

太极数字人教师赋能国际中文教育的实证研究

孙雨彤 朱勇 *

摘要

人工智能技术迭代下，数字人教师为国际中文教育发展带来新机遇，但仍缺乏由客观教学效果衍生的理论机制与实践路径。太极融身体动作与文化词汇于一体，是探索数字人教师设计方案的综合教学范例。本研究因而以此为载体，结合具身认知理论与认知负荷理论构建框架，对 482 名新西兰初级汉语学习者开展实证研究。研究发现：身体、环境与认知的协同交互可动态调节学习者的内外认知负荷与关联认知负荷。基于此机制构建的开发配置“外形 - 动画 - 语音”三合、语言习得“先语后同 - 先母后中”二分、动作习得“慢速 - 分解兼重复 - 多元指令 - 视听强调”四联路径，可显著促进动作当即掌握与语言快速记忆。该路径不光能减轻教师演示负担、提高教学效率，还可扩展至更多文化、语言教学情境，为国际中文教育数智化变革提供更多依据。

关键词：数字人教师 国际中文教育 具身认知理论 认知负荷理论

孙雨彤，华东师范大学国际汉语文化学院，联络电邮：51265400038@stu.ecnu.edu.cn

* 朱勇，北京外国语大学中国语言文学学院教授，联络电邮：zhuyong@bfsu.edu.cn（本文通讯作者）

感谢审稿专家对论文提出的宝贵建议。

一、引言

近年来,人工智能技术随着时代的演进持续迭代发展。数字人教师(Digital Human Teacher)作为AI在虚拟现实层面的具象化呈现,已在多领域广泛应用,其教育性、交互性特点更契合汉语二语教学的情境与交互需求(王春辉,2024)。《教育强国建设规划纲要(2024—2035年)》提出“推动语言文字与信息技术深度融合”,《关于加强数字中文建设的意见》亦明确指出应以智能技术赋能“中文+文化”传播,为数字人教师的融入提供政策指引。

现有研究虽然表明数字人教师可构建语言环境、规划学习路径(魏小东等,2024),但是仍存在显著不足。第一,数字人教师赋能语言与动作习得的内在机制尚未得到系统揭示。少量研究虽已印证数字人教师具备教学吸引力(周世豪等,2025),其外形、语音可促进知识保持与迁移(吴长城等,2024)、降低无关认知负荷(Zhang et al., 2025),但仍以探讨主观感受与分析少量技术要素为主,缺乏多要素结合的客观效果分析(Lin et al., 2024),且动作与语言双习得机制研究仍属空白。而国际中文教育许多文化教学如太极、书法等,本质上都是身体动作与语言词汇的一体化任务,揭示其习得机制,才能解释并优化数字人赋能汉语教学的具体过程。第二,数字人教师的应用较为分散,面向日常教学的数智化路径尚未形成。部分研究虽已总结标准体系(李青等,2024)、实践路径(郭亚军等,2024)、技术架构(余超凡等,2025),但多数比较宏观、难以落地,国际中文教育领域可操作的实践路径同样亟待探索。只有形成一条直接面向日常课堂的数字人设计路径,才能为一线汉语教学提供可复制方案。

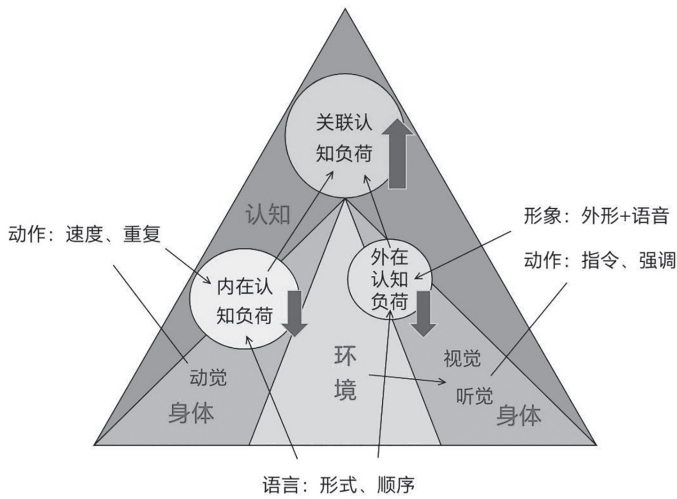
为弥补上述研究缺口,搭建能够有效管理学习者认知资源,并能深刻揭示学习过程中身体、环境与认知互动关系的理论框架十分必要。认知负荷理论与具身认知理论的融合为此提供了理想视角。认知负荷理论通过管理工作记忆资源优化学习效果,由内在、外在、关联¹三类认知负荷构成,其核心目标是将无效认知负荷转化为有效认

1 内在认知负荷是指由学习材料的难度水平带来的负荷。外在认知负荷是由与学习过程无关的活动引起的,不是学习者建构图式所必须的,因而又称无效认知负荷。关联认知负荷是指有利于图式建构的认知负荷(庞维国,2011)。

知负荷 (Paas et al., 2020)。具身认知理论包含环境、身体、认知三大核心要素，认为认知植根于身体感知系统与环境的实时互动 (张旭，2024；罗雪英等，2024)。现有数字人研究已开始尝试应用上述理论，吴长城等 (2024)、Zhang et al. (2025) 运用认知负荷理论对比不同数字人的教学效果，余越凡等 (2025) 运用具身认知理论强调学习者与数字人的积极意向性互动，但均未将两种理论融合。邓国民等 (2019)、张宁等 (2021) 以两种理论为基础探索多媒体学习；Zou et al. (2025) 结合二者生成“关联—整合”模型，指出身体动作与教学内容的高关联、高整合可有效减少认知负荷。然而，这些研究缺乏对教学效果的客观检验，也尚未将两种理论的核心要素进行系统性整合，导致对不同类别认知负荷的动态变化、人机交互的感知过程分析不足。

鉴于此，在既有研究基础上，将认知负荷理论与具身认知理论中的三个要素进行深度整合，构建“具身—负荷”整合框架 (见图 1)。

图 1 “具身—负荷”整合框架



该框架基于认知负荷理论与具身认知理论的深度融合，明确三类认知负荷与具身交互三要素的对应关系：环境要素如外形、语音、指令、强调等作用于外在认知负荷，身体要素如速度、重复等作用

于内在认知负荷，认知要素如图式建构、语义表征作用于关联认知负荷，三者协同决定认知资源分配效率。需指出的是，优化“速度”“形象”等要素可降低相应内外认知负荷已得到前人验证（Sweller et al., 1998；Mayer, 2005），本文则主要聚焦两类负荷的转化效率差异，以及认知负荷与身体感知系统的交互。具体研究路径如下：

内在认知负荷转化路径。本质是“身体—认知”的协同匹配。针对学习材料本身的复杂度（Sweller et al., 2019），通过调节身体、认知要素，适配学习者动觉感知与认知加工能力，使学习内容处于“最近发展区”，推动动作图式建构和语义表征建立。

外在认知负荷转化路径。本质是“环境—身体”的感知适配。针对无关信息的呈现方式（Sweller et al., 2019），通过优化环境要素，匹配学习者的视听与动觉感知系统，减少环境与身体互动中的无效认知消耗，释放工作记忆空间，为认知层面的深度加工提供资源保障。

基于上述理论框架，本研究选取“太极”这一典型的“动作—语言”整合任务作为教学切入点，对新西兰初级汉语水平学习者展开实证教学，通过多维度分析形象、动作、语言测试结果探索以下问题：

1. 习得机制上，学习者在数字人教师引导下的内在认知负荷与外在认知负荷的转化效率有什么差异？其与身体感知系统又是如何具体交互的？
2. 教学实践上，数字人教师赋能国际中文教育的路径应如何设计？

二、研究设计

（一）实验对象

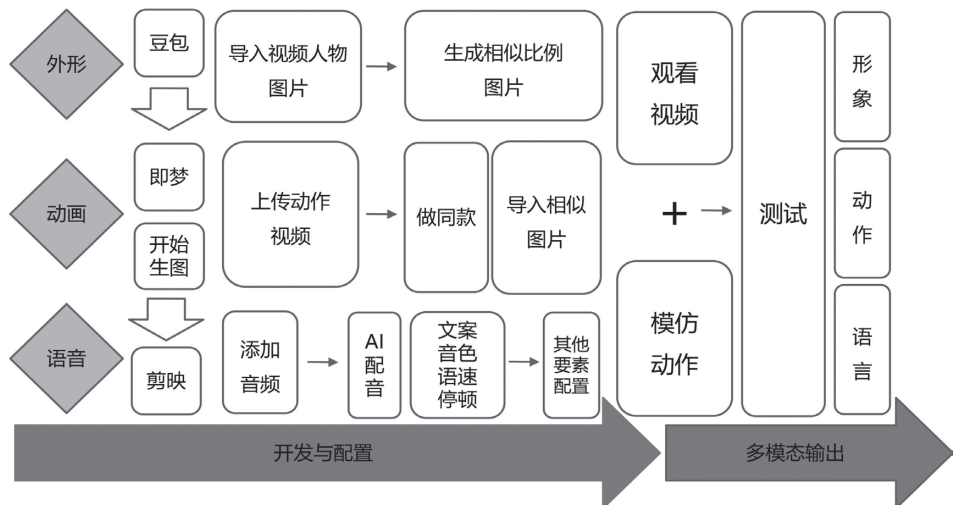
本研究采用整群随机抽样方法，招募482名新西兰某小学汉语学习者，女生243名，男生239名，年龄7-10岁（ $M=9.1$, $SD=0.8$ ），新HSK初级水平（1-2级）。研究获得学校、学生及家长同意并通过伦理审查。实验前通过问卷调查确认所有被试太极动作及相关词汇均为

零基础。被试随机分配至各对照组，独立样本 t 检验显示各组前测基础无显著差异 ($p>0.05$)。此次实验选择新西兰初级学习者作为研究对象的原因有二：其一，研究者于当地孔子学院的实践，可为研究提供可靠的实验环境与生态效度；其二，新西兰作为典型英语文化圈，其学习者的汉语习得挑战具有跨文化代表性，可为数字人教学模型构建可检验、可迁移的清晰范本。

（二）实验过程

为保证研究过程规范，本文参考中国工业和信息化部发布的《虚拟数字人指标要求和评估方法》，将关于太极数字人的实验过程拟定如下（见图 2）：

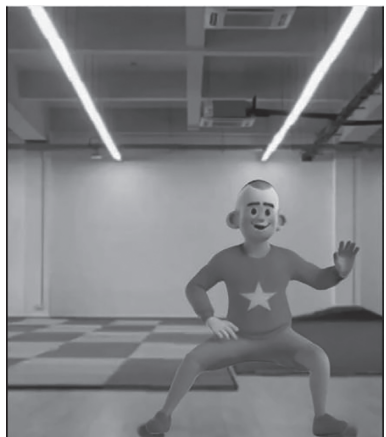
图 2 实验过程图



1. 开发与配置

目前支持自行定制外形与动作的数字人设计工具十分有限。经调研，选用生成速度和质量俱佳的“即梦”作为动画工具，以专业太极教师的教学视频为动作模板。为保证动作高度相似，用“豆包”生成与模板人物外形契合的图片，随后进入“即梦”的“动作模范”版块，上传动作视频并导入图片“做同款”，生成太极数字人视频（见图 3）。

图 3 太极数字人生成效果图



经 SSIM 测试²，形象相似度为 0.97；经 Dartfish 测试³，动作时长与姿势角度几乎完全一致，效果良好。最后将动画导入“剪映”，根据教学需求配置相应要素。

2. 多模态输出

本次实验采用控制变量法和对比实验法，结合前文“具身—负荷”整合框架以及国际中文教学方法，从形象、动作、语言三个维度设计要素对照组，每组控制该维度其他变量为原始变量，具体要素如下（见表 1）：

表 1 实验分组情况

	认知负荷类型	身体感知系统	要素	对照组	
形象	外在认知负荷	视觉 + 听觉	外形 + 语音	真人教师	数字人教师
动作	/			控制组	
	内在认知负荷	动觉	速度	慢速	极慢速
			重复	有重复	分解兼重复
	外在认知负荷	视觉 + 听觉	指令	单一指令	多元指令
			强调	视觉强调	视听强调

2 结构相似性指数，一种基于人眼视觉系统，从亮度、对比度和结构信息综合评估图像相似性的指标。

3 一款专业的运动视频分析软件，提供动作比较分析、动作轨迹自动跟踪、数据统计及生成分析报告等功能，广泛应用于体育教学、技术训练等领域。

语言	内在认知负荷 外在认知负荷	视觉 + 听觉 + 动觉	形式	语动同步	先语后同	先语后动
		视觉 + 听觉	顺序	先中后英		先英后中

第一，形象维度。对比真人教师与数字人教师在外形和语音方面的教学效果。第二，动作维度。首先设置实验控制组，即原速 - 无重复 - 无指令 - 无强调；速度上，对比慢速（ $\times 0.7$ ）与极慢速（ $\times 0.5$ ）；重复上，对比有重复（重复 1 次）与分解兼重复（第 1 遍跟做完整动作，第 2 遍分解后跟做并伴随指令自做，第 3 遍再次跟做完整动作）；指令上，对比单一指令（方向词汇）与多元指令（身体部位词汇 + 方向词汇 + 其他细节词汇）；强调上，对比视觉强调（放大镜）与视听强调（放大镜 + 提示音）。第三，语言维度。形式上，对比语动同步（只进行含语言指令的动作教学）、先语后同（先语言教学，后含语言指令的动作教学）与先语后动（先语言教学，后无语言指令的动作教学）；顺序上，对比先中后英（先中文后英文翻译）与先英后中（先英文后中文翻译）。

将被试按上述要素分为 16 个实验组，每组 25 人。依托学校每周的中华才艺课开展教学，一周共 8 节课，一节课约 1 小时，共 10 周。每周所有课的教学内容、教学场地相同，仅被试与测试要素存在差异。每节课安排 2 组被试，测试由 1 名教师执行，单个被试测试时长约 1 分钟。被试观看数字人教师视频并完成模仿后即时测试，动作维度每节课 6 个动作，答对 1 个得 1 分；语言维度每节课 4 个词语，答对 1 个得 1 分。测试完成后统计习得数据，采用 SPSS 26.0 开展数据分析，系统探讨太极数字人教学效能，评估其学习效果。

（三）数据分析

本文首先基于太极数字人形象、动作和语言的相关要素构建测试方案。然后，开展测试方案检验。结果显示，太极数字人形象、动作和语言三个维度题项的拟合优度全部达标，且具备良好的聚合效度与区别效度，Cronbach's α 系数分别为 0.95、0.98 和 0.96，测试信度较

好，说明上述题项可有效测量太极数字人的教学效果。

三、研究结果

（一）形象测试

本次实验秉承“以学习者为中心”的教学原则，先对学生的太极数字人外形、语音偏好进行调查，发现“大耳朵图图”的偏好度最高（54%）。因此结合上文设计步骤，生成以“大耳朵图图”为外形、语音的太极数字人。与此同时，将真人教师原视频作为对照组，采用李克特量表（非常不喜欢 -1 分，不喜欢 -2 分，不一定 -3 分，喜欢 -4 分，非常喜欢 -5 分），让 25 个学生对两个形象分别进行打分。

之后，将学生分成 2 个被试组，即真人教师组和数字人教师组，每组控制该维度其他变量为原始变量，同时进行动作测试，实验结果如下（见表 2）：

表 2 形象测试结果

数字人教师类型	偏好度		动作	
	平均分	标准差	平均分	标准差
真人教师外形 + 语音	2.73	0.52	1.40	0.69
数字人教师外形 + 语音	4.15	0.96	1.64	0.48

分析表 2 数据可知：外形、语音方面，学生更偏爱“数字人教师”（ $M=4.15$ ），其动作测试得分（ $M=1.64$ ）也略高于“真人教师组”（ $M=1.40$ ）。Pearson 相关系数检验显示，动作得分与偏好度呈显著正相关（ $r=0.509$ ， $p<0.05$ ）。这表明，受被试偏爱的数字人教师外形与语音，可在一定程度上促进动作习得。不同于前人研究多聚焦外语视听学习（Zhang et al., 2025）与知识讲解场景（吴长城等，2024），本研究证实 AI 语音与形象的协同设计对动作模仿的即时表现同样具有积极作用，为其教学应用提供了新的实证参照。

（二）动作测试

1. 单一动作要素教学效果

动作测试部分，设计 9 个被试组，即实验控制组与慢速组、极慢速组、有重复组、分解兼重复组、单一指令组、多元指令组、视觉强调组、视听强调组，每组控制该维度其他变量为原始变量，实验结果如下（见表 3）：

表 3 单一要素动作测试结果

数字人教师类型		平均分	标准差	Cohen's <i>d</i>	95% 置信区间 <i>F</i>	线性回归		
						<i>p</i>	<i>R</i> ²	
实验控制组		1.20	1.47	/	[0.40, 2.01]	/		
速度	慢速	3.21	1.51	1.35	[2.47, 3.95]	8.14	0.007	0.16
	极慢速	2.76	1.65	1.00	[1.90, 3.62]			
重复	有重复	4.82	1.22	2.68	[4.20, 5.44]	60.33	0.000	0.59
	分解兼重复	5.01	0.89	3.20	[4.55, 5.47]			
指令	单一指令	3.19	1.21	1.48	[2.57, 3.81]	15.30	0.000	0.27
	多元指令	3.38	1.13	1.66	[2.80, 3.96]			
强调	视觉强调	1.56	1.38	0.25	[0.82, 2.30]	0.87	0.357	0.02
	视听强调	1.83	1.41	0.44	[1.07, 2.59]			

数据显示，各动作教学要素效果差异显著。“重复”平均分最高、标准差最低，效果突出且习得稳定，其中“分解兼重复”（ $M=5.01$ ， $SD=0.89$ ， $d=3.20$ ，95%CI[4.55,5.47]）优于“有重复”（ $M=4.82$ ， $SD=1.22$ ， $d=2.68$ ，95%CI[4.20,5.44]），更能促进精准掌握；回归显示其影响极显著（ $F=60.33$ ， $p<0.001$ ， $R^2=0.59$ ），解释力过半，为核心影响因素。“强调”平均分最低、标准差最高，效果微弱且个体差异大，“视觉强调”（ $M=1.56$ ， $d=0.25$ ，95%CI[0.82,2.30]）与“视听强调”（ $M=1.83$ ， $d=0.44$ ，95%CI[1.07,2.59]）仅略高于控制组（ $M=1.20$ ），回归显示其无显著影响（ $F=0.87$ ， $p=0.357$ ， $R^2=0.02$ ）。“指令”与“速度”效果居中，“多元指令”（ $M=3.38$ ， $d=1.66$ ，95%CI[2.80,3.96]）优于“单一指令”（ $M=3.19$ ， $d=1.48$ ，95%CI[2.57,3.81]），“慢速”

($M=3.21$, $d=1.35$, 95%CI[2.47,3.95]) 优于“极慢速”($M=2.76$, $d=1.00$, 95%CI[1.90,3.62])。综上, 各要素对动作习得的影响强度呈现清晰梯度: 重复>指令>速度>强调。其中与内在认知负荷相关的“重复”要素的影响强度远高于与外在认知负荷相关的“指令”和“强调”。

2. 复合动作要素教学效果

鉴于要素组合类型较多, 本文从各数字人教师类型中选择效果最优的单一要素, 根据认知负荷类型组合为三类复合教学方案, 即“慢速+分解兼重复”“多元指令+视听强调”“慢速+分解兼重复+多元指令+视听强调”。每类方案再次选取 25 名与前文条件相同的学生作为被试, 在内容、时长等实验条件一致的前提下进行动作测试, 以考察要素复合对于动作习得的效应。将所得数据与前文单一要素“慢速”“分解兼重复”“多元指令”“视听强调”进行对比, 结果如下表所示 (见表 4):

表 4 复合要素动作测试结果

认知负荷类型	数字人教师类型	平均分	标准差	Cohen's d	95% 置信区间	变异系数 (CV)
/	实验控制组	1.20	1.47	/	[0.40, 2.01]	121.18%
内在认知负荷	慢速+分解兼重复	5.21	0.82	3.37	[4.78, 5.64]	10.85%
	慢速	3.21	1.51	1.35	[2.47, 3.95]	36.78%
	分解兼重复	5.01	0.89	3.20	[4.55, 5.47]	12.24%
外在认知负荷	多元指令+视听强调	3.56	1.26	1.72	[2.90, 4.22]	30.53%
	多元指令	3.38	1.13	1.66	[2.80, 3.96]	31.02%
	视听强调	1.83	1.41	0.44	[1.07, 2.59]	85.44%
内在认知负荷+外在认知负荷	慢速+分解兼重复+多元指令+视听强调	5.36	0.67	3.65	[5.01, 5.71]	9.94%

分析表 4 可知, 要素复合的教学效果整体优于单一要素, 且呈现明确梯度: 全要素复合组>内在负荷复合组>外在负荷复合组。内在

认知负荷维度，“慢速+分解兼重复”复合后平均分达 5.21 ($d=3.37$, 95%CI[4.78,5.64])，高于单一“慢速”和“分解兼重复”，变异系数降至 10.85%，明显低于二者的 36.78% 和 12.24%，两类核心内在调节要素产生了显著协同效应，教学效果与普适性同步提升。外在认知负荷维度，“多元指令+视听强调”复合后平均分虽升至 3.56 ($d=1.72$, 95%CI[2.90,4.22])，变异系数降至 30.53%，但仍显著高于内在负荷复合组。当四类要素全部复合时，教学效果达到峰值，平均分最高 ($M=5.36$, $d=3.65$, 95%CI[5.01,5.71])，标准差和变异系数 (9.94%) 均为全组最低。值得注意的是，“强调”要素虽在单因素分析中对动作得分无显著影响，但与其他要素复合后能产生正向协同效应。

(三) 语言测试

语言测试部分，形式维度设计 3 个被试组，即动作教学组（有语言指令）、先语言教学后动作教学组（有语言指令）、先语言教学后动作教学组（无语言指令）；顺序维度设计 2 个被试组，即先中文后英文组、先英文后中文组。每组控制该维度其他变量为原始变量，实验结果如下（见表 5）：

表 5 语言测试结果

	数字人教师类型	平均分	标准差
形式	语动同步	2.20	0.75
	先语后同	3.30	0.64
	先语后动	2.80	0.64
顺序	先中后英	3.00	0.77
	先英后中	3.20	0.60

数据显示，形式维度，“先语后同”的平均分 ($M=3.30$) 显著高于“语动同步” ($M=2.20$) 和“先语后动” ($M=2.80$)，且标准差最小 ($SD=0.64$)。顺序维度，“先英后中” ($M=3.20$) 略优于“先中后英” ($M=3.00$)。综上，语言-动作组合形式采用“先语后同”、词汇呈现顺序采用“先英后中”的教学表现相对更优，且习得稳定性更强。

四、研究讨论

（一）理论机制

1. 具身交互助推负荷动态平衡

（1）强化“身体—认知”协同加工，降低内在认知负荷为主

虽然很多学者认为内在认知负荷在既定学习条件下较难改变 (Sweller et al., 2019)，但本研究发现相较于降低外在认知负荷，降低内在认知负荷对改善学习效果的影响更为显著。此外，身体呈现节奏与认知加工规律协同契合，更利于降低内在认知负荷。以上结论在动作与语言习得两个维度均得到验证。

动作方面，“重复”要素的效应量 ($R^2=0.59$) 远高于“指令”与“强调”，说明调节内在认知负荷是动作习得的主要环节。在此基础上，“分解兼重复”策略将复杂动作切分为符合工作记忆容量的独立单元，在学习者动觉模仿与认知加工过程中 (Mayer et al., 2003)，经由动作重复推动图式自动化与整合 (Sweller et al., 1998)，精准匹配认知节奏以最大限度降低内在负荷；“慢速”有助于学习，而“极慢速”会割裂动作整体性、增加加工片段，反而会提升内在负荷，进一步佐证了身体节奏需匹配认知规律的观点。语言方面，“语动同步”易导致语言与动作共享工作记忆资源，引发资源竞争致使负荷过载，而“先语后同” ($M=3.30$) 通过分阶段加工阻断冲突，有效降低语言、动作各自的内在负荷，使结果显著优于“语动同步” ($M=2.20$)。同时，“先语后同”的教学序列契合学习者认知分加工节奏，能进一步强化内在负荷调控效果；“先英后中” ($M=3.20$) 优于“先中后英” ($M=3.00$)，也表明以母语为支架降低目的语负迁移、匹配语言认知基础，有助于优化内在认知负荷。

（2）优化“环境—身体”感知匹配，降低外在认知负荷为辅

在优先调节内在认知负荷的基础上，外在认知负荷的合理优化同样不可忽视。具身认知理论指出，认知植根于环境与身体感知系统的实时互动 (张旭, 2024; 罗雪英等, 2024)，这意味着当环境信息与

身体执行的认知任务高度相关时，外在负荷的降低能更为有效地转化为关联负荷的提升。这些结论在研究结果中同样可以得到印证。

动作方面，“指令”对动作得分也具有很大影响（ $R^2=0.27$ ），表明外在认知负荷的调节对于动作习得同样重要。此外，“多元指令”将身体部位、方向等词汇与动作执行直接对应，教学效果优于“单一指令”；“视听强调”的单独效果不佳，但与其他要素复合后产生正向协同效应；四要素复合时平均分（ $M=5.36$ ）和变异系数（9.94%）均达最优，上述内容都反映出外在要素需服务于动觉习得的核心任务，其亦可从认知负荷“成本—效益”视角进行解释（Skulmowski et al., 2022），即某些教学要素虽然会带来少量外在认知负荷，但若能使学习者将注意力更精准地聚焦于关键信息，其产生的学习效益依旧能超出成本。形象方面，学习者对大耳朵图图的偏好度与动作得分呈显著正相关（ $r=0.509$, $p<0.05$ ），这也可说明匹配身体感知、认知偏好的环境信息设计能削弱适应性障碍，降低外在负荷，提升学习效果。

2. 负荷优化赋能具身认知加工

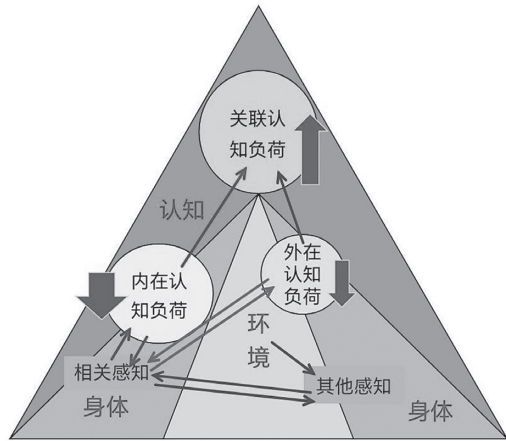
认知负荷的优化会反向赋能具身认知三要素的交互深度，与正向调节形成一定闭环：内在认知负荷的降低可深化身体与认知的协同加工，外在认知负荷的降低能强化环境与身体的感知匹配，两类认知负荷共同推动具身三要素的深度整合。

内在认知负荷降低后，充足的工作记忆空间能支撑深度加工，令身体从机械模仿升级为可自动化且同语义表征绑定的动作图式。“原速”“无重复”导致负荷过载，被试只能被动跟随，运用“分解兼重复”策略降低负荷后，身体动作不再是单纯肢体运动，可逐步形成自动化图式；语动同步组表现欠佳，很大程度源于负荷超标引发语言与动作加工争抢认知资源（He et al., 2024），依托“先语后同”策略拆分加工环节、合理减负，能让身体动作承载语言记忆，实现二者深度联结。外在认知负荷降低后，学习者无需消耗认知资源处理无关信息（Zhang et al., 2025），环境要素与身体感知系统的互动效率显著提升。这一点可从真人教师组与数字人教师组的对比中验证：外形、语音等

环境要素契合认知偏好时，适应性障碍带来的外在负荷被削弱，学习者能更为专注地接收动作指令，视听输入与动觉执行的匹配度大幅提高；同样，单一视听强调效果不佳，主要也是因为其增加额外外在负荷，干扰环境与身体的正常互动，只有外在负荷调控至合理水平时，视听强调才能精准引导注意力，发挥积极作用。

基于上述实证结果，本研究进一步构建“具身—负荷”协同机制（见图4），以可视化方式呈现环境、身体、认知三要素与三类认知负荷的双向作用关系。与图1的理论框架不同，图4的“具身—负荷”协同机制是基于实证数据提炼的作用路径，揭示了数字人教师赋能国际中文教育的内在原理。“具身—负荷”协同机制包含两个核心作用路径：（1）正向路径：身体的相关感知要素主要调节内在认知负荷，环境中的其他感知要素与身体的相关感知要素调节外在认知负荷，认知要素促进关联认知负荷提升。（2）反向路径：内在认知负荷的降低深化身体与认知的协同加工，外在认知负荷的降低强化环境与身体的感知匹配，两类认知负荷共同推动具身三要素的深度整合。这种双向作用形成了一个动态闭环，以降低内在认知负荷为主、降低外在认知负荷为辅，共同促进动作和语言习得。这一机制阐明了数字人教师赋能国际中文教育的内在原理，可为相关教学实践提供理论支撑。

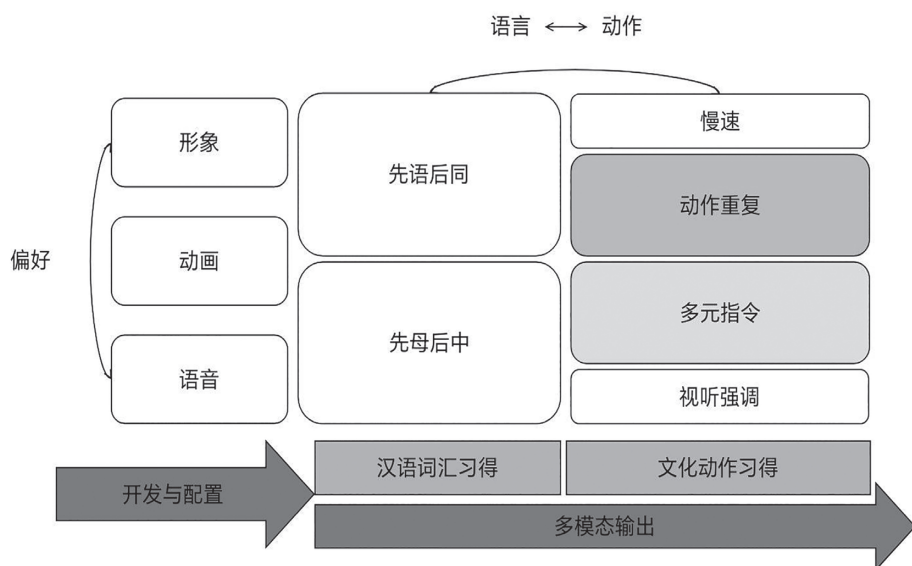
图4 “具身—负荷”协同机制



（二）实践路径

前文分析结果表明，契合学习者偏好的数字人形象、语音与动画深度融合，能够有效提升教学效果，文化动作与汉语词汇相互适配，可推动语言与文化双习得。基于上述结论，本文提出国际中文教育数字人教师生成路径（见图5），整体划分为开发配置、多模态输出两大阶段，实现双向统筹、协同优化。

图5 “三合二分四联”路径



1. 路径详解

（1）开发与配置：“三合”

从形象、动画、语音三方面对数字人教师进行设计。本研究选用新西兰小学生喜爱的“大耳朵图图”作为数字人角色，并利用即梦软件生成动画。设计者可结合用户群体实时调整：成年学习者可选用更经典、专业的形象，老年学习者可配以更清晰、标准的语音。上述设计由用户偏好决定，旨在最大程度降低适应性问题造成的外在认知负荷，激发学习兴趣，加强学习者与环境的交互。

（2）多模态输出

①汉语词汇习得：“二分”

实现“文化动作—汉语词汇”双习得，需将语言教学与动作教学适当分开，以保证各自的主体地位，分别降低两部分的内在认知负荷。同时，在动作教学部分以习得词汇作为指令语，通过多模态输入与不同类型负荷的融合，实现词汇与动作习得效果的同步提升。语言教学部分，应以学习者母语为支架，先显示母语翻译，再展示中文词汇，使目的语负迁移造成的外在认知负荷最小化。

②文化动作习得：“四联”

通过多组对比实验，构建适用于新西兰小学生的“慢速($\times 0.7$)—分解兼重复—多元指令—视听强调”四方面设计模式，以放慢速度、多次重复降低动觉模仿的内在负荷，强化动作图式，并将视觉、听觉与动作内容相融合，共同促进动作技能的持续提升。

上述模式可根据教学对象灵活调整：成年学习者可调节词汇数量与指令难度，老年学习者可进一步放慢速度。

2. 路径价值

①减轻教师负担，提升教学效果

数字人教师可辅助教师完成部分内容的教授与演示，有效减轻工作负担，同时解决了教师展示方向相反、动作有误、口令难以兼顾等问题，在提高教学效率的同时完善教学效果。

②拓展数字人教师的汉语教学应用场景

数字人技术在书法、戏曲、中医等中华文化技能传承中具有广阔应用潜力，可对实践路径进行迁移。例如，书法教学中，数字人教师可通过“慢速+分解兼重复+多元指令（如‘横画起笔轻顿’）+视听强调（箭头标注运笔轨迹+提示音）”，同步实现书法动作掌握与“起笔、行笔、收笔”等汉语词汇记忆。其他类型数字人教师亦可借鉴：词汇教学数字人可采用“先母后中”模式，结合动觉巩固词汇；知识讲解数字人可借助“多元指令”“视听强调”激发兴趣，以“内容重复”加深印象，进一步优化教学效果。

五、未来展望

（一）扩大样本量、延长实验周期

后续研究可进一步扩大样本覆盖范围，采用多中心、分层抽样设计，覆盖不同年龄段、不同汉语水平和不同地域的学习者群体，确保样本的多样性。同时，将实验周期适当延长，在实验结束后增设追踪测试，以“动作复现率”“词汇回忆准确率”为指标，评估数字人教师对长期学习记忆的影响，并通过记录“持续参与率”分析其对学习黏性的长期作用。

（二）评估维度多样化

未来可对更多要素组合实验验证，以进一步检验和丰富现有结论。同时，构建多维度评估指标体系，如通过半结构化访谈评价学习者对太极“阴阳平衡”内涵的理解；设计“用汉语描述3个书法动作”等任务，从“词汇准确性”“语句连贯性”角度评分。还可引入眼动、脑电等定量数据与访谈记录、汉语描述文本等定性数据，全面反映数字人教师的综合赋能价值。

（三）优化数字人教师形象

目前“即梦”要求导入图片的人物大小、轮廓与视频模板高度相似，限制了卡通形象的完整复刻和多元文化符号的覆盖。对此，可加强与AI研发团队的技术合作，优化图像适配算法，降低对输入图片轮廓和尺寸的限制，支持局部调整，以生成外形更精准、动画更流畅的数字人教师。

（四）丰富实时互动、沉浸式等功能

在现有基础上，可集成Kinect传感器等动作捕捉技术与语音识别技术，构建“实时反馈系统”。动作层面，通过传感器比对学习者和数字人教师的动作偏差，即时弹出提示，如“手臂抬高5°”；语言层

面，通过语音识别判断发音准确度，生成评分与纠正建议。此外，可融合 VR/AR 构建沉浸式场景，让学习者在虚拟环境中跟随数字人教师练习，强化“身体与环境实时互动”的具身体验和学习临场感。

六、结语

本研究以具身认知理论与认知负荷理论为基础，利用数字人教师进行太极教学，对 482 名新西兰初级汉语学习者展开实验，核心结论如下：将学习者偏好度较高的形象、语音与数字人动画相结合，能有效降低习得的外在认知负荷；语言教学以“先语后同—先母后中”为最佳设计模式，动作教学“慢速—分解兼重复—多元指令—视听强调”的设计模式教学效果显著，二者彼此独立又相互促进，实现身体与环境有效交互、内外认知负荷相互协同，推动主要认知的不断精进与完善。本文基于以上结论构建“具身—负荷”协同机制与“三合二分四联”路径，在国际中文教育数字人教师于动作技能教学及数智化实践层面有所创新。

在今后的研究中，可通过延长实验周期、添加追踪测试、增多对比类型、丰富研究方法等，对数字人教师的复习策略、习得机制进一步探索。此外，可以集成动作捕捉、语音识别与 VR/AR 技术，开发具备实时纠错功能、高具身程度的多功能数字人教师；还可根据学习者的表现，动态调整教学策略，开发自适应算法以适配不同群体需求，进而生成更加完整的“复习—教学—总结—测评”的数字人资源链条。此外，应推动数字人成果与国际中文教育资源库对接，助力“中文+文化”的全球化、标准化传播，为数字技术赋能国际中文教育提供更具实践价值的方案。

参考文献

- 邓国民、阎婷 (2019): 面向智能时代的跨媒体学习方式、理论与资源环境——基于国际多媒体学习研究的启示,《远程教育杂志》, 37(03),61-69。
- 郭亚军、庞义伟、周家华等 (2024): ChatGPT 赋能图书馆虚拟数字人: 技术优势、应用场景与实践路径,《图书馆论坛》, 44(08),69-79。
- 李青、刘勇 (2024): 教育虚拟数字人标准体系设计及其路径规划,《现代教育技术》, 34(07),70-80。
- 罗雪英、尚硕 (2024): 沉浸、互动、想象: 具身认知视域下的古籍+VR 开发实践探析,《出版发行研究》, (06),54-62。
- 庞维国 (2011): 认知负荷理论及其教学涵义,《当代教育科学》, (12),23-28。
- 王春辉 (2024): 数字人进场的语言生活,《外语学刊》, (06),41-50。
- 魏小东、习鹏飞 (2024): 大语言模型驱动下的虚拟数字人第二语言教育应用研究,《电化教育研究》, 44(05),75-82。
- 吴长城、胡双武、蒋雨江 (2024): GenAI 驱动的教育数字人架构设计与实证研究,《现代教育技术》, 34(09),26-36。
- 余超凡、赵志群 (2025): AIGC 时代数字人赋能人机协同学习: 价值、框架与路径,《远程教育杂志》, 43(02)。
- 张宁、Miguel Baptista Nunes、李俊炀 (2021): 面向中华古籍阅读推广与文化传播的 VR 系统模型构建与实现,《图书情报工作》, 65(13),12-24。
- 张旭 (2024): 具身认知理论视域下古籍资源游戏化开发策略研究——以“故纸修复师·碎丹青”为例,《图书馆学研究》, (11),93-100。
- 周世豪、陈天序 (2025): 基于课堂体验的汉语二语学习者对数字人教师技术的接受度探究 // 北京大学对外汉语教育学院。2025 对外汉语博士生论坛暨第十八届对外汉语教学研究生学术论坛论文集, 中央民族大学国际教育学院, 373-380。
- Chandler P., & Sweller J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233-246.
- Deng L.X., Zhou Y., Cheng T.L., Liu X., Xu T., Wang X. (2022). My English teachers are not human but I like them: Research on virtual teacher self-study learning system in K12. *Lecture Notes in Computer Science*, 13329, 1-15.
- He C.Y., Chen Y.Y., Phang C.R. (2024). Exploring embodied cognition and brain dynamics under multi-tasks target detection in immerse projector-based augmented reality (ipar) scenarios. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32, 3476-3486.
- Lin Y.P., & Yu Z.G. (2024). Learner perceptions of artificial intelligence-generated pedagogical agents in language learning videos: Embodiment effects on technology acceptance. *International Journal of Human-Computer Interaction*, 41, 1606-1627.
- Mayer R., & Moreno R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43-52.
- Mayer R.E., & Fiorella L. (2005). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. *The Cambridge handbook of multimedia learning*, 1, 183-200.
- Paas F., & Van Merriënboer J.J.G. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, 29, 394-398.
- Rotgans J.I., & Schmidt H.G. (2011). The role of teachers in facilitating situational interest in an active-learning classroom. *Teaching and Teacher Education*, 27, 37-42.
- Skulmowski A., & Rey G.D. (2017). Measuring cognitive load in embodied learning settings. *Frontiers in Psychology*, 8, 1191.

- Skulmowski A., & Xu K.M. (2022). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review*, 34, 171-196.
- Sweller J., van Merriënboer J.J.G., Paas F.G.W.C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.
- Sweller J., van Merriënboer J.J.G., Paas F.G.W.C. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 261-292.
- Tesán L., González D., Martins P. (2025). Thermodynamics-informed graph neural networks for real-time simulation of digital human twins. *Computational Mechanics*, 76, 1-22.
- van Merriënboer J.J.G., & Sweller J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17, 147-177.
- Zhang X.Y., Zhang T.Y., Wei K.L. (2025). Cognitive load suppresses explicit learning while sparing implicit learning in visuomotor adaptation. *Journal of Neurophysiology*, 134, 1133-1145.
- Zhang Y.D., Lucas M, Bem-haja P, Pedro L. (2025). AI versus human-generated voices and avatars: Rethinking user engagement and cognitive load. *Education and Information Technologies*, 30, 22547-22566.
- Zou L.Y., Zhang Z.H., Mavilidi M. (2025). The synergy of embodied cognition and cognitive load theory for optimized learning. *Nature Human Behaviour*, 9, 877-885.

An Empirical Study on Taichi Digital Human Teachers Empowering International Chinese Language Education

SUN, Yutong ZHU, Yong*

Abstract

With the rapid advancement of artificial intelligence, digital human teachers are creating new opportunities for international Chinese language education; however, theoretical mechanisms and practical pathways grounded in objective pedagogical effectiveness remain underdeveloped. Taichi, which integrates physical movements with cultural vocabulary, serves as a comprehensive teaching paradigm for designing digital human teachers. Taking Taichi as the medium, this study constructs a framework combining embodied cognition theory and cognitive load theory and conducts an empirical investigation involving 482 beginner Chinese language learners in New Zealand. The results reveal that the synergistic interaction of body, environment, and cognition dynamically regulates learners' intrinsic, extraneous, and germane cognitive loads. Building on this mechanism, the developed pathway consists of a three-in-one configuration of "appearance, animation, and voice", a two-step language acquisition approach of "speech before synchronization and native language before Chinese," and a four-dimensional movement acquisition approach characterized by "slow speed, decomposition and repetition, multimodal instructions, and audio-visual emphasis." This pathway significantly enhances immediate movement mastery and rapid language memorization. Beyond reducing the burden of teacher demonstrations and improving instructional efficiency, the pathway can be extended to broader cultural and linguistic teaching contexts, providing more foundations for the engineering

transformation of international Chinese language education.

Keywords: digital human teacher, international Chinese language education, Embodied Cognition Theory, Cognitive Load Theory

SUN, Yutong, School of International Chinese Language and Culture, East China Normal University.

* ZHU, Yong, School of Chinese Language and Literature, Beijing Foreign Studies University. (corresponding author)