

数智时代中文学习效果的 课堂智能测评框架构建研究

Research on the Construction of an Intelligent Classroom Assessment Framework: Effective Chinese Language Learning in the Digital Age

鲁俊锋¹

Junfeng LU

Lima Senior High School, Lima, Ohio, USA

美国俄亥俄州莱马市莱马中学

lu882007@126.com

DOI: <https://doi.org/10.5281/zenodo.22012258>

摘 要 随着数智技术与教育深度融合, 国际中文教育领域正经历测评范式的变革。课堂智能测评正成为实现国际中文教育智能化、精准化、个性化测评的关键支撑。在国际中文教学的实践中, 智能测评已呈现出多模态融合、个性化响应、即时反馈与多元协同四个维度核心特征。基于核心特征, 本研究从技术基础与核心算法、测评维度与模型构建、教学应用与干预反馈、伦理与效度及教师发展四个维度提出构建数智时代中文学习效果的课堂智能测评框架, 并在此基础上进一步探讨构建过程中所遇到的挑战 and 对策, 旨在为构建科学高效的中文教学课堂智能测评框架提供理论参考和实践指导。

关键词 数智时代; 中文学习; 国际中文教育; 学习效果; 课堂智能测评

Abstract With the deep integration of digital technology and education, the field of international Chinese education is undergoing a revolutionary transformation in its assessment paradigm. Intelligent classroom assessment is becoming a key support for achieving intelligent, precise, and personalized assessment in international Chinese education. In international Chinese teaching practice, intelligent assessment has shown four core characteristics: multi-modal integration, personalized response, immediate feedback, and multi-faceted collaboration. Based on these core characteristics, this study proposes a framework for constructing intelligent classroom assessment of Chinese learning outcomes in the digital era from four dimensions: technical foundation and core algorithms, assessment dimensions and model construction, teaching application and intervention feedback, ethics, validity and teacher development. This study further explores the challenges and countermeasures encountered during the construction process, hoping to provide theoretical reference and practical guidance for building a scientific and efficient intelligent classroom assessment system for Chinese teaching.

收稿日期: 2025-09-14; 修改日期: 2025-12-25; 录用日期: 2026-02-13

作者简介: ¹ 鲁俊锋, 美国俄亥俄州莱马市莱马中学教师。

Keywords Digital age; Chinese language Learning; International Chinese language education; Learning outcomes; Intelligent classroom assessment

引言

当前,新一轮科技革命正风起云涌,以人工智能、大数据、云技术、物联网为代表的前沿技术对人们的生产生活产生了重大影响,加快推动人类社会迈向人机协同、跨界融合、共创分享的智能时代。随着数智技术在教育领域的深度渗透,国际中文教育测评方式正加速迈向“数智化”新阶段。传统课堂测评存在“高延时、滞后性、情境脱嵌”等弊端,难以满足数智时代“大规模与个性化”与“过程和结果”并重的双重诉求(陈凯泉等,2019),国际中文学习因其跨文化、跨文字、跨模态复合特征,对测评体系的精准性、文化适应性和伦理合规性提出了更高要求。课堂智能测评应运而生,其核心是通过数据驱动的方式,实现对学生学习过程和学习效果的动态、精准、多维评估与反馈。课堂智能测评不仅关乎评价方式的技术革新,更涉及教育理念的深层变革。它强调评价的发展性功能,主张将测评融入教学过程中,为教与学的持续优化不断提供服务。开展中文学习效果的课堂智能测评研究是数智时代改变测评方式的学科之需,也是国际中文教育提质增效的内在要求。本研究系统梳理课堂智能测评的理论基础、研究现状及核心特征,提出构建中文学习效果的课堂智能测评框架,并进一步阐述在构建过程中所遇到的挑战和优化策略,以期构建一个多维度、全过程、动态化、个性化的智能测评生态系统,为国际中文教师提供精准的教学决策支持,促进国际中文教育高质量发展。

一、数智时代课堂智能测评理论基础和研究现状

(一) 理论基础

数智时代中文课堂智能测评体系的构建根植于多学科理论的交叉融合,其核心理论主要源于以下四个方面。第一、建构主义学习理论。建构主义学习理论建立在维果茨基的“社会建构理论”和皮亚杰的“自我构建理论”之上,是这两者在学习、教学领域的有机融合和运用(吕天营,2025)。该理论强调学习是基于已有经验主动建构意义的过程,而非被动接受知识。在中文课堂中,学习者的语言习得过程具有动态性、情境性和个体差异性。智能测评需突破传统标准化测试的局限,通过实时捕捉学习者在真实的交际情境中的语言输出、互动行为及认知策略运用,动态评估其意义构建能力与语言运用水平。第二、形成性评价理论。1967年美国评价学专家Scriven提出形成性评价的概念。形成性评价是通过诊断教育活动中存在的问题,为正在进行的教育活动提供信息,以提高活动质量的评价(王薇,2025)。该理论主张测评与教学过程的深度融合,强调通过持续、即时的反馈促进教与学的双向调整,数智技术为形成性评价提供强大的工具支持,使“测评即学习”成为可能。课堂智能测评系统能够实时采集课堂多模态数据,例如语音、表情、操作行为等,通过算法模型即时生成可视化分析报告,帮助教师动态

调整教学策略，同时为学习者提供个性化学习建议。第三，认知诊断理论。认知诊断理论是认知心理学与心理测量学深度融合的产物，其核心在于通过建模个体内部认知结构与加工过程，精准识别学习者知识状态、技能掌握水平及潜在认知缺陷。与传统测验仅关注总分排名不同，该理论将学习效果细化为特定认知属性（如中文词汇、语法规则、语篇理解等）的掌握程度，并借助数学模型生成学习者在各属性上的诊断报告，为个性化教学提供科学依据。第四、智能教育理论。智能教育理论以人工智能技术为核心驱动力，通过自适应学习系统、智能辅助等技术载体，实现教育供给的个性化和智能化。其核心技术支撑包括机器学习、自然语言处理、计算机视觉等，以“技术赋能的教育公平”理念为基础。在中文学习的课堂智能测评中，智能教育理论的应用呈现出多技术融合的特征。智能教育理论推动课堂测评从“人工主观评价”转向“人机协同评价”，既保障教师对跨文化语境理解等复杂要素的质性评估，又通过技术手段实现基础指标的客观量化。

（二）研究现状

数智化背景下的国际中文教育深度融合了教学与技术，体现了智慧教育具有的技术与教育双重属性（马瑞凌等，2024），是未来国际中文教育的典型业态和发展目标。相比传统教育，智慧教育能带来更加简单、高效、智能和个性化的教学体验，是教育信息化发展的必然阶段，备受研究者关注。近年来学者对智慧课堂中智能测评的相关研究呈现出多层次、多视角特征，笔者发现主要有以下四个方面的研究路径。

第一、在技术基础和核心算法层方面，随着技术和算法的持续更新迭代，近年来研究者们不断结合新技术来辅助实现课堂智能测评。借助人工智能和大数据技术，可实现智能化学习诊断、动态化资源适配和“管家式”学习支持服务三大功能的智能化过程性评价（姜丽萍、王立，2021）。利用数字技术和人工智能技术可解决国际中文教育中本土化、多元化、均衡化实施的难题（刘玉屏、欧志刚，2022）。借助计算机视觉技术可以检测视频帧图像人脸数目、轮廓特征、主体动作幅度等数据，使用特殊模型推理主体行为特征，实现对课堂教学视频中行为的智能识别（周鹏霄等，2018）。现研究还包括结合人工智能和机器学习技术对师生行为进行过程性评价（魏晖，2025），借助人工智能监控系统（Parambil et al., 2022）、神经网络识别系统（Mery et al., 2019）、传感器技术（Abbasy & Enrique, 2017）、物联网技术（Abdel-Basset et al., 2018）、机器人技术（Jones & Castellano, 2018）、虚拟现实技术（Górski et al., 2016）和混合现实技术（Kasapakis et al., 2024）等技术，以实现智能检测、智能识别、智能记录学生课堂上各种学习影响因素变化，为课堂提质增效。

第二、在测评纬度和模型构建方面，课堂智能测评的相关研究目前包含多个核心维度。张鹏（2025）提出基于智慧化的三维立体评价体系，从学生的生物特征状态评价、知识获取度评价、自评与互评参与度评价三个方面评价学生的课堂学习效果。其次，测评维度已从三维立体

评价拓展至对学生学习态度的多维度感知。例如,多模态学习分析技术的应用,使得对学生行为、认知与情感三个维度的投入度进行综合分析成为可能(侯成坤,2020)。此外,评价维度还涵盖语言能力、互动交际能力、情感与参与度等,并涉及自适应测评等研究方向。在模型构建方面,已有成果包括智慧课堂的教学结构理论模型(孙曙辉、刘邦奇,2015)、国际中文智慧教学模式的过程模型(马瑞凌等,2022)、智慧课堂数据挖掘分析框架与应用模型(孙曙辉等,2018)、智慧教学情景感知模型(王东青等,2018)以及课堂教学评价模型(邓伟等,2023)。

第三、在教学应用与反馈干预方面。在教学应用方面,目前课堂智能测评系统已深度融入国际中文教学课前、课中和课后全流程。在课堂教学环节中,研究表明利用数智技术可实现精准施教(刘邦奇,2023),并能促进课堂教学语言的智能化评价(冯秀梅,2025)。针对具体的中文教学要素,该技术可覆盖语音、汉字、语法、词汇、听力、口语、阅读、写作等十个方面,实现全方位智慧化教学,从而显著提升数智技术对国际中文教学的应用价值(马瑞凌,徐娟,2024)。基于汉语二语文本可读性的特征集合,借助特征算法,杜月明等(2022)实现了汉语阅读文本可读性的自动评估。借助大语言模型技术,国际中文教师可搭建多种虚拟场景,促进语言学习者交互性和沉浸感(郑艳群,2025)。薛嗣媛等(2023)详细介绍了写作智能测评研究的数智技术方法,阐述了写作智能测评研究的任务模式和评估方式。在反馈干预方面,其核心是通过数据驱动实现精准教学决策,例如,教师在课堂智慧测评中可借助数智技术纠正语音、汉字书写和开展针对性语法和表达干预等。

第四、在伦理与教师专业发展方面。ChatGPT为教学带来了便利,但值得深思其在教育领域引起的伦理、隐私保护和数字鸿沟问题(周洪宇、李宇阳,2023)。袁羲(2023)提出人工智能技术在提供智能化服务时,数据隐私、偏见性和所有权三种伦理问题值得重视。尚智和和闫禹宏(2023)将伦理问题主要归因于教师主体性丧失和情感交流及道德的淡漠。朱光辉、王喜文(2023)提出,可通过法律法规应对科技创新引发的伦理问题。在教师专业发展方面,李栋(2020)指出,大数据与人工智能等新技术的应用推动教师教育进入数据驱动、精准化、个性化的“智能研修”新阶段。

当前,关于数智时代中文学习效果的课堂智能测评研究在测评纬度的多元化和模型构建的技术创新方面取得了一定研究进展,促进了从经验主义到向数据驱动与经验融合的范式转变,但还存在很多研究不足,例如多模态数据融合的精度和评价模型的场景适应性问题,以及如何平衡技术赋能与教育伦理、效度的关系。针对以上研究不足,综合多维度研究困境,开展中文学习效果的课堂智能测评框架研究具有重要的理论意义和现实价值。

二、数智时代课堂智能测评核心特征

随着数智技术的飞速发展，以人工智能为代表的数智技术为中文课堂测评带来了颠覆性变革。中文课堂作为学生学习中文的主要场所，其智能测评呈现出诸多新型特征，其多模态融合、个性化响应、即时反馈与多元协同等特征，正重塑国际中文教育的评价生态。

（一）多模态数据融合的动态化评价

国际中文学习涉及听、说、读、写、译等多种技能，单一模态测评难以全面反映学习的效果。多模态智能测评系统突破传统纸笔或口头测试局限，综合运用语音识别、图像处理、自然语言处理、情感计算、计算机视觉等技术，实现口语语音流、汉字书写轨迹、课堂互动、微表情姿势、情感状态、小组讨论度等多模态数据的同步采集与分析，从而多维度综合评估学习者能力。例如，语音识别可分析发音准确性与语调自然度，自然语言处理可评估作文语法、词汇与篇章连贯性。系统还能实时记录多维数据，构建动态学习画像，实现动态化评价，显著提升测评效果。

（二）数据和智能双驱动的个性化测评

借助人工智能与大数据分析技术，智能测评系统能够识别学习者个体差异与学习轨迹，进而为每位学习者生成个性化测评报告与学习建议。通过分析学生作业提交模式、错误类型分布及知识掌握速度等数据，系统可自动输出差异化学习诊断报告；同时，通过采集与分析学习者的行为数据、认知数据及情感数据，智能测评技术能够构建动态更新的学习者知识图谱。以北京语言大学“国际中文智慧教学系统”为例，该系统利用大语言模型，实时解析学生口语表达中的语法错误与发音缺陷，生成涵盖“语音、词汇与句法”三个维度的能力雷达图，精准定位个体薄弱环节。

（三）智能评价反馈的即时性与生成性

传统测评存在结果滞后与反馈延迟等问题，智能测评系统通过算法模型可实现实时反馈。在课堂教学过程中，智能测评系统实时采集学习者的学习数据，如课堂回答问题的情况、参与小组讨论的表现、在线测试的答题情况等，对这些数据进行即时分析和处理，生成实时测评结果。教师能够根据实时测评结果，及时了解学习者对中文知识的掌握程度和学习过程中存在的问题，进而调整教学策略和教学进度，为学习者提供及时的指导和反馈。同时，这种反馈提供的生成性学习资源，如定制化微课、强化练习题，也便利了学习者的自主修正与能力提升。

（四）评价主体的多元化协同

在传统课堂中，教师评价主要是单一教师主体。随着数智技术的快速发展，数智技术推动了课堂评价主体由教师单一主导向“人机协同”的范式转变。在测评环节中，智能系统承担了数据采集、初步分析与量化评价等任务，例如系统量化统计学生课堂发言次数、互动质量等数据，教师聚焦于质性解读、情感关怀与教学决策。教师结合专用经验分析其语言运用的得体性

与文化适应性，形成算法诊断与教师智慧互补的协同评价模式。这种人机协同的分工合作既提升了评价效率，又保障了教育的人文温度。

三、数智时代中文学习效果的课堂智能测评框架

在数智技术深度赋能教育领域的背景下，国际中文教育正从经验驱动向数据驱动。课堂智能测评体系作为教学闭环的核心环节，亟需突破传统测评的滞后性与单一性，构建覆盖基于“认知-情感-行为”多维度和“诊断-反馈-干预”全流程的智能动态评价框架。基于数智时代背景，构建数智时代中文学习效果的课堂智能测评框架，需要从技术基础与核心算法、测评维度与模型构建、教学应用与反馈干预、伦理与效度及教师发展四大框架做起。

（一）技术基础与核心算法：多模态数据融合与智能分析

数智时代背景下，课堂智能测评框架的构建依赖坚实的技术基础与精准的核心算法。在多模态数据采集和融合处理方面。当前，先进的课堂智能评测系统能够捕获和整合语音、文本、视频、眼动、脑电等多源异构数据，实时掌握学生学习过程中的语言输出、表情变化、互动行为等原始数据。借助语音识别技术可实时转录学生口语表达；结合自然语言处理技术可深度分析学生的中文输出，包括作文中的语法准确性、词汇丰富度、语义连贯性、以及口语表达中的发音准确度、语流流畅度；借助计算机视觉技术可捕捉学生的课堂行为特征，如举手次数、小组讨论参与度、课堂面部表情与肢体动作，辅助判断学生学习投入度；借助眼动仪可追踪学生中文阅读轨迹，诊断中文词汇识别效率。

在核心算法与模型架构方面，课堂智慧测评系统已集成多项核心算法及大模型。机器学习算法（包括监督学习、无监督学习及半监督学习等）主要用于对学生学习行为、成绩及情感状态进行分类、预测与聚类分析；深度学习算法（包括卷积神经网络、循环神经网络等）则用于处理复杂的非线性关系与时序数据，以提升模型准确性。自适应学习算法是另一项关键技术，其基于行为模型生成个性化学习规划，同时可感知学生状态并给予适配回应，配合多模态输入交互，共同建构智能学习场景。在模型架构层面，系统依托通义大模型的语义分析能力，可拆解作文的立意深度、逻辑结构及文化隐喻等维度，并据此生成个性化评语。

（二）测评维度与模型构建：从单一语言能力到综合素养评价

在构建中文课堂智能测评框架时，需确立科学全面的测评维度与模型。应突破传统“听、说、读、写”的局限，建立以语言知识、语言能力、学习策略、思维品质和文化认同为核心的新型测评体系。其中，语言知识涵盖语音、词汇、语法及表达理解；语言能力包括听说写译、语篇连贯、跨文化交际与文化理解；学习策略涉及元认知策略（如自我计划、监控与反思）及资源利用策略（如工具使用与同伴互助）；思维品质聚焦逻辑、批判与创造性思维；文化认同则考察学生对文化符号的理解、价值观接纳及跨文化适应力。

在模型构建方面,动态认知智能诊断模型是精准评估的关键。课堂智能测评框架构建采用动态的、多层次、多指标的综合评价方法。首先,基于上述不同测评维度,构建一个综合评价模型,以实现对学生的学习效果进行全面、客观的评估。其次,由于学生的个体差异和学习情况动态性,应构建动态适应性评价模型,根据每个学生的差异,动态调整评价指标和标准,实现个性化评价。最后,应构建预测性评价模型,利用教育历史大数据和机器学习算法,预测学生未来的学习表现和潜在问题,为提前干预提供依据。

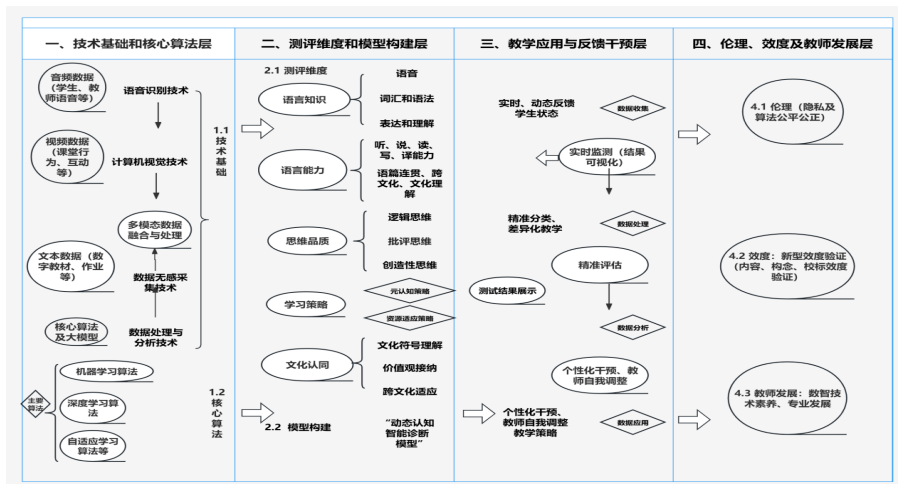
(三) 教学应用与反馈干预:从数据洞察到教学优化

智能测评系统的最终价值体现在教学应用中,主要包括实时监测、精准评估与个性化干预三大方面,同时促进教师教学自我调整。在实际教学场景中,系统可实时反馈学生学习状态,例如当多数学生对某语法点理解困难时,教师可即时调整节奏,补充讲解或组织互动练习;同时依托聚类算法进行精准评估,将学生划分为不同水平群体,便于教师设计差异化练习。在个性化干预方面,系统通过构建多主体、多层次的反馈干预体系,实现数据采集、处理、分析与可视化呈现,直观展示学生优劣势并提供改进建议。智能测评与课堂教学的深度融合,强化个性化干预效果,也辅助教师动态优化教学策略,从而推动“测评—反馈—干预—提升”良性循环的持续运行。

(四) 伦理与效度及教师发展:促进保障体系可持续性

构建课堂智能测评体系需重视伦理问题、效度保证和教师发展是不可忽视的重要方面。伦理层面,系统涉及多模态数据收集与分析,必须严格保护学生数据隐私,确保个人信息安全,同时坚持算法公平公正,避免评价偏见,并保障评价机制透明可解释;效度层面,需建立科学验证机制,采用多维多元策略,涵盖内容效度(评估内容代表中文学习构念)、构念效度(测评工具测量认知与情感等理论构念)及效标效度(测评结果预测未来学习表现与实际能力)三重验证;教师发展层面,则需着力提升教师数字素养,帮助其熟练运用智能测评技术以优化教学实践,从而确保体系的整体有效性与可持续性。

图1 国际中文学习效果的课堂智能测评框架



由图1可知,国际中文学习效果的课堂智能测评框架通过多模态数据与智能算法,实现对学习者认知、情感、行为的多维动态评估,为教学干预提供实时的、个性化的科学依据,从而提升国际中文教学质量、促进学习效率,并保障教育评价的准确性、全面性与公平性。

四、中文学习效果的课堂智能测评核心挑战与对策

当前,数智技术正日益深度融入国际中文教育,课堂智能测评作为提升教学精准度的核心环节,其构建过程面临技术适配、教学协同、伦理规范、效度保障等四个维度挑战,亟需系统性对策予以破解。

(一) 构建中文学习效果的课堂智能测评体系核心挑战

1. 技术适配性挑战:中文特性与多模态数据的耦合难题

首先,在中文特性方面,作为表意文字的中文学习涉及语音、语义、文化理解等多模态信息,当前技术在自然课堂环境中难以高效捕获全维度数据。例如,汉语口语表达的音韵特征、跨文化交际中非语言行为等关键指标,常因数智设备局限或技术算法缺陷导致缺失。其次,关于多模态数据方面,在课堂场景中,基于测评基础的语音、文本、视频、眼动等多模态数据在时空方面和语义上是关联的,然而现有算法在处理非结构化数据时存在显著局限,且多模态数据融合难度大,例如无法将“皱眉表情”与“语法错误率”有效关联以分析学生困惑点。

2. 教学协同性挑战:技术工具与教学实践脱节

首先,智能测评系统与国际中文教育教学场景协同性不足,主要体现在三个方面。一是智能测评工具多聚焦测评语言能力,而忽略文化素养与跨文化交际能力的测评。二是教师使用智慧测评系统消解了教师的主体性,基于量化指标的课堂智能测评评价削弱了教师对非智力因素的洞察力,例如使用智能测评系统的老师识别课堂中学生文化认知冲突识别率下降。同时,技术依赖还导致教师灵活调整评价策略的能力下降,教师沦为技术附庸抑制了教学情境智慧的动态决策。三是教师技术应用能力不足和智能测评技术在中文教学场景中的应用存在明显的局限性。国际中文教师在使用智慧测评系统工具时多数仅仅存在能完成测评系统的基础操作,智慧测评系统本身往往过度依赖数据、缺乏过程性解释。

3. 效度保障及教师专业发展挑战:动态场景下的测评结果可靠性

中文学习的复杂性和教学场景差异正深刻影响着课堂智能测评效度,导致两重困境。一是动态测评的实时性与准确性失衡。课堂智能测评系统因为数据采集与处理的复杂性、算法模型的局限性、效度与信度的保障、教师适应性问题等因素促使课堂动态测评的实时性与准确度失衡。二是效标关联效度偏低。课堂智能测评系统存在过于关注技术指标、过分强调数据化、标准化时,由于测评效度验证机制缺失,会导致学生真实的中文语言能力和测评结果出现很大偏差。同时,也存在教师数智素养水平不一,影响了智能测评测评效果和质量。

4. 伦理风险与公平挑战：跨文化场景下的隐私和公平性问题

课堂智能测评在数据获取和算法设计层面存在伦理陷阱。首先，在数据隐私保护方面，课堂智能测评系统涉及采集学生隐私、敏感数据，例如人脸照片、声纹特征、面部表情、操作记录、课堂行为等，其数据采集、存储和使用面临着法律风险，现在欧盟已将课堂监控视频存储纳入特殊保护范畴，同时人工智能技术在国际中文教育中的普遍应用导致缺乏完善的伦理审查和监管机制。汉语二语学习者涉及多国文化背景，数据伦理问题更为复杂。其次，在算法设计方面，测评模型的训练数据存在算法偏见与文化误判。课堂训练数据的文化单一性导致评估误差，文化维度评估更存在深层偏见，如在海外中文教学中将当地学生谦辞使用频率低简单归因为礼貌意识不足，忽略西方学生直率表达的文化合理性。

（二）构建中文学习效果的课堂智能测评框架应对策略

1. 技术适配优化：构建中文适配的多模态测评架构

数智时代的课堂智能测评高度依赖人工智能、大数据、语音识别及自然语言处理等前沿技术，然而这些技术在中文学习效果测评中的适配性仍存不足。为降低课堂智能测评系统对数据的过度依赖并规避失真风险，需充分观照中文的独特语言属性，如汉字结构的复杂性、声调系统的差异性、语法组织的灵活性，从而针对性地攻克技术处理中的各项难点，构建与中文特征相匹配的多维度数据驱动生态。构建多模态感知技术体系、扩展数据采集纬度、优化感知技术和情境感知计算。在智能测评实践中可通过部署智能化语音终端、计算机视觉技术及可穿戴设备，形成“声、文、像、情”四维数据采集网络，优化语音识别技术，提高处理音近字，复杂声调变化，复杂方言背景等难题，减少自然语言处理技术分析中文文本的语义，语境和情感时因中文词语多义性，歧义导致的偏差。为提高中文学习课堂智能测评的适配性，首先应加强针对中文独特语言特点的技术研发，着力攻克中文语音识别、自然语言处理中的关键技术瓶颈，有效应对汉字复杂性、声调变化及语义歧义等问题，从而提升系统的准确性与稳定性。其次，需充分考量学习者的文化背景差异，依据不同群体的认知与语言习惯，优化测评内容与应用场景，确保测评过程的公平性与结果的有效性。

2. 强化教学协同：推动技术工具与教学实践深度融合

为实现智能测评与教学的有效协同，需多管齐下，共建“人机协同”智能测评新生态。第一，重构智能测评内容体系。课堂智能测评体系需涵盖语言能力、学习过程、文化素养三大维度。在语言能力纬度，聚集听、说、读、写四种能力。在学习过程检测方面，聚集学生的学习行为与策略，包括参与度，学习策略，情感态度。在文化素养方面，注重检测学生的跨文化交际能力，文化知识掌握度。第二，强化教师在智慧测评系统中的主体性，正确认识数智技术在课堂测评中的作用，教师在测评环节中应发挥主体作用，减少对科技的依赖，灵活调整评价策略，基于教学情景智慧动态决策。第三，国际中文教师应提升教师技术素养，理性认识智慧测

评系统在中文教学场景中的局限性，倡导“人本智能”理念，量化人机协作的最佳平衡点，构建兼具智能精度与人文温度的国际中文教育新测评生态。

3. 效度及专业发展保障：确保动态场景下的测评结果可靠性

在效度方面，建立科学的效度验证机制对智能测评系统至关重要。多维度的测评效度验证包括内容效度验证、结构效度验证、预测效度验证以及收敛效度验证。首先、在内容效度验证方面，要确保评价内容的科学性和合理性；其次、在结构效度和预测效度验证方面，智能测评结果要能有效预测学生未来学习表现和实际能力，智能测评要能准确测量评估构念，例如认知水平、情感状态、社交能力等。再次，在教师发展方面，应提高教师的数据素养和技术应用能力，熟练掌握智慧测评系统的使用方法，提高智能测评效果和质量。

4. 走出伦理与公平困境：保障跨文化场景下的隐私与公平

在构建和应用中文学习效果课堂智能测评体系的过程中，伦理问题是不可忽视的。教师在采集和使用学生数据时，需确保数据使用的透明性和安全性，注重保护学生隐私，确保学生个人信息安全。同时，算法公平性也是重要伦理考量，应避免算法偏见对特定群体学生的不公平评价，减少对于学习者的学习数据和量化结果依赖，多关注学习者的心理状态，情感需求等质性因素，增加对学习者的关怀，完善伦理规范。

结语

数智技术为国际中文教育课堂测评带来革命性变化，推动国际中文教育从高速发展阶段迈向高质量发展。本文以课堂智能测评核心特点为基础，对课堂智能测评框架、现存挑战和发展策略进行了初步探索。目前，在数智测评框架构建中还存在精准、个性、协同、优化、创造和思维等方面值得思考的问题。未来研究需立足中文语言的特殊性，推动课堂智能测评从“技术赋能”向“生态重构”的跨越，加强技术、教育与伦理的协同推进，以技术融合为基，以伦理规范为界、以教师发展为要，构建新型中文学习效果的课堂智能测评体系，真正实现数智技术赋能下的教育公平与质量提升。

参考文献

- Abbasy, M. B., & Enrique V. Q. (2017). Predictable influence of IoT (Internet of Things) in the higher education. *International Journal of Information and Education Technology*, 7(12), 914–920. <https://doi.org/10.18178/ijiet.2017.7.12.995>
- Abdel-Basset, M., Manogaran, G., Mohamed, M., & Rushdy, E. (2019). Internet of things in smart education environment: Supportive framework in the decision-making process. *Concurrency and Computation Practice and Experience* 31, e4515. <https://doi.org/10.1002/cpe.4515>
- Górski, F., Buń, P., Wichniarek, R., Zawadzki, P., & Hamrol, A. (2017). Effective design of educational virtual reality applications for medicine using knowledge-engineering techniques. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(2), 395–416. <https://doi.org/10.12973/eurasia.2017.00623a>

- Jones, A., & Castellano, G. (2018). Adaptive robotic tutors that support self-regulated learning: A longer-term investigation with primary school children. *International Journal of Social Robotics*, 10(3), 1–14. <https://doi.org/10.1007/s12369-017-0458-z>
- Kasapakis, V., Gavalas, D., Dzardanova, E. (2024). Mixed Reality. In: Lee, N. (eds) *Encyclopedia of Computer Graphics and Games*. Springer, Cham. https://doi.org/10.1007/978-3-031-23161-2_205
- Mery, D., Mackenney, I., & Villalobos, E. (2019). Student attendance system in crowded classrooms using a smartphone camera. In 2019 IEEE *Winter Conference on Applications of Computer Vision (WACV)* (pp. 857–866). IEEE. <https://doi.org/10.1109/WACV.2019.00096>
- Parambil, M. M. A., Ali, L., Alnajjar, F., & Gochoo, M. (2022). Smart Classroom: A deep learning approach towards attention assessment through class behavior detection. *Advances in Science and Engineering Technology International Conferences*. 1–6. <https://doi.org/10.1109/ASET53988.2022.9735018>
- 陈凯泉,张春雪,吴玥玥,刘璐. (2019). 教育人工智能(EAI)中的多模态学习分析、适应性反馈及人机协同. *远程教育杂志*, 37(05), 24–34. <https://doi.org/10.15881/j.cnki.cn33-1304/g4.2019.05.004>.
- 邓伟,杨晓丹,高倩倩. (2023). 人工智能支持下的课堂教学评价模型研究. *中国教育信息化*, 29(08), 3–14. <https://doi.org/10.3969/j.issn.1673-8454.2023.08.001>
- 杜月明,王亚敏,王蕾. (2022). 汉语水平考试(HSK)阅读文本可读性自动评估研究. *语言文字应用*, (03), 73–86. <https://doi.org/10.16499/j.cnki.1003-5397.2022.03.012>.
- 冯秀梅,蒋雨辰,王依婷. (2025). 课堂教学语言智能化评价:效用,应用与优化. *教师教育学报*, 12(4), 51–61. <https://doi.org/10.13718/j.cnki.jsjy.2025.04.006>.
- 姜丽萍,王立. (2021). 智慧教育视域下中文学习平台的构建:特征,功能与实现路径. *国际中文教育(中英文)*, 6(04), 91–99.
- 李栋. (2020). 人工智能时代的教师发展:特质定位与行动哲学. *电化教育研究*, 41(12), 5–11. <https://doi.org/10.13811/j.cnki.eer.2020.12.001>.
- 侯成坤. (2020). 基于多模态融合的课堂教师教学行为自动识别研究(硕士学位论文). 华中师范大学. <https://doi.org/10.27159/d.cnki.ghzsu.2020.001237>.
- 刘玉屏,欧志刚. (2022). 本土化,多元化,均衡化:人工智能在国际中文教育中的应用探析. *民族教育研究*, 33(1), 162–169. <https://doi.org/10.15946/j.cnki.1001-7178.2022.01.020>.
- 刘邦奇. (2023). 智慧课堂引领教学数字化转型:趋势、特征与实践策略. *电化教育研究*, 44(8), 71–79. <https://doi.org/10.13811/j.cnki.eer.2023.08.010>.
- 吕天营,俞金波. (2025). 基于建构主义学习理论的“情境-行动”式创业课程体系变革. *宁波教育学院学报*, 27(04), 73–77.
- 马瑞陵,蔡建永,曹钢,于淼,徐娟. (2024). 基于实践的国际中文智慧教学认识深化. *天津师范大学学报(社会科学版)*, (6), 9–16. <https://doi.org/10.20247/j.issn1671-1106.2024.06.002>.
- 马瑞陵,曹钢,徐娟. (2022). 国际中文智慧教学模式的过程模型与课堂活动设计. *数字化国际中文教育*. 70–77.
- 马瑞陵,徐娟. (2024). 国际中文写作智慧教学模式构建与应用探索. *语言教学与研究*, (02), 12–23. <https://doi.org/10.3969/j.issn.0257-9448.2024.02.002>
- 孙曙辉,刘邦奇. (2015). 基于动态学习数据分析的智慧课堂模式. *中国教育信息化*, (22), 21–24.
- 王薇. (2025). 基于形成性评价理论的超常儿童评价框架与策略研究. *中国特殊教育*, (02), 56–66. <https://doi.org/10.3969/j.issn.1007-3728.2025.02.007>
- 薛嗣媛,周建设,任福继. (2023). 写作智能评测研究综述和发展前景. *中文信息学报*, 37(02), 1–14. <https://doi.org/10.3969/j.issn.1003-0077.2023.02.001>

- 袁羲,吴应辉. (2023). ChatGPT Plus给国际中文教育带来的机遇、风险及应对策略. *云南师范大学学报(对外汉语教学与研究版)*, 21(3), 53-62. <https://doi.org/10.16802/j.cnki.ynsddw.2023.03.011>
- 张鹏,董宇轩. (2025). 智慧化三维立体评价体系在课堂教学效果评价中的应用研究. *科教文汇*, (6), 36-40. <https://doi.org/10.16871/j.cnki.kjwh.2025.06.008>
- 郑艳群.(2025).迈向新生态:大语言模型背景下国际中文教育的跨学科变革. *天津师范大学学报(社会科学版)*, (3), 31-40. <https://doi.org/10.20247/j.issn1671-1106.2025.03.003>
- 周洪宇,李宇阳. (2023). ChatGPT对教育生态的冲击及应对策略. *新疆师范大学学报(哲学社会科学版)*, 44(4), 102-112. <https://doi.org/10.14100/j.cnki.65-1039/g4.20230224.001>
- 周鹏霄,邓伟,郭培育, & 刘清堂. (2018). 课堂教学视频中的S-T行为智能识别研究. *现代教育技术*, 28(6), 54-59. <http://doi.org/10.3969/j.issn.1009-8097.2018.06.008>
- 朱光辉,王喜文. (2023). ChatGPT的运行模式、关键技术及未来图景. *新疆师范大学学报(哲学社会科学版)*, 44(4), 113-122. <https://doi.org/10.14100/j.cnki.65-1039/g4.20230217.001>

(责任编辑:张鹭遥)