者主观经验等影响,易产生评价误差或评分者效应^[11],且往往无法在短时间内完成评定工作,难以及时地提供评价结果与反馈。

三是评价结果的应用与反馈效果欠佳。表现性评价常被作为一种过程性或总结性评价手段,评价者倾向于将表现性评价结果直接等同于分数或等级,这在一定程度上忽略了表现性评价的其他作用,如提供详细反馈、关注自身任务表现及其改进、激发学生持续参与等^[12],从而影响了评价反馈的针对性。

三、技术赋能表现性评价的实践框架

技术为表现性评价提供了智能化的新机遇,也为其解决实践难题提供了有效途径。为了更好地利用技术解决上述表现性评价在实践中存在的问题,构建与未来的人才需求、创新人才成长规律相适应的学生评价体系,需要进一步探索技术赋能表现性评价的实践取向。在技术与表现性评价融合创新的过程中,应依循以下理念:一是创新评价方式多元化。表现性评价的多元化包括评价目标、表现性任务、评价主体以及评价功能的多元化,其核心聚焦于促进个体的发展。二是发展数据驱动的评价。使用人工智能、大数据、学习分析等技术采集学生的多元学习表现数据,强调数据在评价中的基础地位,并通过数据分析以优化评价效果。三是教学、学习和评价一致性发展。评价不再是教或学之后的一个孤立环节,而是贯穿于整个教学和学习过程之中^[13],共同促进学生的发展。

在此基础上,研究进一步分析了技术赋能表现性评价的理论基础,旨在为评价实施提供理论依据。“证据中心设计”(Evidence-Centered Design, ECD)是一个基于证据推理原则的系统化教育评价方法,致力于解决复杂技能等高阶思维能力的测量问题,可为评价提供设计开发框架,并倡导积极整合相关技术以获取证据,开展基于证据的推理,被认为是一种数据驱动的教育评价范式^[14]。ECD理论主要由学生模型、证据模型和任务模型三个核心模型构成。其中,学生模型

明确了需要通过评价来测量的目标,以促进可观察行为的操作化;证据模型明确了哪些特点需要被呈现及其所占权重,以及对应的评价规则等;任务模型是指设计基于情境的测评任务以诱发学生在评价指标方面的行为表现,进而获取证据模型所需要的证据。ECD理论与表现性评价相契合,可用于指导高阶思维能力和素养的表现性评价设计^[15]。

综上所述,本研究针对表现性评价在实践中存在的主要问题,提出了技术赋能表现性评价的实践框架,如图1所示。

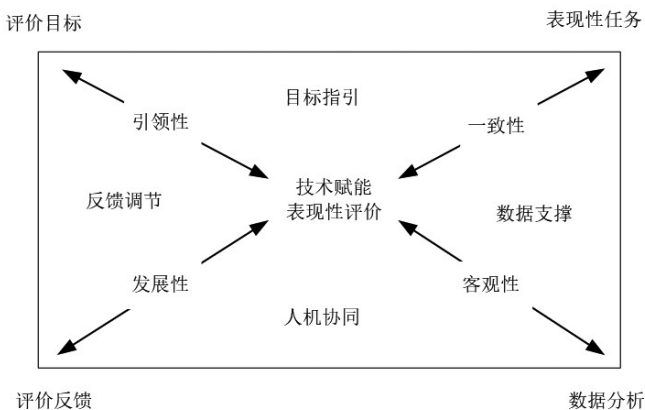


图1 技术赋能表现性评价的实践框架

(一)坚持以育人目标为指引,确保评价设计与实施的一致性

在深化新时代教育评价改革背景下,表现性评价显著的特征之一就是强调指向高阶思维能力与素养发展的育人目标,并将其作为评价的出发点和落脚点。育人目标是教育活动的核心,表现性评价的设计必须围绕如何更好地实现这些目标展开,才能真正发挥育人作用。从ECD理论来看,表现性评价的设计需要先确立指向高阶思维与素养发展的评价目标,并将其转化为具体的行为表现,进而才能开展基于证据的推理过程,其证据的获取、分析与推断应围绕能够直接或间接体现学生高阶思维能力与素养发展的行为表现或信息。基于ECD理论的表现性评价是一个紧密相连的链条,每个环节都是围绕高阶思维能力与素养目标而展开。育人目标是评价实施的出发点和归宿,决定了表现性评价的方向和内容,确保评价内容和方式都服务于学生高阶思维能力与素养目标发展。同时,对多模态数据的探索与应用^[16],不仅为表现性评价提供了强大的技术支持,也为提升表现性评价育人成效奠定了数据基础。例如,朱莎等以信息素养为评价目标,基于ECD理论构建了学生信息素养任务模型,通过基于xAPI的过程性数据采集与存储、观测变量的抽取与处理等,以实现数据驱动的学生信息素

养评价^[17]。

(二)提高数据采集质量,保障评价的客观性

表现性评价的关键在于获取学生的学习表现数据。传统的数据采集方式因资源和技术限制,其数据采集范围窄且效率低,难以充分支持过程评价与复杂学习结果的量化。技术赋能表现性评价可通过智能化采集、大数据、学习分析等技术,对学生学习表现的过程性、多模态数据进行全方位、多层次、立体化的采集,从而更科学、全面、客观地呈现学生的真实情况。技术的效用主要体现在以下四个方面:第一,拓展数据采集的广度。学生学习表现数据不再局限于学生单次提交的作品,还应涵盖学生的课内外、线上线下行为表现、认知与情感等多维数据。学习表现数据的不同来源可以在支持数据互补验证的基础上,全方位理解学生在特定领域的真实水平,确保评价结果的准确性与可靠性。第二,挖掘数据采集的深度。通过丰富表现性任务的呈现方式、利用相关技术系统平台等途径实时采集细粒度的多元学习表现数据,包括学生在线行为和历次作品评分等结构化数据,以及文本、视频和图片等非结构化数据,可为评价提供丰富、有价值的的数据资源,有助于保障数据的完整性与客观性,并提高数据分析与解释的精确度。第三,突破数据采集的时间限制。无论是学生的行为数据,还是在技术系统平台中的交互数据,都可以通过技术手段实现实时采集,获取动态数据流,从而更全面地了解学生在特定领域的变化情况。第四,提高数据采集的效率与准确性。相比于人工数据采集和处理工作,技术赋能表现性评价可通过智能教学平台、大数据分析技术实现伴随式、智能化的数据采集与处理,既减轻了人力和时间成本,也提高了数据采集的效率与准确性。例如,笔者所在研究团队基于小雅智能教学平台实时、自动地采集大学生在专业课程中的13个在线学习行为指标,包括课程任务完成情况、在线讨论、资源浏览、学习时长等,以探究混合式学习情境下大学生在线学习行为特征^[18]。

(三)推动人机协同支持的数据分析,提供发展性评价结果与反馈

如何对获取的数据进行分析以准确衡量学生在特定领域的实际水平,这是技术赋能表现性评价的重点。学习表现数据分析的准确性和科学性直接影响评价的有效性和质量。随着人工智能、大数据和机器学习等技术的发展,表现性评价正在逐渐从人工评价转向人机协同的评价,以发掘数据潜能,提高评价效率与效果。人机协同的评价依靠技术、算法和模型快速

处理海量数据并揭示潜在趋势,然后再由专家基于初步分析结果提出见解和解决策略。这一过程强调体现人的价值,旨在突破技术至上和算法偏见的限制,将人类智慧与机器智能相结合,充分发挥各自的优势,驱动多元学习表现数据分析的实施^[19]。例如,宛平等利用生成式人工智能技术构建了人机协同评价实践模式,并以主观作业评价作为实践案例加以应用,研究发现人机协同评价与教师评价结果之间具有较高的一致性^[20]。另外,通过人机协同支持的数据分析,评价者能够高效地为学生提供发展性评价结果,即关注学生的个体差异,能够呈现其在特定领域的具体优劣势及其成长轨迹^[21],如学生个人发展报告等。教师可利用技术的实时监测和分析功能对采集到的连续性数据进行分析,及时捕捉学生多元学习表现数据的变化趋势,并借助生成式人工智能等技术,进一步为学生提供及时、详细、针对性的反馈与建议。

(四)促使学生开展反馈调节,引领个体自我发展

技术赋能表现性评价的核心在于如何促使学生开展反馈调节以推动自身学习与自我发展。表现性评价与发展性评价在最终目的、功能取向方面较为一致,且更注重过程表现,强调采集和分析关键数据以反映学生的实际水平,针对其优势与不足提供具体建议,以促进学生发展^[22],但是在具体实施过程中,对于评价结果与反馈的应用仍缺乏有效的认识与策略。因此,推动学生合理应用评价结果与反馈以促进其发展是技术赋能表现性评价的重要环节。技术赋能表现性评价主要借助技术手段设计开发评价反馈系统,为学生提供实时的、发展性的评价反馈,发挥评价的育人导向作用,实现以评价促进学习、以评价促进发展。这种评价反馈系统融合了外部评价和自我评价的结果,丰富了评价结果的内容与维度,并将评价反馈作为学生参与后续表现性任务的基础,从而优化评价结果的应用,发挥评价的导向、诊断、调控和改进等作用。具体而言,以生成式人工智能为代表的人工智能技术能够协助教师为学生及时提供发展性评价结果与反馈,并鼓励学生依据相应的评价规则进行自我评价。在外部评价和自我评价的结果与反馈共同作用下,促使学生更清晰地认识自身的表现水平、优势和不足,激发学习动机,并通过自我调节等方式改善学习实践,将这些改进应用到未来的学习实践之中以形成良性循环。例如,杭州市天长小学在语文学科评价中基于学生学习表现数据和学生综合评价系统,生成了学生综合素养发展画像及评价报告,并通过建立学生作品样例库、系统推送与线上展示等方式,激发学生的学习积

极性^[23]。

四、技术赋能表现性评价的实践路径

(一)构建技术赋能表现性评价的支持环境,支撑评价工作的开展

技术赋能表现性评价支持环境的构建可为表现性评价实践提供稳定支持。技术赋能表现性评价支持环境将推动表现性评价从传统经验主导的评价转向数据驱动的评价,使评价过程更科学、客观、准确和高效。研究将从数据采集、处理分析以及评价反馈三个主要技术应用场景出发,支持课程教学中技术赋能表现性评价的实践应用。一是整合人工智能、大数据和智能化采集等技术,对学生在参与任务时所产生的行为表现数据、认知数据、情感数据等进行全过程、伴随式采集。这些数据不仅展示学生的特定领域表现及其结果成效,还反映其变化与成长。此外,利用大数据、智能技术建立统一的数据平台,汇聚全过程、海量的多元学习表现数据,以解决表现性评价中数据限制、量化困难以及客观公正不足等问题^[24]。二是依托学习分析、数据挖掘、人工智能算法等技术对多元学习表现数据进行智能处理、数据深度挖掘和诊断分析,提高评价效率。这种客观和量化的数据分析有助于评价者对具体表现指标进行科学分类,建立不同类别水平所对应的标准参照数据集,精准、客观地描绘学生在特定领域的表现。同时,还可利用自动分析和训练来构建预测模型,从而对学生的个人作品等非结构化数据进行智能评分。三是结合人工智能、可视化技术等,将评价结果进行可视化呈现,生成学生个体的数字画像及个性化发展报告。这不仅让教师、学生、教育管理者和家长能够及时、直观地理解评价结果,还有助于激励学生在学习中开展自我调节。

(二)完善技术标准体系框架,促进评价的规范与成熟

标准体系框架是优化资源配置、提升效率的手段,同时也是衡量评价质量和发展程度的基准。本研究认为,可通过加快制定技术标准、数据标准、质量标准和管理规范,强化对标准的认识和应用,推动评价工作向规范化和成熟化发展。第一,技术标准是技术赋能表现性评价中数据采集、处理、分析及反馈所使用技术的性能效率、运行稳定性和系统兼容性等基本规范。技术标准的研制可通过遵循国家技术标准、借鉴国际技术标准、制定地方技术标准和参考行业技术标准四种形式^[24]。另外,技术标准体系的研制也需结合技术成熟以及教、学、评需求适时调整,以形成一种

动态的、相对稳定的体系结构^[25]。第二,数据标准是保障学习表现数据内外部使用、交换的一致性和准确性的评价规范依据。在数据标准的支持下,多元学习表现数据的采集、存储、汇聚、管理和保护成为可能,有利于促进不同技术系统之间的学习表现数据交换^[26],并在一定程度上实现教学评价体系中多类型数据的汇聚。第三,质量标准是对评价内容的描述与规范。它需与技术赋能表现性评价的特征保持高度一致,并为后续研究实践提供一个实用的分析框架,具体涉及评价设计、信效度、评分一致性等,以及监控评价实施过程的质量,如学生参与度、与课程教学的关联程度、公平性等议题,进而确保评价的有效性、可靠性与可推广性^[27]。第四,管理规范是服务于技术赋能表现性评价的系列配套支持政策、制度等举措的总称,包括评价任务的方案制定和组织管理、评价者的责任划分、评价结果的应用策略和组织机制等,可进一步确保资源的高效配置,让评价者有章可依。

(三) 优化评价过程与技术工具,推动评价的可持续发展

优化评价过程与技术工具对推动技术赋能表现性评价的稳步实施与长久发展至关重要。其主要目的是确保技术赋能表现性评价在不同教育阶段、学科、专业和应用场景中,针对各评价主体能够合理有序地开展,同时充分发挥评价的导向、诊断、调控和改进等作用。技术赋能表现性评价的目的是促进学生的学习和发展,这就意味着评价应从课程整体视角进行审视与思考,即与教学活动紧密相连,融合成一个和谐的整体,实现教学、学习和评价三者的协调配合^[28]。为此,教师与相关研究者需要着力打破评价与教学分离的局限,在“教—学—评—一致性”的背景下找准技术赋能表现性评价的角色与定位,在整体上明晰技术赋能表现性评价的功能与作用。在此基础上,围绕当前教育评价改革的需求,贯彻国家相关政策文件要求,针对最新修订的课程方案与标准、教育理念、教学设计、学习方式、评价设计与实施等领域开展试点研究与技术攻关,通过行动研究、基于设计的研究等方法,探索分析技术赋能表现性评价的实际效果与作用体现,并据此对评价过程与技术工具进行迭代优化,不断完善技术赋能表现性评价的实践框架,探究技术赋能表现性评价在不同学段、学科、课程中的有效实施策略。

(四) 建立技术赋能表现性评价运行机制,确保评价的有效运行

高效推进技术赋能表现性评价,关键在于建立完善的运行机制,保障其效率和质量。技术赋能表现性

评价不单是一种评价方式,也是我国教育评价体系的重要组成部分,可在一定程度上为教育系统内各级决策提供有效支撑,如指导课堂教学改进、为更广泛的综合素质评价及国家与区域政策制定提供参考。因此,我们应坚持理论与实践应用两手抓,以课程教学为切入点,建立“评价—反馈—改进”机制,持续提升实际应用的效果,并为技术赋能表现性评价的未来发展提供稳固支撑。首先,构建以学生发展为中心的科学的科学评价体系,充分发挥技术的优势,支持技术赋能表现性评价的设计与实践。该评价体系可涵盖评价的理论框架、目标定位、实践路径和潜在数据来源等内容,指导评价的设计与任务呈现、评价实施等。其次,加强对师生参与评价过程的有效监控,建立互动反馈机制。借助技术环境、自我报告以及课堂观察等手段实时监控学生的表现、参与程度与体验,以便评价者能够及时掌握学生的学习状态、学习动机和可能遇到的困难,并积极采取相应的措施帮助其克服困难。同时,对评价任务内容、流程及其效果进行监控,为后续评价的改进与完善提供依据^[29]。在此基础上,建立并完善学生与评价者之间的互动反馈机制,为学生提供接受外部反馈、提出疑问和需求的稳定渠道,协助评价者优化评价过程。最后,多元评价主体共同参与以推动评价的持续改进。多元评价主体通过评价结果与学生反馈等内容定期进行效果检验与评估,通过多轮评价实践与干预动态调整评价实践,涵盖目标设定、任务呈现、数据采集与分析、反馈机制、技术优化与评价实施迭代等方面,共同构建符合我国实际的、高质量的技术赋能表现性评价体系。

五、结束语

技术的快速发展为解决表现性评价所面临的诸多挑战提供了新的机遇。本研究从学生评价改革推进的现实需求出发,探讨了表现性评价的价值取向,针对表现性评价在实践中存在的学习表现数据采集尚不全面、评价过程主观性较强且耗时耗力、评价结果的应用与反馈效果欠佳等问题,构建了技术赋能表现性评价的实践框架并阐述了四大实践路径。然而,本研究仅是技术赋能表现性评价研究的起点,要真正实现科学、专业和客观的评价仍须久久为功。未来技术赋能表现性评价的研究需紧扣新时代教育评价改革路向定位,探寻与学生评价、教师评价等领域改革重点任务的契合点,充分发挥技术的优势,推动表现性评价改革创新,为深化教育评价改革和提升教育教学质量奠定坚实基础。

[参考文献]

- [1] 郑勤华,陈丽,柴唤友,等. 基于信息技术的表现性评价:内涵、作用点与发展路向[J]. 中国电化教育,2023(3):55-61.
- [2] 吴龙凯,程浩,张珊,等. 智能技术赋能教育评价的时代内涵、伦理困境及对策研究[J]. 电化教育研究,2023,44(9):19-25.
- [3] 刘邦奇,袁婷婷,纪玉超,等. 智能技术赋能教育评价:内涵、总体框架与实践路径[J]. 中国电化教育,2021(8):16-24.
- [4] 赵德成. 表现性评价:历史、实践及未来[J]. 课程·教材·教法,2013,33(2):97-103.
- [5] BLAND L M, GAREIS C R. Performance assessments: a review of definitions, quality characteristics, and outcomes associated with their use in k-12 schools[J]. *Teacher educators' journal*, 2018, 11:52-69.
- [6] POPHAM J W. 促进教学的课堂评价[M]. 促进教师发展与学生成长的评价研究项目组译. 北京:中国轻工业出版社,2003:135.
- [7] ADRI J, ABDULLAH A S. Critical thinking skills in performance-based assessment: instrument development and validation[J]. *Journal of technical education and training*, 2022, 14(1): 90-99.
- [8] 孙宏志,解月光,张于. 核心素养指向下高阶思维发展的表现性评价设计[J]. 电化教育研究,2021,42(9):91-98.
- [9] 周文叶. 表现性评价:指向深度学习[J]. 教育测量与评价,2018(7):1.
- [10] 郑勤华,陈丽,郭利明,等. 理论与技术双向驱动的学生综合素养评价新范式[J]. 中国电化教育,2022(4):56-63.
- [11] WOLFE E W. Identifying rater effects using latent trait models[J]. *Psychology science*, 2004, 46(1):35-51.
- [12] MOON J. A handbook of reflective and experiential learning: theory and practice[M]. London: Routledge,2004.
- [13] 崔允灏,夏雪梅. “教—学—评一致性”:意义与含义[J]. 中小学管理,2013(1):4-6.
- [14] MISLEVY R J, ALMOND R G, LUKAS J F. A brief introduction to evidence-centered design [R]. New Jersey: Research & Development Division Princeton,2003.
- [15] 袁建林,刘红云. 核心素养测量:理论依据与实践指向[J]. 教育研究,2017,38(7):21-28.
- [16] 罗生全,陈卓. 大数据时代教育评价的价值重构与逻辑理路[J]. 贵州师范大学学报(社会科学版),2023(4):116-128.
- [17] 朱莎,吴砥,杨浩,等. 基于ECD的学生信息素养评价研究框架[J]. 中国电化教育,2020(10):88-96.
- [18] YIN X, ZUO M, ZHANG Y, et al. Exploring the characteristics of online learning behavior in blended learning [C]//2023 International Symposium on Educational Technology (ISET). Danvers, USA: IEEE, 2023:59-63.
- [19] 谢幼如,高磊,邱艺,等. 智能技术赋能高质量课堂的评价创新[J]. 电化教育研究,2023,44(12):73-79,94.
- [20] 宛平,顾小清. 生成式人工智能支持的人机协同评价:实践模式与解释案例[J]. 现代远距离教育,2024(2):33-41.
- [21] 柴唤友,郑勤华,胡天慧,等. 基于信息技术的表现性评价:概念解析、构成要素及分类框架[J]. 中国电化教育,2024(2):1-7.
- [22] 霍力岩,黄爽. 表现性评价内涵及其相关概念辨析[J]. 西北师大学报(社会科学版),2015,52(3):76-81.
- [23] 王林慧. 表现性评价量规在低段语文非纸笔测评中的研制与应用——以杭州市天长小学低段语文非纸笔测评为例 [J]. 语文建设,2022(4):62-66.
- [24] 刘邦奇,张振超,王亚飞. 区域教育大数据发展参考框架[J]. 现代教育技术,2018,28(4):5-12.
- [25] 吕红,邱均平. 高等教育质量标准体系基本理论问题研究[J]. 重庆大学学报(社会科学版), 2015,21(5):128-133.
- [26] 吴砥,饶景阳,吴晨. 教育大数据标准体系研究[J]. 开放教育研究,2020,26(2):75-82.
- [27] 周文叶,董泽华. 表现性评价质量框架的构建与应用[J]. 课程·教材·教法,2021,41(10):120-127.
- [28] 崔允灏,雷浩. 教—学—评一致性三因素理论模型的建构[J]. 华东师范大学学报(教育科学版), 2015,33(4):15-22.
- [29] 章兢,廖湘阳. 以学生发展为导向建立高等教育质量评价与监控体系[J]. 中国高等教育,2014(1):32-34,40.

Technology-empowered Performance Assessment: Framework Design and Practice Path

ZHANG Yao¹, ZUO Mingzhang²

(1.Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan Hubei 430079;

2.Collaborative Innovation Center for Informatization and Balanced Development of K-12 Education by MOE and Hubei Province, Central China Normal University, Wuhan Hubei 430079)

[Abstract] Technology provides the conditions to promote the reform of educational assessment, and also brings unprecedented opportunities for the innovative development of performance assessment. Accordingly, the study discussed the value orientation of performance assessment, analyzed the problems of performance assessment such as incomplete data collection of learning performance, subjective and time-consuming evaluation process, and ineffective application and feedback of evaluation results in practice, and then built a practical framework of technology-empowered performance assessment, including adhering to the goal of nurturing people to ensure the consistency in evaluation design and implementation, improving the quality of data collection to guarantee the objectivity of evaluation, promoting data analysis supported by human-machine collaboration to provide developmental results and feedback, and promoting students to carry out feedback regulation to lead individual self-development. On the basis of the practical framework, the study further proposes four major practical paths for technology-empowered performance assessment, i.e., constructing a supportive environment of technology-enabled performance assessment to support the evaluation; improving the standard system framework to promote the standardization and maturity of the evaluation; optimizing the evaluation process and technical tools to promote the sustainable development of evaluation; and establishing an operational mechanism of the technology-empowered performance assessment to ensure the effective operation of the evaluation.

[Keywords] Technology Empowerment; Performance Assessment; Curriculum Instruction; Developmental Assessment; Assessment Reform

(上接第 63 页)

[Abstract] Mastering problem-solving skill in peer interactions at an early age can have a profound impact on the quality of an individual's future work and life. At a time when artificial intelligence technology continues to evolve, the effectiveness of humanoid robot-assisted education on the problem-solving skills of young children in peer interactions needs to be explored. The study selected 70 4- to 5-year-old children from kindergartens as intervention subjects, forming experimental group 1 (n=27), experimental group 2 (n=20), and control group (n=23). The educational experimental method was used to investigate the effect of humanoid robot on children's problem-solving skills in peer interactions. The results showed that the humanoid robot had a significant overall impact on children's problem-solving skills in peer interactions, with clear promotional effects on peer problem attribution, generation of diverse methods, and decision-making in behavioral strategies. But they failed to improve children's skill in predicting behavioral outcomes. The mechanism by which humanoid robots effectively enhance the problem-solving skills of young children in peer interactions was discussed. A three-dimensional guarantee of humanoid robot-assisted education was constructed, consisting of "embodied interaction", "pedagogical interaction", and "social interaction". The self-limitation of humanoid robots in intervening to improve complex social skills such as predicting behavioral outcomes was analyzed. The conclusion of the study is helpful in providing scientific guidance for the implementation of educational and teaching practices based on humanoid robots.

[Keywords] Humanoid Robot; Artificial Intelligence Education; Problem-solving Skills in Peer Interaction; Young Children's Sociality; Educational Intervention