

中国高中生计算思维量表研发与应用

周培杰¹, 徐海霞¹, 胡建栗², 刘艳平², 陈文琼², 周思思²

¹湖北师范大学人工智能与计算机学院, 湖北 黄石

²黄石市第三中学, 湖北 黄石

收稿日期: 2026年3月11日; 录用日期: 2026年4月10日; 发布日期: 2026年4月22日

摘要

本研究基于社会认知理论, 整合计算思维的认知操作与自我效能感双维结构, 开发了一套适用于我国高中生的本土化计算思维量表。通过文献提取、专家效度检验与项目分析, 构建包含分解、抽象、算法、评估、概括等5个认知操作维度和创造力、问题解决、批判性思维、合作等4个自我效能感维度的九因子量表, 共45个题项。经探索性因子分析与验证性因子分析, 一阶九因子模型拟合优异, 各维度信度良好; 二阶模型分析表明, 认知操作与自我效能感两构念高度相关($r = 0.788$)但测量属性存在差异, 不宜强制合并为单一因子。多变量方差分析显示, 高中生计算思维水平不受性别与年级主效应影响, 区域主效应边缘显著($p = 0.025$)但效应量极小; 编程经验对创造力维度具有促进作用($p = 0.003$), 对其他维度无显著影响。

关键词

计算思维, 量表开发, 自我效能感, 探索性因子分析, 验证性因子分析

Development and Application of a Computational Thinking Scale for High School Students in China

Peijie Zhou¹, Haixia Xu¹, Jianli Hu², Yanping Liu², Wenqiong Chen², Sisi Zhou²

¹School of Artificial Intelligence and Computer Science, Hubei Normal University, Huangshi Hubei

²Huangshi Third Middle School, Huangshi Hubei

Received: March 11, 2026; accepted: April 10, 2026; published: April 22, 2026

Abstract

Based on social cognitive theory, this study integrated the cognitive operations and self-efficacy dual-

文章引用: 周培杰, 徐海霞, 胡建栗, 刘艳平, 陈文琼, 周思思. 中国高中生计算思维量表研发与应用[J]. 教育进展, 2026, 16(4): 1041-1051. DOI: 10.12677/ae.2026.164749

dimensional structure of computational thinking to develop a localized computational thinking scale suitable for Chinese high school students. Through literature extraction, expert validity testing, and item analysis, a nine-factor scale was constructed, comprising 45 items across five cognitive operation dimensions (decomposition, abstraction, algorithm, evaluation, and generalization) and four self-efficacy dimensions (creativity, problem-solving, critical thinking, and collaboration). Exploratory factor analysis and confirmatory factor analysis demonstrated excellent fit for the first-order nine-factor model, with good reliability across all dimensions. Second-order model analysis indicated that the two constructs of cognitive operations and self-efficacy were highly correlated ($r = 0.788$) but exhibited distinct measurement properties, suggesting they should not be forcibly merged into a single factor. Multivariate analysis of variance revealed that high school students' computational thinking levels were not affected by main effects of gender or grade; a marginally significant regional main effect was observed ($p = 0.025$) but with negligible effect size. Programming experience showed a facilitative effect on the creativity dimension ($p = 0.003$) but no significant impact on other dimensions.

Keywords

Computational Thinking, Scale Development, Self-Efficacy, Exploratory Factor Analysis, Confirmatory Factor Analysis

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

当前,我们正处于信息技术深刻重塑的时代,计算思维(Computational Thinking, CT)已与阅读、写作和算术一样,是每一位现代公民所必须掌握的基本能力。我国普通高中信息技术课程标准也将计算思维列为核心素养之一,并且计算思维是信息技术学科核心素养培养的重要议题[1]。

然而,计算思维作为一种内隐的、复杂的思维技能,其培养效果的评估面临着显著的现实困境。对计算思维的直接测量往往需借助多种方法,如基于编程任务的项目成果分析、基于系统环境的过程性数据采集、基于标准试题的认知能力测试,以及基于量表的主观认知与自我效能评估[2]。这些方法虽各有侧重,但普遍存在实施成本高、耗时长、对硬件与软件环境依赖强等问题,且容易受任务情境、评分者一致性、学生动机等外部因素干扰,导致传统方法难以支撑大规模、高效度的计算思维评估常态化实施。在此背景下,开发一种易于实施、信效度良好且适用于大规模群体测量的标准化评估工具,成为推动普及与深化计算思维培养的关键。

2. 文献综述

计算思维最早由周以真提出,是指用计算机科学的基础概念去求解问题、设计系统和理解人类的行为,涵盖了一系列反映计算机科学之广泛性的思维工具[3]。然而,由于其内隐性与复杂性,如何有效评估成为教育研究与实践的关键挑战。

社会认知理论(Social Cognitive Theory)为理解及评估 CT 提供了重要的理论视角。Bandura 指出,自我效能感(self-efficacy)——即“个体对自身能否组织和执行特定任务所需能力的信念”,是预测个体行为选择、努力程度和坚持性的关键因素[4]。

已有研究证实,自我效能感常被作为衡量计算思维的一个方面。因为编程活动会影响学生的自我效

能感,自我效能感也会对计算思维产生积极影响[5][6]。因此,通过测量个体对自身运用计算思维相关技能的信念和信心(即计算思维自我效能感),能够为我们间接评估和预测其计算思维认知操作的发展水平提供一个有效且高效的前置性诊断指标[7]。

现有计算思维量表存在两种不同的建构取向:

能力信念导向: Korkmaz 等人(2017)开发的量表[8],以 ISTE 框架(创造力、算法思维、合作、批判性思维、问题解决)为理论依据[9]。其题项设计(如“我喜欢那些对自己大多数决定有把握的人”)更偏向于个体对自身计算思维能力的信念(即自我效能感),而非直接评估认知操作。此类量表更贴近“效能预期”,反映的是主观能力感知。Kukul(2019)在 Korkmaz 的 CTS 量表基础上,开发了 CTSES 计算思维自我效能感量表[10],进一步细化了能力信念导向的测量维度。

认知操作导向: Tsai 等人(2021)开发的量表,基于 Selby 的框架[11](分解、抽象、算法、评估、概括),聚焦计算思维的五大认知操作能力[12]。其题项(如“我通常会思考一个问题是否可以被分解”)则直接指向具体认知行为,更符合计算思维的本质界定。

我国学者在引进国际量表的同时,也积极开展了本土化探索。张屹等人开发了适用于我国小学生群体的计算思维量表,题项设计充分考虑了中国学生的学习情境与表达习惯[13]。白雪梅等人则基于国外成熟量表,采用探索性因子分析和验证性因子分析对国外已有的计算思维量表的信效度进行研究[14]。这些研究提示,其测量需充分考虑文化适应性。例如,中国教育情境中的“合作”维度往往更强调集体主义文化下的协同行为与共享责任,这与西方量表个体主义取向下的合作表述可能存在差异。因此,量表的本土化并非简单翻译,而需在语义表达、情境设置及价值导向上进行重构,以更准确反映中国学生的思维特质与行为表现。

综上所述,整合认知操作与自我效能感的双重取向,并深入进行文化适应性的本土重构,是开发适用于我国高中生的计算思维量表的关键路径。本研究拟在此基础上,构建一个既符合认知本质、又契合我国高中教育实情的多维测量工具,为计算思维的大规模评估与教学改进提供实证依据。

3. 研究方法

3.1. 研究对象

本研究采用多阶段分层整群抽样法,选取华中地区湖北省黄石市、华南地区广东省江门市作为抽样地理区域(两地分别代表中部地级市、珠三角沿海地级市,兼顾中部教育均衡试点区域与华南教育发达区域特征,经济、教育发展水平具有区域代表性),严格按照区域类型、办学性质、办学层次分层抽取样本学校。

首先以“城市/县城/农村”为区域分层依据,结合省示范、市示范、普通高中的办学层次差异,每分层维度各抽取 1~2 所学校,最终选取 6 所高中,包含 4 所地级市市区学校(广东省 2 所、湖北省 2 所)、1 所县城学校(湖北省)和 1 所农村学校(广东省)。在确定样本学校后,采取整群抽样的方式选取学生,即每所学校内,以自然班为单位,随机抽取高一至高三各一个班级的全部学生作为调查对象。这种抽样方式保证了样本的随机性,并便于问卷的集中发放与回收。

本样本的城乡分布与我国 2024 年普通高中教育统计年鉴中对应省份的城乡学生比例进行对比,发现本样本的城市学生比例略高,而农村学生比例略低,在县城学生比例上基本持平。这表明样本在一定程度上对城市学生略有侧重,但其整体结构与区域总体情况仍具有较高的可比性。

3.2. 量表初始设计

为全面解析计算思维的发展路径,规避“能力-信心偏差”,本研究对现有国际主流量表进行系统

性整合与本土化重构，量表设计流程分为文献提取、语义聚类和本土化调整三个步骤。

Table 1. Formal scale of computational thinking
表 1. 计算思维正式量表

维度	题项
分解：将问题分解为可管理的小部分来解决的倾向，包括分析结构和整理要素。	<p>我通常会思考一个问题的整体结构。</p> <p>我通常会思考如何将一个大问题拆分成几个小问题。</p> <p>我能列出问题中给定的条件和所需的结果。</p> <p>我会分析复杂系统中各个部分之间的相互作用和依赖关系。</p> <p>我能通过画图来整理解决问题的思路。</p>
抽象：聚焦解决问题的关键信息的倾向，包括提取本质、识别模式和关注细节。	<p>我通常会尝试找出问题的核心要点。</p> <p>我通常会尝试分析不同问题之间的共同模式。</p> <p>在处理难题时，我会首先弄清问题的根本原因。</p> <p>我能清晰区分问题中的关键信息和可以忽略的次要细节。</p>
算法：规划问题解决方案的倾向，包括对数学过程/算法建模的兴趣和执行细节。	<p>我会为复杂任务设计清晰有序的执行步骤。</p> <p>我可以轻松发现图形之间的关系。</p> <p>我能将语言描述的问题转化为数字形式。</p> <p>我善于发现任务中的重复模式，并想办法简化或自动化流程。</p> <p>设计方案时，我会考虑不同情况并制定对应的步骤(例如：如果 A 发生，就执行 B)。</p> <p>我能用结构化的方式(如列表、流程图)阐述问题及其解决过程。</p>
评估：考虑可用资源以找到问题最佳解决方案的倾向，包括比较选项和规划效率。	<p>我会设定明确的标准来比较和判断不同解决方案的优劣。</p> <p>我致力于寻找当前情况下最有效、最可行的解决方案，而不一定是完美的。</p> <p>在选择方案时，我会综合考虑时间、成本和效果等多方面因素。</p> <p>我能够客观评估一个解决方案的有效性和潜在缺陷。</p> <p>在比较不同选项做决定时，我会用系统的方法来分析。</p>
概括：识别特定问题解决方案的模式并将其应用于类似问题的倾向，包括经验迁移。	<p>我能将从一个领域学到的知识或技能，转化应用到另一个不同的领域中去。</p> <p>解决一个问题后，我会总结其中的方法，看看能否应用到其他场合。</p> <p>我致力于寻找能解决一整类问题的通用策略或模式。</p> <p>当我在某个问题上陷入困境时，我会尝试把以前的解决方案应用到新的情况中。</p>
创造力：相信自己能从多角度生成创新想法和线索的信心。	<p>我会从不同角度看待问题，以产生多种想法。</p> <p>遇到新问题时，我总能想出好的想法。</p> <p>当我在某个问题陷入困境时，我会从周围环境中寻找解决线索。</p> <p>在着手解决问题时，我相信自己对“正确”和“错误”的直觉。</p> <p>在遇到新情况时，我能灵活应对处理。</p> <p>当遇到问题时，我会先停下来想清楚，然后再行动。</p>
问题解决：相信自己能想象、解释和解决问题的信心。	<p>我能轻松在脑海中可视化问题的解决方案。</p> <p>在制定解决问题的计划时，我相信自己能执行它。</p> <p>即使面对陌生难题，我也相信自己最终能攻克它。</p>

续表

	我相信如果投入足够时间和努力,我能解决大多数问题。 我能清楚地解释我的问题解决方案。
批判性思维:对自己能否成功运用批判性思维完成特定任务的自信判断。	我相信自己能透过复杂现象,发现问题的根本原因。 要反对他人的意见,必须提出合理的理由。 我乐于学习具有挑战性的事物。 在分析问题时,我相信自己能够保持客观和中立,不轻易受个人偏见或情绪影响。 我相信我能够进行精确和逻辑严密的思考。 我相信自己的判断力,对自己做出的决定很有把握。
合作:在合作环境中产生想法和感到舒适的信心与倾向。	我喜欢和同学组成小组,围绕同一个问题或项目学习。 我相信团队合作能比独自学习更快更好地解决问题。 和大家一起学习时,我能提出更多有价值的想法。 在合作学习中,和小组伙伴一起学习让我感到轻松。

首先,从现有量表中提取初始题项池,结合陈兴治[15]和 Bers [16]的理论框架生成备选项,共搜集 71 个题项。然后使用 NVivo 软件进行内容分析,计算语义相似度(阈值 > 0.75,基于词向量模型),优先保留本土化表述(如张屹量表中适应中国学生情境的题项),筛选后得到 59 个题目。

其次,进行专家效度验证:邀请 5 位专家(3 名信息技术课程教师与 2 位教育学教授)采用 Delphi 法两轮评审。第一轮内容效度比率(CVR)计算结果显示,59 个题项中,CVR > 0.8 的题项有 50 个,保留率为 85%;第二轮焦点小组讨论进一步调整文化适应性。最终形成初始量表共 50 题,采用五点制测量。

为确保量表的结构效度和内部一致性,我们进行了实证校验。首先,对初始量表进行预测试(N = 362)。整体克隆巴赫 α 系数为 0.898,表明初步内部一致性良好;KMO 值为 0.872,巴特利特球形度检验 $p < 0.001$,显示数据适合因子分析。随后,采用探索性因子分析(EFA,主成分分析结合正交旋转)和信度分析,对量表的结构效度和内部一致性进行检验。EFA 基于以下标准进行项目净化:因子载荷阈值设为 0.40(低于此值的项视为对维度贡献不足);交叉载荷阈值设为 0.30(高于此值的项视为维度混淆);此外,考虑项与维度核心构念的理论一致性。

不符合上述标准的题项直接删除。对于因子载荷尚可但语义存在瑕疵、或对维度内涵贡献独特的较差题项,则进行了针对性语义优化与表述精炼后予以保留。预试后量表修订为 48 题。

为避免引入新偏差,净化后进行二次试测(N = 346),克隆巴赫系数为 0.967,EFA 结果显示 KMO = 0.959, Bartlett $p < 0.001$,确认结构稳定。试测后进一步删除 3 题,最终形成正式量表 45 题,具体见表 1。

4. 量表正式发放

量表正式发放后回收得到 847 份有效数据,按奇偶序号分为两个子样本:子样本 A (n = 424)用于探索性因子分析,子样本 B (n = 423)用于验证性因子分析。两组样本在性别、年级、区域分布上无显著差异($p > 0.05$),具有可比性。

4.1. 探索性因子分析

首先对数据进行适用性检验,结果显示 KMO 值为 0.968, Bartlett 球形检验的近似卡方值为 36418.675 (自由度 = 990, $p < 0.001$),达到极其显著的水平。这两项指标共同表明,样本数据非常适合进行因子分

析, 变量间存在共享的潜在结构, 能够提取出有意义的公因子。总量表的克隆巴赫系数为 0.967, 表明整个量表具有极高的内部一致性信度。

然后采用主成分分析法并配合最大方差法进行正交旋转, 以特征值大于 1 为标准提取公因子。各因子的特征值、方差解释率及累积解释率见表 2。分析结果共提取出 9 个公因子, 与量表理论构建的维度完全吻合。累积方差解释率为 81.229%, 意味着这 9 个因子共同解释了所有题项大部分的信息变异。

Table 2. Cumulative variance explained

表 2. 累积方差解释率

成分	总计	方差百分比	累积百分比
1	4.993	11.096	11.096
2	4.966	11.035	22.132
3	4.962	11.028	33.160
4	4.114	9.143	42.303
5	4.082	9.072	51.374
6	4.041	8.979	60.353
7	3.162	7.027	67.381
8	3.143	6.985	74.366
9	3.088	6.863	81.229

旋转后的成分矩阵显示, 所有题项均清晰地负载于其预设的理论维度上, 且因子负荷量均较高(>0.70)。不存在显著的交叉负荷现象(即没有一个题项在两个及以上因子上的负荷同时大于 0.4), 九个维度之间具有良好的区分度。

在提取出清晰的九因子结构后, 我们进一步计算了每个因子的组合信度(CR)与平均方差抽取量(AVE), 以检验各维度的内部一致性与收敛效率(见表 3)。

Table 3. AVE and CR of each component

表 3. 各成分 AVE 与 CR

潜变量	题数	AVE	CR
分解	5	0.649	0.891
抽象	4	0.625	0.870
算法	6	0.651	0.918
评估	5	0.636	0.897
概括	4	0.611	0.863
创造力	6	0.676	0.926
问题解决	5	0.679	0.913
批判性思维	6	0.680	0.927
合作	4	0.645	0.879

结果表明, 所有九个因子的平均方差抽取量(AVE)值均大于 0.6, 均远超 0.50 的门槛值; 组合信度(CR)值均大于 0.8, 远高于 0.70 的临界标准。这充分说明, 各题项能有效汇聚于其所要测量的潜变量, 具

有出色的收敛效率。

4.2. 验证性因子分析

为全面检验量表的结构效度,本研究依次构建了三种竞争性模型并进行比较:(1)一阶九因子基准模型(M1);(2)二阶双因子模型(M2,包含认知操作CT1与自我效能感CT2两个相关二阶因子);(3)二阶单因子模型(M3,所有维度归属一个高阶因子)。各模型的拟合指数、因子负荷及维度间相关分别汇总于表4。

Table 4. Summary of Model Fit Indices

表 4. 各模型拟合指数汇总

模型	χ^2/df (CMIN/DF)	SRMR	RMSEA	CFI	TLI (NNFI)	NFI	PNFI
M1	1.009	0.0171	0.003	1.000	1.000	0.975	0.895
M2	1.527	0.133	0.025	0.986	0.986	0.961	0.915
M3	2.085	0.192	0.036	0.972	0.97	0.947	0.903

综合比较, M1 (一阶九因子模型)拟合最优,支持量表九维度的独立性; M2 虽可接受,但认知操作维度的二阶聚合存在问题; M3 拟合较差,不支持计算思维作为单一高阶因子的理论构想。这一结果提示:认知操作与自我效能感是两个相关但不同质的构念,不宜简单合并,建议在实际应用中分别使用两个子量表或采用九维度计分方式。

5. 结果与讨论

5.1. 研究结果

为考察学校所在区域、高中年级、有无编程经验及性别对计算思维的影响,本研究采用多变量方差分析进行分析。基于 847 名高中生的有效数据,其样本特征分布如表 5 所示。多元检验结果显示(表 6),学校所在区域主效应显著,但年级、有无编程经验及性别主效应均不显著。多元检验中,仅区域与年级的交互效应显著。

Table 5. Distribution of sample characteristics

表 5. 样本特征分布情况

特征	类别	人数	百分比
性别	男	439	51.80%
	女	408	48.20%
有无编程经验	有	314	37.07%
	无	533	62.93%
年级	高一	340	40.10%
	高二	314	37.10%
	高三	193	22.80%
学校所在区域	城市	479	56.55%
	县城	259	30.58%
	农村	109	12.87%

单变量检验进一步显示,区域与编程经验、年级与性别的交互效应在特定维度显著,详见表7。交互效应方面,区域与年级的交互效应在多元检验中即显著($p=0.024$),并在抽象、算法、概括三个维度表现出单变量显著性,提示这一交互效应具有较强的稳健性。相比之下,区域与编程经验、年级与性别的交互效应虽在单变量检验中显著,但未通过多元检验,提示这些效应可能局限于特定维度,整体解释力相对有限。

从效应量来看,所有显著效应的偏 η^2 均小于0.02,属于极小效应。模型整体解释力有限,各维度 R^2 介于0.029~0.047之间,调整后 R^2 多为负值,表明人口特征变量对计算思维变异的解释力极其微弱。

Table 6. Multivariate tests (Overall effect)

表 6. 多变量检验(整体效应)

效应	Wilks' lambda	F	df	误差 df	p
截距	0.074	1008.637	10	802	<0.001
学校所在区域	0.959	1.714	20	1604	0.025*
年级	0.978	0.882	20	1604	0.612
有无编程经验	0.982	1.481	10	802	0.141
性别	0.993	0.545	10	802	0.859
区域 × 年级	0.929	1.492	40	3042.946	0.024*
区域 × 编程经验	0.975	1.036	20	1604	0.415
年级 × 性别	0.976	0.965	20	1604	0.503
其余交互	-	-	-	-	>0.05

Table 7. Tests of between-subjects effects (Significant results only)

表 7. 主体间效应检验(仅呈现显著结果)

源	因变量	F	p	偏 η^2
学校所在区域	抽象	3.849	0.022	0.009
年级	抽象	3.078	0.047	0.008
有无编程经验	创造力	8.634	0.003	0.011
区域 × 年级	抽象	3.563	0.007	0.017
	算法	2.416	0.047	0.012
	概括	2.649	0.032	0.013
区域 × 编程经验	算法	3.609	0.028	0.009
	分解	3.302	0.037	0.008
年级 × 性别	评估	3.501	0.031	0.009
	问题解决	3.471	0.032	0.009

5.2. 研究讨论

本研究没有拘泥于国际现有的单一取向,而是整合“认知操作”与“自我效能感”双维框架。经实证检验证实该模型兼具理论合理性与测量有效性。CFA结果显示,认知操作与自我效能感两大构念高度相

关($r=0.788$), 但又在测量上彼此区分。这一方面说明, 计算思维技能与个体对自身能力的信心紧密关联, 掌握技能是建立信心的基础; 另一方面, SRMR 值偏高恰恰从实证角度反映了这两个构念在测量属性上的内在差异: 认知操作题项更偏向对客观行为的频率评估, 而自我效能感题项则更侧重于主观信念的判断, 二者反应机制的不同导致了残差结构的复杂化。

5.2.1. 人口学变量主效应

尽管多变量检验显示, 性别、年级的主效应在统计上不显著, 但这并不能等同于现实中的“无差异”或“已实现完全公平”。统计上的不显著可能受到样本量、效应量大小、测量工具的敏感性以及组间分布等多种因素影响。例如, 年级主效应不显著, 可能暗示着高中阶段常规的学科教学并未在整体上显著促进计算思维各维度的同步发展。这更可能反映了当前课程整合的现状, 即计算思维的培养可能尚未系统性地嵌入各年级的常规教学, 其发展更多依赖于学生的自发探索或特定课程的影响, 而非年级的自然进阶。同样, 性别主效应不显著, 可能部分得益于近年来基础教育阶段在鼓励女生参与 STEAM 活动方面的努力, 但也不应忽视在更深层次的认知风格、职业选择上可能存在的性别差异。因此, 这些“不显著”的发现, 其更重要的启示在于提示我们关注过程性变量(如教学方式、学习动机、学习风格)而非结构性变量(如年级、性别)对计算思维发展的潜在影响。

5.2.2. 显著交互效应

区域与年级的交互效应在抽象、算法、概括三个维度上显著, 表明不同区域(城市、县城、农村)的学生在计算思维发展轨迹上存在差异。这可能源于不同区域教育资源配置与教学质量的差异。例如, 城市学校可能更早、更系统地引入项目式学习或 STEAM 教育, 使得学生在高年级时, 高阶的抽象、算法思维发展更快。而县城或农村学校, 其课程可能更侧重于基础知识的传授, 计算思维的培养更多依赖于教师的个人探索, 导致学生相关能力的发展曲线更为平缓或滞后。这种发展轨迹上的分化, 是教育区域不平衡的深层体现, 值得政策制定者和教育研究者关注。

区域与编程经验的交互效应在算法维度显著, 提示编程训练的效果并非均等, 而是受到地区教育背景的调节。在资源丰富、师资力量强的城市学校, 编程经验可能通过更高质量的教学转化为更强的算法思维能力。然而, 在资源相对匮乏的地区, 即使学生有编程经验, 也可能因为教学指导不足或实践机会有限, 使其对算法思维的促进作用大打折扣。因此, 在推广编程教育时, 不能仅仅关注硬件投入, 更需关注课程质量和师资培训的配套支持。

年级与性别的交互效应在分解、评估及问题解决维度显著, 暗示高中阶段计算思维的发展可能存在性别差异的年级轨迹。例如, 可能在高一时, 男女学生在分解复杂问题上的能力并无显著差异, 但随着年级升高, 由于社会期望、兴趣分流或课程选择(文理分科)的影响, 男生在这些维度上的增长速率可能加快, 从而在高年级产生差异。这提示我们, 在高中阶段应持续关注并鼓励女生参与逻辑推理、问题拆解类的思维活动, 通过创设包容性的课堂环境, 避免潜在的性别刻板印象固化。

6. 研究不足与展望

综上所述, 本研究揭示了计算思维发展的“稳定性”与“差异性”并存的复杂图景。主效应的不显著提示我们, 计算思维作为核心素养, 其发展具有一定的普遍性; 而显著的交互效应则警示我们, 这种发展的具体路径和速率, 深受个体所处的区域教育环境、学习经历以及年级发展阶段等多重因素的交互影响。未来研究应超越对单一人口学变量的考察, 转而关注这些变量如何通过复杂的交互作用, 塑造个体的计算思维发展轨迹, 并探索能够有效促进教育公平、因材施教的教学干预模式。

尽管各个模型在绝大多数拟合指标上均达到优异标准, 但其标准化残差均方根(SRMR)值持续偏高,

这是同时测量两种不同心理构念的预期结果。“认知操作”(CT1: 分解、抽象、算法、评估、概括)旨在评估个体从事计算思维相关认知活动的行为与能力倾向,其题项表述更贴近对客观行为的描述;而“自我效能感”(CT2: 创造力、问题解决、批判性思维、合作)则旨在测量个体对自身运用这些技能完成特定任务的信心与信念水平,其题项更具主观评估色彩。这两种构念在反应模式上存在系统性差异:认知操作的作答更多基于对过往行为的回溯,而效能感的作答则受到个体自信心、情绪状态等内在因素的影响更大。这种测量属性的根本差异,导致两个子模块的题目残差(误差)呈现出不同的协方差模式,从而使得当它们被置于同一个测量模型时,其整体的残差矩阵与模型预测矩阵存在较大差异,直接体现为 SRMR 值升高。

当分别对 CT1 和 CT2 进行单独建模时,两个子模型的 SRMR 均得到改善(CT1: 0.0137; CT2: 0.0166),且均展现出清晰、稳定的因子结构。而将 CT1 与 CT2 作为两个相关但独立的潜变量构建二阶模型时,两者呈现出高度正相关($r=0.788$)。这清晰地表明:计算思维认知操作与计算思维自我效能感是两个高度共生、但概念上可区分、测量上不同质的构念。在实际应用层面,这一发现提示研究者:在使用本量表时,既可将 CT1 与 CT2 视为一个整体来评估个体的计算思维总体水平,也应重视两个子量表的独立价值,分别考查学生在认知操作与自我效能感上的具体表现,从而进行更具针对性的教学干预与培养。

基金项目

2025 年湖北师范大学研究生科研创新项目“解构·重构·创构:基于项目式学习的计算思维三维 5C 培养模式研究”(编号: 2025Z022)。

参考文献

- [1] 张立国, 王国华. 计算思维: 信息技术学科核心素养培养的核心议题[J]. 电化教育研究, 2018, 39(5): 115-121.
- [2] 惠恭健, 兰小芳, 钱逸舟. 计算思维该如何评?——基于国内外 14 种评价工具的比较分析[J]. 远程教育杂志, 2020, 38(4): 84-94.
- [3] Wing, J.M. (2006) Computational Thinking. *Communications of the ACM*, **49**, 33-35. <https://doi.org/10.1145/1118178.1118215>
- [4] Bandura, A. (1986) *Social Foundations of Thought and Action*. Prentice-Hall, Inc.
- [5] Tolba, E.G. (2025) Computational Thinking Skills and Computational Self-Efficacy as Predictors of Academic Well-being. *International Journal of Technology in Education*, **8**, 761-788. <https://doi.org/10.46328/ijte.1180>
- [6] Román-González, M., Pérez-González, J., Moreno-León, J. and Robles, G. (2018) Extending the Nomological Network of Computational Thinking with Non-Cognitive Factors. *Computers in Human Behavior*, **80**, 441-459. <https://doi.org/10.1016/j.chb.2017.09.030>
- [7] Ramalingam, V., LaBelle, D. and Wiedenbeck, S. (2004) Self-Efficacy and Mental Models in Learning to Program. *Proceedings of the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, Leeds, 28-30 June 2004, 171-175. <https://doi.org/10.1145/1007996.1008042>
- [8] Korkmaz, Ö., Çakir, R. and Özden, M.Y. (2017) A Validity and Reliability Study of the Computational Thinking Scales (CTS). *Computers in Human Behavior*, **72**, 558-569. <https://doi.org/10.1016/j.chb.2017.01.005>
- [9] ISTE & CSTA (2011) Operational Definition of Computational Thinking for K-12 Education. https://cdn.iste.org/www-root/Computational_Thinking_Operational_Definition_ISTE.pdf
- [10] Kukul, V. and Karatas, S. (2019) Computational Thinking Self-Efficacy Scale: Development, Validity and Reliability. *Informatics in Education*, **18**, 151-164. <https://doi.org/10.15388/infedu.2019.07>
- [11] Selby, C. and Woollard, J. (2013) Computational Thinking: The Developing Definition. https://eprints.soton.ac.uk/356481/1/Selby_Woollard_bg_soton_eprints.pdf
- [12] Tsai, M., Liang, J. and Hsu, C. (2021) The Computational Thinking Scale for Computer Literacy Education. *Journal of Educational Computing Research*, **59**, 579-602. <https://doi.org/10.1177/0735633120972356>
- [13] 张屹, 莫尉, 张岩, 等. 我国小学生计算思维量表研发与应用[J]. 中国电化教育, 2020(10): 49-57.
- [14] 白雪梅, 顾小清. K12 阶段学生计算思维评价工具构建与应用[J]. 中国电化教育, 2019(10): 83-90.

-
- [15] 陈兴治, 马颖莹. 本土化计算思维评价指标体系的构建与探索——基于 1410 名高中生的样本分析与验证[J]. 远程教育杂志, 2020, 38(5): 70-80.
- [16] Bers, M.U. (2010) The TangibleK Robotics Program: Applied Computational Thinking for Young Children. *Early Childhood Research & Practice*, **12**, 1-20.