

人机协同翻译中学习者认知负荷的动态演变研究

——基于汉语块状语英译的纵向追踪分析

赵振强, 沈越, 刘绍龙

浙江越秀外国语学院应用外语学院, 浙江 绍兴

收稿日期: 2025年12月1日; 录用日期: 2025年12月28日; 发布日期: 2026年1月5日

摘要

本研究基于认知负荷理论, 运用纵向追踪实验与多元数据采集方法, 系统考察高校翻译学习者在人机协同环境下处理汉语块状语英译时的认知负荷变化轨迹及策略调适机制。研究发现: (1) 人机协同模式显著降低了学习者的内在认知负荷, 但初期引发外在认知负荷的短暂升高; (2) 学习者经历“机械依赖 - 批判反思 - 策略内化”三阶段的发展历程, 策略使用频次与使用效能呈现倒U型演进曲线; (3) 元认知调控能力的发展是认知负荷优化的关键, 具备较强元认知监控的学习者能提前2~3周完成策略内化; (4) 基于认知支架设计的渐进式教学干预能有效促进学习者从表层模仿到深层理解的跨越。本研究在理论层面丰富了翻译教学中认知负荷的动态理论, 在实践层面为高校翻译课程的人机协同教学设计提供了发展性评估框架和阶段性干预路径, 对培养人工智能时代的翻译人才具有参考价值。

关键词

认知负荷, 人机协同, 汉语块状语

A Study on the Dynamic Evolution of Learners' Cognitive Load in Human-Machine Collaborative Translation

—A Longitudinal Analysis Based on the English Translation of Chinese Chunk Expressions

Zhenqiang Zhao, Yue Shen, Shaolong Liu

School of Applied Foreign Languages, Zhejiang Yuexiu University, Shaoxing Zhejiang

Received: December 1, 2025; accepted: December 28, 2025; published: January 5, 2026

Abstract

Grounded in cognitive load theory, this study uses longitudinal tracking and multi-method data collection to investigate how translation students adapt their cognitive load and strategies when translating Chinese chunks into English in a human-machine collaborative setting. Key findings show that: (1) Human-machine collaboration significantly lowers intrinsic cognitive load but causes a brief rise in extraneous load initially; (2) Learners progress through three stages-mechanical dependence, critical reflection, and strategy internalization with strategy use forming an inverted U-shaped curve; (3) Stronger metacognitive skills allow learners to internalize strategies 2~3 weeks faster; (4) Scaffolded teaching interventions help learners move from surface imitation to deep understanding. The study contributes to cognitive load dynamics in translation education and offers a practical framework for designing collaborative human-machine translation courses.

Keywords

Cognitive Load, Human-Machine Collaboration, Chinese Chunk

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

人工智能技术的发展正在深刻改变翻译教学的认知模式。汉语块状语作为汉语特色表达的典型代表，以四字格为主要形态承载丰富的文化内涵，其英译过程涉及语义解构、文化转换、语用重构等多层次认知操作，对学习者的认知资源提出较高要求。传统教学中，学习者需独立完成全流程翻译，认知负荷高度集中且分配失衡，导致学习效率低下。人机协同模式的引入为优化认知负荷提供了契机，但教学实践发现，学习者在面对机器译文时往往陷入过度依赖或完全排斥两种极端。人机协同环境如何改变认知负荷结构？学习者如何发展有效策略？不同认知水平学习者呈现何种差异化轨迹？这些问题亟待回答。鉴于此，本研究采用 16 周纵向追踪实验，系统考察高校翻译专业学生处理汉语块状语英译时的认知负荷变化与策略发展轨迹，旨在揭示人机协同环境下学习者的认知适应过程与发展规律，为翻译教学改革提供理论依据和实践路径。

2. 文献综述

2.1. 认知负荷理论在翻译教学中的应用

认知负荷理论由 Sweller (1988) 提出，强调人类工作记忆容量有限，有效学习需要合理分配认知资源 [1]。该理论区分内在负荷(由任务本身的复杂性决定)、外在负荷(由不当的教学设计引起)和相关负荷(用于图式构建与自动化的有效负荷)三类。近年来，认知负荷理论逐渐拓展至翻译教学研究领域，成为分析译员认知过程的重要理论框架。苏雯超等(2021)基于该理论探讨了口译教学的任务难度设计，发现渐进式

难度调节能有效优化学习者的认知资源分配，减轻外在负荷并促进技能自动化[2]。翟秋蘭(2025)进一步考察了笔译教学中不同教学模式对学习者认知负荷的影响，证实范例学习通过提供清晰的结构和可参照的认知路径，能显著降低外在负荷并提升相关负荷，从而提高翻译质量和学习效率[3]。然而，现有研究多关注传统人工翻译模式，对技术介入下如人机协同翻译环境中认知负荷的动态变化、技术工具对内在与外在负荷的调节机制以及多任务处理对认知资源分配的影响仍缺乏系统且深入的实证分析。

2.2. 人机协同翻译的学习效应研究

人机协同翻译对学习效果的影响已成为翻译教育领域的研究热点。Trojczak (2022)通过控制组实验发现，接受系统译后编辑训练的学生在翻译准确性和效率方面显著优于未受训练组，但研究同时指出需要培养学生对机器输出的批判性技术素养，避免盲目接受机译结果[4]。郑颖等(2025)的教学实验证实，人机协同训练能有效提升学生处理文化负载表达和语境适应能力，通过对比分析、错误识别和创造性重构等环节增强跨文化交际能力，但研究发现在训练初期存在学生过度依赖机译输出的问题，需要教学干预来平衡人机协作关系[5]。杨艳霞等(2025)从认知视角探讨了机器翻译介入对学习者策略发展的影响，提出应系统培养学生在预处理、实时协作和译后编辑三个阶段的机译评估能力、修改决策能力和质量监控意识，从而最大化人机协同的认知增益[6]。这些研究为理解协同翻译模式的教学价值提供了实证证据，但对学习者在长期训练中认知负荷的纵向演变规律、不同能力阶段策略发展的结构性特征以及人机互信关系建立的认知机制仍缺乏多维度、深层次的系统考察。

2.3. 翻译学习者的策略发展研究

翻译策略发展是翻译教学的核心目标，关系到学习者能否有效解决翻译过程中的实际问题。Chesterman (2016)将翻译策略系统划分为语法策略、语义策略以及语用策略三个层级，为策略教学提供了理论框架[7]。鲁华山等(2025)通过为期一学期的纵向追踪研究，发现学习者的策略使用呈现出从表层结构模仿到深层意义重构、从单一策略试误到多元策略系统整合的发展路径，并指出元认知能力——包括策略选择意识、过程监控与效果评估——是影响个体策略发展速度与质量的关键变量[8]。吕倩兮等(2025)借助屏幕录制与回溯性访谈，细致分析了高、低水平学习者在译后编辑任务中的策略使用差异，发现高水平学习者不仅能灵活调用语法、语义和语用层面的多种策略，还能在问题识别、策略选择与执行效果间形成有效的自我调节与监控机制[9]。然而，现有研究多局限于传统或纯人工翻译情境，对人机协同翻译环境中策略发展如何与动态变化的认知负荷相互作用、技术工具对策略选择的影响机制以及人机协作能力与策略发展阶段的适配关系仍缺乏系统性的理论构建与实证探讨。

2.4. 汉语块状语翻译的教学研究

块状语翻译是翻译教学的重难点。游凤艳(2022)指出，四字格翻译要求学习者同时具备语义解构、文化转换和修辞再现能力，学习者普遍存在文化意识薄弱、策略单一等问题[10]。李小芳(2025)发现，机器翻译系统在处理文化负载四字格时准确率仅为 45%，远低于一般词汇，学习者在编辑时的修改集中于语义补充和文化适应[11]。赵振强等(2024)通过教学实验证实，基于人机协同的块状语翻译训练能提升学生的文化敏感性，但需要系统的策略指导[12]。

综观现有研究，学界在认知负荷理论、人机协同学习、策略发展等领域积累了丰富成果，但仍存在不足：缺乏针对人机协同环境下学习者认知负荷动态变化的纵向追踪研究；块状语翻译的策略发展研究多基于横向对比，对发展过程的阶段性特征关注不足；定量与定性研究的整合不足。本研究针对上述不足，采用混合方法系统考察学习者在人机协同块状语翻译中的认知负荷调适与策略发展机制。

3. 理论基础与研究设计

3.1. 理论框架

本研究以认知负荷理论(Cognitive Load Theory)为核心理论支撑，整合策略发展理论(Strategy Development Theory)与脚手架教学理论(Scaffolding Theory)，构建汉语块状语人机协同英译学习的认知分析框架。

认知负荷理论由 Sweller 等人发展，其核心观点是人类工作记忆容量有限(约 7 ± 2 个信息单元)，有效学习需要合理分配认知资源[13]。在翻译学习中，内在负荷源于块状语本身的语义复杂性和文化负载，外在负荷来自不当的教学设计或技术界面，相关负荷指促进图式建构的深层认知加工。人机协同模式通过机器承担部分形式转换任务，理论上可以降低内在负荷，释放认知资源用于高层次加工，但技术介入也可能引入新的外在负荷。

策略发展理论强调学习者策略使用经历从无意识到有意识、从低效到高效的发展历程。Anderson 的 ACT 理论提出，技能习得包括认知阶段、联结阶段和自动化阶段[14]。在翻译学习中，学习者最初依赖显性规则和外部辅助，通过反复练习逐步将策略内化为自动化程序。人机协同环境改变了这一发展路径：机器译文作为“认知外部化”，既可能成为学习支架，也可能阻碍策略的内部化。

脚手架理论源于 Vygotsky 的最近发展区概念，强调通过临时性支持帮助学习者完成超出当前能力的任务[15]。在人机协同翻译教学中，脚手架表现为机器译文提供初始框架、教师指导引导策略发展、同伴互动促进反思调整。有效的脚手架应遵循渐进撤除原则，避免学习者形成依赖。

3.2. 研究问题

基于上述理论框架，本研究提出以下三个核心问题：

- 1) 人机协同环境下，学习者在处理汉语块状语英译时的认知负荷如何分布与演变？不同类型认知负荷的变化轨迹如何？
- 2) 学习者的翻译策略发展经历哪些阶段？各阶段的策略使用特征与认知负荷的关系如何？
- 3) 如何设计有效的教学干预促进学习者的认知负荷优化与策略内化？

3.3. 研究设计

3.3.1. 参与者

本研究在某高校翻译专业开展 16 周的纵向追踪实验。参与者为 36 名大二学生(男 8 名，女 28 名)，平均年龄 19.7 岁，均已完成《综合英语》《基础翻译》等课程，具备基本的翻译能力。所有参与者在实验前未接受过系统的人机协同翻译训练。

3.3.2. 实验材料

实验材料包含 80 个高频汉语四字格，选自政论、文学、商务三个领域。根据语义复杂度和文化负载，将四字格分为低、中、高三个难度等级，每个等级约 27 例。每周安排 5 个四字格翻译任务，难度采用螺旋上升设计。

3.3.3. 实验流程

研究采用“基线测试(第 1 周)→渐进式训练(第 2~14 周)→延迟测试(第 16 周)”的三阶段设计。

(1) **基线测试阶段(第 1 周)**：学生完成 10 个四字格的人工翻译任务和 10 个人机协同翻译任务，建立认知负荷和策略使用的基准线。

(2) **渐进式训练阶段(第 2~14 周)**：每周包含三个教学环节：① 任务实践：学生使用 DeepSeek 翻译系统生成初译，在机译基础上进行编辑优化；② 策略讲解：教师每两周讲解一类翻译策略(语义分析策

略、文化适应策略、修辞重构策略等);③ 反思讨论: 学生分享译后编辑过程中的决策点和困惑点, 开展同伴互评。

(3) 延迟测试阶段(第 16 周): 学生完成与基线测试同等难度的新四字格翻译任务, 评估学习效果的保持。

3.3.4. 数据收集

研究采用多元数据采集方法:

(1) 认知负荷测量。每次翻译任务后, 学生填写 Paas (1992) 的 9 点量表[16], 评估主观认知负荷(1 = 非常容易, 9 = 非常困难)。使用眼动仪(Tobii Pro Spectrum)记录第 1、8、16 周的关键任务, 采集注视时间、瞳孔直径等生理指标作为认知负荷的客观测量。

(2) 策略使用数据。使用屏幕录制软件记录译后编辑过程, 编码修改操作类型(增加、删减、替换、重组)。每 4 周进行一次回溯式访谈, 让学生解释关键决策点的思考过程。

(3) 翻译质量评估。采用 BLEU 自动评价和人工评分(5 分制)相结合的方法, 评估译文质量。人工评分由两名教师从准确性、流畅性、文化适应性三个维度打分, 一致性系数 0.87。

(4) 元认知问卷。采用改编的 Schraw 和 Dennison 的元认知意识问卷[17], 在第 1、8、16 周测量学生的元认知水平。

3.3.5. 数据分析

定量数据采用 SPSS 27.0 进行分析: 重复测量方差分析(Repeated Measures ANOVA)考察认知负荷和翻译质量的时间变化趋势; 相关分析探讨认知负荷、策略使用、元认知水平与翻译质量的关系; 聚类分析识别不同的学习者发展类型。定性数据采用 NVivo 12 进行编码分析: 开放编码识别策略类型和使用情境; 主轴编码建立策略与认知负荷的关联; 选择性编码构建策略发展的阶段模型。

4. 数据呈现与分析

4.1. 定量结果

4.1.1. 认知负荷的动态演变

表 1 呈现了学习者在 16 周内主观认知负荷的变化趋势。

Table 1. Subjective cognitive load scores at different stages ($M \pm SD$)

表 1. 不同阶段主观认知负荷得分($M \pm SD$)

测试时间	内在负荷	外在负荷	相关负荷	总体负荷
第 1 周	7.2 ± 1.1	4.8 ± 1.3	3.5 ± 0.9	5.2 ± 0.8
第 4 周	6.5 ± 1.0	5.9 ± 1.2	4.2 ± 1.0	5.5 ± 0.7
第 8 周	5.8 ± 0.9	4.2 ± 1.1	5.3 ± 0.8	5.1 ± 0.6
第 12 周	5.1 ± 0.8	3.6 ± 1.0	5.8 ± 0.7	4.8 ± 0.5
第 16 周	4.9 ± 0.7	3.2 ± 0.9	6.1 ± 0.8	4.7 ± 0.5

重复测量方差分析显示, 总体认知负荷在 16 周内呈显著下降趋势($F(4, 140) = 23.67, p < 0.001$), 从基线的 5.2 分降至 4.7 分, 下降幅度 9.6%。但不同类型认知负荷呈现差异化变化模式:

(1) 内在负荷。从第 1 周的 7.2 分持续下降至第 16 周的 4.9 分, 降幅达 31.9% ($F(4, 140) = 45.32, p <$

0.001)。事后检验显示,第8周是关键转折点,此后下降速度放缓。这表明随着学习者对块状语义结构和英译规律的熟悉,材料本身的复杂性感知显著降低。

(2) **外在负荷**。呈现“先升后降”的倒U型曲线。第1周为4.8分,第4周升至5.9分(增幅22.9%),随后逐步下降至第16周的3.2分。第4周的峰值反映了学习者在初期适应人机协同界面、学习评估机译输出时的额外认知投入。随着技术熟练度提升和决策模式的程序化,外在负荷逐步降低。

(3) **相关负荷**。呈现持续上升趋势,从第1周的3.5分增至第16周的6.1分,增幅74.3% ($F(4, 140) = 38.91, p < 0.001$)。相关负荷的增加意味着学习者将更多认知资源投入到促进理解、技能建构与图式形成的深层加工,这符合认知负荷理论的核心观点,是学习有效性的积极信号。

表2的眼动数据进一步验证了主观评估的结果。

Table 2. Eye-tracking indicators at key time points ($M \pm SD$)

表2. 关键时间点的眼动指标($M \pm SD$)

测试时间	总注视时间(秒)	平均注视时长(毫秒)	平均瞳孔直径(mm)	回视次数
第1周	124.3 ± 18.2	285.6 ± 32.4	4.82 ± 0.41	18.3 ± 4.2
第8周	98.7 ± 15.6	238.9 ± 28.1	4.21 ± 0.38	12.6 ± 3.5
第16周	87.4 ± 14.2	215.3 ± 24.7	3.95 ± 0.35	9.8 ± 3.1

总注视时间从第1周的124.3秒降至第16周的87.4秒,降幅29.7% ($F(2, 70) = 34.82, p < 0.001$),表明学习者的加工效率显著提升。瞳孔直径作为认知负荷的生理指标,从4.82mm降至3.95mm,降幅18.0% ($F(2, 70) = 28.54, p < 0.001$),客观验证了认知负荷的下降。回视次数的减少(从18.3次降至9.8次,降幅46.4%)说明学习者对机译输出的评估更加高效,决策更加果断。

4.1.2. 翻译质量的发展轨迹

表3呈现了学习者译文质量的变化趋势。

Table 3. Translation quality scores at different stages ($M \pm SD$)

表3. 不同阶段翻译质量得分($M \pm SD$)

测试时间	BLEU	准确性	流畅性	文化适应性	综合得分
第1周	0.521 ± 0.082	3.12 ± 0.64	3.28 ± 0.57	2.85 ± 0.72	3.08 ± 0.53
第4周	0.558 ± 0.078	3.45 ± 0.61	3.52 ± 0.54	3.18 ± 0.68	3.38 ± 0.51
第8周	0.624 ± 0.071	3.89 ± 0.58	3.87 ± 0.52	3.76 ± 0.64	3.84 ± 0.48
第12周	0.673 ± 0.065	4.21 ± 0.55	4.18 ± 0.49	4.12 ± 0.59	4.17 ± 0.45
第16周	0.695 ± 0.062	4.38 ± 0.52	4.35 ± 0.47	4.29 ± 0.56	4.34 ± 0.43

重复测量方差分析显示,综合得分从第1周的3.08分提升至第16周的4.34分,增幅40.9% ($F(4, 140) = 67.45, p < 0.001$)。BLEU得分的提升(从0.521到0.695,增幅33.4%)与人工评分高度一致,验证了学习效果的可靠性。

值得注意的是,三个维度的提升幅度存在差异:文化适应性得分从2.85分提升至4.29分,增幅50.5%,显著高于准确性(40.4%)和流畅性(32.6%)。这表明人机协同模式在促进文化维度能力发展方面具有特殊优

势，可能与机译在文化转换上的系统性不足激发学习者深层思考有关。

4.1.3. 认知负荷与翻译质量的相关

表 4 的相关分析揭示了认知负荷各维度与翻译质量的关系。

Table 4. Correlation coefficients between cognitive load and translation quality

表 4. 认知负荷与翻译质量的相关系数

变量	综合得分	准确性	流畅性	文化适应性
内在负荷	-0.68***	-0.71***	-0.58***	-0.63***
外在负荷	-0.52***	-0.48***	-0.49***	-0.44***
相关负荷	0.74***	0.69***	0.61***	0.78***
总体负荷	-0.41***	-0.38***	-0.36***	-0.42***

注: *** $p < 0.001$ 。

内在负荷与翻译质量呈显著负相关($r = -0.68, p < 0.001$)，表明材料复杂性是制约质量的重要因素。外在负荷同样与质量负相关($r = -0.52, p < 0.001$)，说明不当的认知干扰损害学习效果。相关负荷与质量呈显著正相关($r = 0.74, p < 0.001$)，特别是与文化适应性的相关最强($r = 0.78, p < 0.001$)，验证了深层认知加工对文化维度能力发展的关键作用。

4.2. 定性分析

4.2.1. 策略发展的三阶段模型

基于屏幕录制和访谈数据的编码分析，本研究归纳出学习者策略发展的三阶段模型。

阶段一：机械依赖期(第 1~4 周)

该阶段学习者表现出对机器译文的高度依赖，策略使用呈现明显的表层特征。屏幕录制显示，学习者平均用时仅 12.3 秒评估机译输出，修改操作主要集中于拼写纠正、词汇替换等表层调整，占总修改量的 73.5%。访谈数据揭示了这一阶段的认知特征：

- ① “我觉得机器翻译已经挺好了，主要就是改改个别词，让它读起来通顺一点。”（学生 S03，第 2 周访谈）
- ② “有些四字格我不太懂什么意思，但机器给出了翻译，我就直接用了，感觉应该没问题。”（学生 S17，第 3 周访谈）

这一阶段的认知负荷特征表现为：内在负荷虽高(7.2 分)，但学习者通过依赖机译回避了深层认知加工，相关负荷低(3.5 分)；同时，适应新的协同模式导致外在负荷快速上升至峰值(5.9 分)。译文质量较低(综合得分 3.23 分)，特别是文化适应性维度得分最低(3.02 分)，反映了机械依赖导致的文化意识缺失。

阶段二：批判反思期(第 5~10 周)

该阶段是策略发展的关键转折期，学习者开始对机译输出进行批判性评估。策略使用频次显著增加：平均每个四字格的修改操作从第 1~4 周的 3.8 次增至 7.6 次，增幅 100%。修改操作的类型也发生质的变化，深层调整(如语义重构、文化补偿、修辞优化)占比从 26.5% 提升至 58.3%。访谈数据显示，教师的策略讲解和同伴讨论促使学习者产生元认知觉醒：

- ① “老师讲了‘风雨同舟’的文化内涵后，我才意识到机器翻译‘weather storms together’缺少了‘同甘共苦’的深层意义，应该补充成‘through thick and thin’。”（学生 S12，第 6 周访谈）

②“我发现同学A总是能找到机器译文的问题，比如‘巧夺天工’翻译成‘exquisite’太简单了。我开始学着像他那样批判性地看机译。”（学生S24，第8周访谈）

这一阶段的认知负荷调适特征明显：内在负荷持续下降(从6.5降至5.8分)，反映了语义理解能力的提升；外在负荷快速下降(从5.9降至4.2分)，表明技术熟练度的提高；相关负荷显著上升(从4.2升至5.3分)，验证了深层认知投入的增加。译文质量稳步提升(综合得分从3.38分升至3.84分)，文化适应性维度的进步最为显著(从3.18分升至3.76分)。

阶段三：策略内化期(第11~16周)

该阶段学习者的策略使用趋于自动化和灵活化。修改操作频次呈现倒U型曲线的下降段，从第10周的峰值7.6次降至第16周的5.2次，但修改效率显著提高。访谈显示，学习者形成了系统化的评估框架：

①“现在我看到四字格机译，会自动从三个角度评估：语义是否完整、文化信息是否传达、表达是否自然。这个过程变得很快，几乎是本能反应。”（学生S08，第14周访谈）

②“我不再纠结于每个词的修改，而是整体评估译文是否达到交际目的。有时候机译虽然不完美，但能接受就不改了，把精力放在真正有问题的地方。”（学生S29，第15周访谈）

这一阶段的认知负荷达到最优平衡：内在负荷降至低位(4.9分)，外在负荷最小化(3.2分)，相关负荷保持高位(6.1分)，反映了学习者将认知资源高效配置于促进理解的深层加工。译文质量达到较高水平(综合得分4.34分)，三个维度均衡发展。

4.2.2. 典型案例的微观分析

表5为本研究选取的典型案例，展示不同阶段学习者的策略使用特征。

Table 5. Three-stage comparison of student S21's approach to “开拓进取” translation

表5. 学生S21处理“开拓进取”的三阶段对比数

阶段	机译输出	学生译文	修改类型	修改时长	质量得分
第2周	pioneer and forge ahead	pioneer and forge ahead	无修改	8秒	3.0
第7周	pioneer and forge ahead	blaze new trails and advance with determination	深层重构	125秒	4.2
第14周	pioneer and forge ahead	break new ground and forge ahead	精准优化	52秒	4.5

案例分析显示，学生S21在第2周完全接受机译，未进行任何修改。第7周时，在批判反思期的影响下，学生进行了大幅度重构，用“blaze new trails”替换“pioneer”以强化“开拓”的动态意象，用“advance with determination”替换“forge ahead”以增加情感强度。但过度修改导致时间成本高(125秒)，且表达略显冗长。到第14周，学生达到策略内化水平，能够在保留机译合理框架的基础上进行精准优化，用“break new ground”替换“pioneer”既保持了简洁性又提升了表现力，修改效率显著提高(52秒)，质量进一步提升(4.5分)。

4.2.3. 元认知水平的中介作用

表6元认知水平问卷数据显示，学习者的元认知能力在16周内显著提升。

元认知总分从第1周的2.98分提升至第16周的4.17分，增幅39.9% ($F(2, 70) = 52.34, p < 0.001$)。进一步分析发现，元认知水平在认知负荷优化中起中介作用。将学习者按第8周元认知得分分为高分组(>4.0分，n=12)和低分组(<3.5分，n=13)，追踪两组的策略内化进程。

Table 6. Changes in metacognitive levels ($M \pm SD$)
表 6. 元认知水平的变化($M \pm SD$)

测试时间	计划能力	监控能力	评估能力	总体元认知
第 1 周	3.12 ± 0.68	2.95 ± 0.71	2.87 ± 0.74	2.98 ± 0.59
第 8 周	3.86 ± 0.62	3.74 ± 0.65	3.68 ± 0.67	3.76 ± 0.54
第 16 周	4.28 ± 0.58	4.15 ± 0.61	4.09 ± 0.63	4.17 ± 0.52

高元认知组在第 10 周即达到策略内化标准(修改操作频次开始下降但质量持续提升)，而低元认知组直到第 13 周才达到相同标准，两组差距约 3 周。访谈揭示了差异机制：

- ① “每次翻译完，我会反思哪些策略有效、哪些无效，下次遇到类似情况就知道怎么做了。”(高元认知组学生 S05)
- ② “我经常感觉自己在原地打转，每次遇到问题都像第一次碰到，没有形成固定的应对方法。”(低元认知组学生 S19)

这表明元认知能力强的学习者更善于从经验中提取规律、建构策略图式，从而加速策略内化进程，优化认知负荷分配。

5. 发现与讨论

5.1. 人机协同模式对认知负荷结构的重塑

本研究的核心发现在于揭示了人机协同模式对学习者认知负荷结构的系统性重塑。不同于传统翻译教学中认知负荷的相对稳定分布，人机协同环境下三类认知负荷呈现明显的动态演进特征。

内在负荷的持续下降(降幅 31.9%)验证了协同模式的核心优势。机器译文作为“认知脚手架”，降低了材料本身的复杂性感知。在传统模式下，学习者面对“风雨同舟”等文化负载高的四字格时，需要独立完成语义解构、文化阐释、对应选择等多重任务，工作记忆承受巨大压力。而在协同模式下，机译提供了初步对应(如“stand together through storms”)，为学习者建立了思考起点，将认知资源从“从零生成”转向“评估优化”，显著降低了任务复杂度。

外在负荷的倒 U 型演进(先升 22.9% 后降 46.6%)揭示了技术介入的双刃剑效应。初期的负荷升高源于学习者需要适应新的工作流程，即如何操作翻译系统、如何阅读机译输出、如何决策修改策略等。这些“技术学习成本”构成了额外的外在负荷，在第 4 周达到峰值。但随着技术熟练度提升和决策模式的程序化，外在负荷快速下降，到第 16 周甚至低于传统模式的基线水平。这一发现对教学设计具有重要启示：教师需要在初期提供充足的技术支持和操作培训，帮助学习者快速度过“适应期”，避免技术障碍阻碍学习进程。

相关负荷的持续上升(增幅 74.3%)是最积极的信号，表明学习者将人机协同所释放的认知资源有效转化为促进理解的深层加工，实现了认知资源的优化再分配。在策略内化期，学习者不再将精力消耗于形式转换的繁琐操作，而是聚焦于文化内涵的深度挖掘、语用效果的精准判断、修辞手段的创新运用。相关负荷的增加意味着图式建构、知识整合等高层次认知活动的强化，这不仅显著提升了译文的交际适应性与文化适应性，更从认知机制层面印证了翻译能力发展的内在核心，为人机协同教学模式的有效性提供了关键证据。

从认知负荷理论的视角看，人机协同模式实现了认知资源的优化再分配：通过降低内在负荷和外在负荷，为相关负荷的增加腾出空间，使学习者能够进行更深层次的认知加工。这种结构性重塑是传统教

学模式难以达成的，体现了技术赋能教育的本质价值。需要指出的是，本研究发现的认知负荷优化效果是在特定条件下(汉语块状语英译任务、16周干预周期、特定机器翻译系统)观察到的。内在负荷降低31.9%这一数据不应被简单解读为因果关系的直接证据，因为学习效应、任务熟悉度提升等因素也可能产生影响。此外，不同机器翻译系统的质量差异可能对认知负荷结构产生不同影响，本研究基于DeepSeek系统的发现需要在更多技术平台上得到验证。

5.2. 策略发展的非线性动态过程

本研究揭示的“机械依赖-批判反思-策略内化”三阶段模型，为理解人机协同环境下的策略发展提供了新视角。不同于传统模式下策略从无到有的线性累积，协同模式下的策略发展呈现明显的非线性特征。

机械依赖期普遍存在的“认知惰性”现象是人机协同教学面临的首要挑战。学习者倾向于将机译视为“正确答案”，回避深层思考，这与最小努力原则相符。在翻译学习过程中，批判性评估机译输出需要高强度认知投入，而直接接受机译则能最大化节省认知资源、减轻思维负担。但这种“省力”策略实质上阻碍了有效策略图式的主动建构，导致翻译能力发展陷入停滞状态。

批判反思期的策略频次激增(增幅100%)和操作类型的质变，标志着学习者元认知的觉醒。这一转折得益于两方面因素：一是教师的策略讲解提供了认知工具，使学习者具备了评估机译的分析框架；二是同伴互动创造了社会学习情境，“看到”他人的批判性思维激发了模仿动机。当学习者观察到同伴对机译进行有效改进并获得积极反馈时，会增强自我效能感，尝试相似的策略。

策略内化期的修改频次下降但质量提升，体现了自动化处理的典型特征。根据Anderson的ACT理论，技能习得从陈述性知识向程序性知识的深度转化后，执行速度大幅加快、认知负荷明显降低。学习者不再需要有意识地提取和应用策略规则，而是能够“直觉性”地识别机译问题并快速修正。这种自动化使认知资源得以释放，用于更高层次的创造性加工。

值得关注的是，策略发展呈现显著的个体差异。元认知水平高的学习者提前2~3周完成内化，说明元认知是加速发展的关键变量。在翻译学习中，元认知能力强的学习者能够准确评估自身的策略效能，及时调整无效策略，从成功经验中提取可迁移规律。这种“学会学习”的能力使其在相同学习经历中获得更高的成长效率。然而，本研究的三阶段模型是基于特定学习群体(本科二年级翻译专业学生)建立的，对于不同认知水平、学习背景的学习者是否具有普适性尚需进一步检验。此外，由于缺乏对照组设计，无法完全排除自然成熟效应的影响。策略使用频次的倒U型曲线与翻译质量提升之间虽呈现相关关系，但从相关到因果还需要更严格的实验设计来验证。

5.3. 教学干预的阶段性适配原则

本研究的教学实践验证了基于认知支架的渐进式干预的有效性，同时也揭示了教学设计需要遵循阶段性适配原则。

针对机械依赖期，教学干预的重点是“减少依赖、提升意识”。通过展示机译的典型错误(特别是文化负载表达的误译)，帮助学习者认识到机器的局限性。本研究在第3周组织的“机译错误分析”活动收效显著：学生在找出20个四字格机译问题后，对机器的盲目信任明显降低。这一干预降低了学习者的认知惰性，为批判反思期的到来奠定了基础。

针对批判反思期，教学重点转向“提供工具、引导实践”。策略讲解不应是抽象的规则灌输，而应结合具体案例展示“如何识别问题-如何分析原因-如何选择策略-如何实施修改”的完整流程。本研究采用的“教师示范→学生练习→同伴互评→反思讨论”循环模式，有效促进了策略从外部指导向内部图式的转化。需要注意的是，这一阶段不宜过度干预，过多的规则可能增加外在负荷，反而阻碍学习者自主探索适合自己的策略组合。

针对策略内化期，教学重点是“减少支架、促进自主”。随着学习者能力提升，教师应逐步撤除显性指导，鼓励学习者独立应对复杂任务。本研究在第12周后减少了策略讲解频次，增加了开放性翻译任务，给予学习者更多自主决策空间。延迟测试的结果(质量保持在高位)证实，策略已成功内化为学习者的稳定能力，不再依赖外部支持。

这一阶段性适配原则与Vygotsky的最近发展区理论高度契合。有效教学应在学习者的现有水平与潜在水平之间架设支架，支架的高度和密度需要随学习者能力提升而动态调整。过早撤除支架导致学习者“够不着”目标，过晚撤除则阻碍自主能力发展。教师需要敏锐捕捉学习者的发展阶段标志(如策略使用频次的拐点)，及时调整教学策略。

5.4. 对翻译教学改革的启示

本研究的发现对高校翻译教学改革具有多重启示：首先，应重构翻译能力培养目标，将“人机协同能力”纳入核心培养目标，包括技术应用能力、批判性评估能力、策略性干预能力和元认知调控能力。其次，教学设计应以认知负荷优化为核心原则，通过清晰的操作指导、合理的任务节奏减少外在负荷，通过问题驱动、案例分析提升相关负荷。第三，实施分阶段差异化教学，根据学习者所处发展阶段提供针对性支持。第四，强化元认知能力培养，通过反思日志、自我评估、同伴互评等方式培养学习者的元认知策略。

6. 结论与展望

本研究初步发现，在人机协同环境下，学习者认知负荷呈现出阶段性动态演变的特征，为人机协同模式对翻译学习“认知生态”的影响提供了初步证据。技术介入将学习者从低层次的形式性操作中解放，使认知资源逐步由材料难度与界面适应转向图式建构与文化判断。这一迁移以元认知调控为枢纽，并可通过过程性信号(如操作频率的拐点、回视与决策时长的下降)加以识别。由此可见，翻译教学优化的核心不在“一味降负”，而在“精准配负、促成深加工”，以阶段性支架与同伴反馈共同驱动能力生长，并为文化维度的译介质量提升提供稳定支撑。

本研究存在以下局限性：其一，样本量相对较小($n = 36$)且来源单一(仅限某高校翻译专业学生)，限制了研究结论的外部效度和推广性；其二，缺乏对照组设计(如传统教学模式组、不同机器翻译系统对比组)，难以完全排除成熟效应和其他混淆变量的影响；其三，本研究仅使用DeepSeek单一机器翻译系统，不同翻译系统的质量差异可能对认知负荷产生不同影响；其四，研究聚焦于汉语块状语这一特定翻译任务类型，对其他文本类型(如文学翻译、科技翻译)的适用性有待验证；最后，16周的追踪周期虽能观察中期发展轨迹，但对于策略内化的长期保持效果和职业迁移能力仍需延续性研究。

我们期待今后的相关研究可进一步在以下三方面推进：一是开展多中心、跨年级与跨任务的证据累积，引入神经与过程挖掘等多模态指标，刻画“负荷-策略-质量”的因果链路；二是建设可调式协同平台，依据实时学习者画像动态调节支架密度与任务颗粒度，系统比较LLM共写、记忆库融合与机译后编辑等不同协同形态的教学差异；三是将“人机素养”纳入师生能力框架，兼顾伦理、公平与可迁移性，通过毕业后追踪检验其对职业表现与持续学习的长期效应。

基金项目

本文系浙江省教育厅科研项目“文化翻译视域下新闻话语对外传播效能的优化策略研究”(项目编号：Y202455035)的后期研究成果；浙江省教育厅科研项目“大语言模型对中国形象类块状语英译能力评测与优化研究”(项目编号：Y202558054)的阶段性研究成果。

参考文献

- [1] Sweller, J. (1988) Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, **12**, 257-285. https://doi.org/10.1207/s15516709cog1202_4
- [2] 苏雯超, 李德凤, 曹洪文. 论口译认知负荷的眼动研究[J]. 外语学刊, 2021, 45(3): 109-114.
- [3] 翟秋兰. 大语言模型(LLM)驱动笔译教学: 基于认知负荷理论的文化因素提示词设计研究[J]. 人文与社科亚太学刊, 2025, 2(1): 104-113.
- [4] Trojsczak, M. (2022) Translator Training Meets Machine Translation—Selected Challenges. In: Lewandowska-Tomaszczyk, B. and Trojsczak, M., Eds., *Language Use, Education, and Professional Contexts*, Springer International Publishing, 179-192. https://doi.org/10.1007/978-3-030-96095-7_11
- [5] 郑颖, 朱玉彬, 彭昌. 人机协同翻译模式在景点介绍英译中的应用研究[J]. 巢湖学院学报, 2025, 27(4): 104-113.
- [6] 杨艳霞, 刘润泽, 陈莹. 人机协同时代翻译学习者机器翻译素养量表的编制研究[J]. 外语界, 2025, 46(5): 77-85.
- [7] Chesterman, A. (2016) Memes of Translation: The Spread of Ideas in Translation Theory. John Benjamins, 89-112. <https://doi.org/10.1075/btl.123>
- [8] 鲁华山, 裴涛, 申云化. 英语学习者元认知意识与二语听力成绩的关系: 自我效能感的中介作用[J]. 外语导刊, 2025, 48(1): 80-90, 159-160.
- [9] 吕倩兮, 姜兆坤. 人机协同反馈对学习者投入与修订质量的影响——基于文学翻译译后编辑的实证研究[J]. 当代外语研究, 2025, 45(5): 156-169.
- [10] 游凤艳. 美学视角下《2020 年国务院政府工作报告》中四字格的翻译研究[J]. 佳木斯职业学院学报, 2022, 38(2): 74-76.
- [11] 李小芳. 跨文化语境下神经网络机器翻译的不足——以《围城》中的文化负载词为例[J]. 语言文字教学与研究, 2025, 2(1): 61-63.
- [12] 赵振强, 刘绍龙. 面向 AI 机译的时政块状话语对应研究[J]. 教育研究前沿(中英文版), 2024, 8(4): 73-83.
- [13] Sweller, J. (2011) Cognitive Load Theory. *Psychology of Learning and Motivation*, **55**, 37-76. <https://doi.org/10.1016/B978-0-12-387691-1.00002-8>
- [14] Anderson, J.R. (1982) Acquisition of Cognitive Skill. *Psychological Review*, **89**, 369-406. <https://doi.org/10.1037/0033-295X.89.4.369>
- [15] Vygotsky, L.S. (1978) *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, 79-91.
- [16] Paas, F.G. (1992) Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. *Journal of Educational Psychology*, **84**, 429-434. <https://doi.org/10.1037/0022-0663.84.4.429>
- [17] Schraw, G. and Dennison, R.S. (1994) Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, **19**, 460-475. <https://doi.org/10.1006/ceps.1994.1033>