

人工智能替代焦虑及其影响因素： 基于中国社会状况综合调查的研究

徐超凡¹, 李正风^{2,3}

(1. 中国科学院科技战略咨询研究院, 北京 100190;

2. 清华大学社会科学学院, 北京 100084;

3. 清华大学科学技术与社会研究中心, 北京 100084)

摘要: 人工智能技术的快速普及在深刻重塑劳动力市场的同时,也引发了劳动者深层的“替代焦虑”,探索该焦虑的生成机制对数字时代的社会风险治理具有重要意义。本文基于中国社会状况综合调查数据,综合采用多元线性回归(OLS)与机器学习模型,考察人工智能替代焦虑的影响因素及其作用机制。研究发现:第一,我国劳动者的人工智能替代焦虑总体处于中低水平且呈现显著的结构分化,平台工作者与体制外劳动者的焦虑水平更高;第二,OLS回归揭示了整体层面的普遍性保护机制,体制内就业和较高的社会经济地位能显著降低个体的替代焦虑;第三,针对高焦虑群体的机器学习预测进一步发现,个人收入与受教育年限也是识别高焦虑个体的核心特征,表明基础经济资源与人力资本的匮乏是触发深度技术焦虑的关键。交互效应分析显示,个体户口和教育的协同、收入与阶层地位的拮抗联合驱动替代焦虑的生成。研究结果深化了对数字时代技术风险不平等分布的经验理解,为缓解公众就业焦虑及制定包容性公共政策提供了经验依据。

关键词: 人工智能;替代焦虑;技术风险;影响因素;机器学习

中图分类号: TP18;C912.6 **文献标识码:** A **文章编号:** 1005-0566(2026)09-0083-11

Artificial intelligence replacement anxiety and its determinants: Evidence from the Chinese social survey

XU Chaofan¹, LI Zhengfeng^{2,3}

(1. *Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China;*

2. *School of Social Sciences, Tsinghua University, Beijing 100084, China;*

3. *Center for Science, Technology and Society, Tsinghua University, Beijing 100084, China)*

Abstract: The rapid diffusion of artificial intelligence (AI) is profoundly reshaping the labor market while triggering deep-seated “replacement anxiety” among workers. Exploring the mechanisms underlying this anxiety is of great significance for social risk governance in the digital era. Based on data from the Chinese Social Survey (CSS), this study combines ordinary least squares (OLS) regression with machine learning models to examine the determinants of AI replacement anxiety and its underlying mechanisms. The findings reveal the following: First, the overall level of AI

基金项目: 国家自然科学基金项目“主要国家科研资助机构改革动向与资助布局研究”(L2524003);中国科协创新战略研究院科研项目“中国科技工作者发展权益的社会学理论问题研究”(2026-RCS-1);国家社会科学基金重大项目“深入推进科技体制改革与完善国家治理体系研究”(21ZDA017)。

作者简介: 徐超凡(1998—),男,江苏南通人,中国科学院科技战略咨询研究院助理研究员,研究方向为科学社会学、科学技术与社会。通信作者:李正风。

replacement anxiety among Chinese workers is at a low-to-moderate level but exhibits significant structural differentiation, with platform workers and non-state-sector workers reporting higher levels of anxiety. Second, OLS regression identifies a general protective mechanism, demonstrating that state-sector employment and higher socioeconomic status significantly mitigate individuals' replacement anxiety. Third, machine learning predictions focused on the high-anxiety cohort further indicate that personal income and years of education are core features for identifying highly anxious individuals, suggesting that a lack of basic economic resources and human capital is a key trigger for severe technological anxiety. Interaction analysis further shows that hukou status and education exert a synergistic effect, while income and class status have an antagonistic effect, jointly shaping the formation of replacement anxiety. This study deepens the empirical understanding of the unequal distribution of technological risks in the digital age, providing evidence for alleviating public employment anxiety and formulating inclusive public policies.

Key words: artificial intelligence; replacement anxiety; technological risk; determinants; machine learning

人类历史上的每次重大技术变迁都不可避免地重塑着社会的生产方式与劳动分工。当前,人工智能加速嵌入社会生产过程,在客观层面替代部分常规任务、引发岗位重组与技能再配置^[1-4],也在个体层面形成对未来职业风险的感知^[5-6]。与以往主要替代体力劳动的技术工具不同,当前的人工智能标志着技术对人类劳动的替代正式跨越了生理感官,从物理层面的“肢体”延伸至认知层面的“大脑”^[7]。这种深层次的认知性替代不仅触及了人类智力劳动的核心领域,也极大加剧了劳动者对自身职业前景与不可替代性的深层担忧。社会情绪并非孤立的个人体验,当其跨越阶层群体引发共鸣时,便会进一步影响公众的社会认同与行为取向^[8]。从这一意义上说,“人工智能替代焦虑”作为一种先于真实岗位变动而发生的社会情绪,已成为理解劳动者如何回应技术变革,以及社会如何治理数字转型风险的核心切入点。

尽管学界已初步证实人工智能焦虑是一个可测量的心理现象^[9],发现其会对劳动者的社会经济地位感知^[10]、地位认同^[11]及心理健康^[12]产生负面效应,且在不同社会经济位置人群中存在明显差异^[13],但关于该情绪生成逻辑的系统性经验解释仍存在局限。首先,现有文献多关注公众对人工智能的总体态度或宏观风险感知^[14-15],或主要测算人工智能对职业可替代性的客观概率^[16-18],却缺乏基于全国性劳动者样本,直接追问“哪些具体特征的劳动者更容易产生主观替代焦虑”。其次,在方法论层面,既有影响因素分析多依赖传统线性回归框架^[19]。该方法虽有助于识别单一变量的方向

与显著性,但面对多维资源禀赋、数字技能与复杂就业形态交织而成的真实社会网络,传统模型难以刻画变量间的非线性关联,更无法有效识别和比较多重因素在触发焦虑时的相对重要性^[20]。

围绕上述背景与方法论关切,本文使用 2023 年中国社会状况综合调查数据,聚焦 18~65 岁在业劳动者,重点考察以下三个核心研究问题:第一,当前中国劳动者的人工智能替代焦虑总体处于何种水平,并呈现出怎样的分布特征?第二,哪些社会人口学、社会经济、健康、数字技能与就业特征因素与人工智能替代焦虑存在显著的经验关联?第三,在传统回归分析之外,利用机器学习方法进一步识别关键影响因素的相对重要性及其交互效应。基于此,全文余下结构安排如下:第一部分为文献综述与研究问题的提出,第二部分介绍数据来源、变量设置与研究方法,第三部分报告实证结果,第四部分总结与讨论主要发现并给出政策启示。

一、文献评述与问题提出

(一)人工智能替代焦虑的研究现状

人工智能替代焦虑是人工智能扩散背景下劳动者对自身工作可能被智能技术替代的主观担忧。围绕这一现象,现有研究首先完成了概念识别和测量建构。有研究从一般性的人工智能焦虑入手,强调人工智能焦虑并不是单一情绪,而是包含学习焦虑、控制感下降、职业替代担忧等多个维度的综合性心理反应^[21-22]。还有研究则直接面向工作场景,将替代焦虑界定为员工对未来岗位或任务被智能技术接管的担忧,并将其与工作不安全感、职业压力等概念加以区分^[23]。整体来看,

学界已经逐渐形成一个基本共识,即人工智能替代焦虑并非泛泛的技术态度,而是一种具有明确职业指向的主观风险感知。

在此基础上,研究重心进一步转向人工智能替代焦虑的后果。现有研究表明,这种焦虑会显著影响劳动者的工作体验^[23-24]、地位认同^[10-11]和健康状态^[12,25]。也就是说,人工智能替代焦虑不只是对技术变迁的即时反应,它还会沿着主观态度和客观健康两个方向持续外溢,进而影响劳动者的职业判断和行为选择。这一研究进展使人工智能替代焦虑逐渐从技术接受研究中的附属变量,转变为理解劳动者主观职业风险的重要切口。

不过,现有研究大多以后果为中心来设计问题,把人工智能替代焦虑作为解释其他现象的前提。研究逻辑通常默认替代焦虑已经存在,再考察它会带来什么结果。与之相比,什么样的劳动者更容易形成这种焦虑,哪些条件会触发或强化这种焦虑,目前仍缺乏同样充分的研究积累,这也构成了现有研究在问题设置上的明显空缺。

(二)人工智能替代焦虑影响因素的研究现状

进一步追问人工智能替代焦虑的生成机制,现有研究已识别出若干关键影响因素,例如熟悉度、信任、隐私担忧、知识储备和政治社会信任,都会影响人工智能风险感知^[14,26]。但这些因素如何共同作用于替代焦虑的形成,仍缺乏系统性的整合;换言之,学界已识别出一批影响人工智能风险感知的重要变量,却尚未形成关于替代焦虑影响因素的统一解释框架。

其次,现有研究已经从心理特质、认知机制和技术属性三个层面展开了较为具体的识别。一类研究强调个体差异的重要性,认为具体焦虑体验对人工智能消极态度的解释力高于一般人格特质^[27]。另一类研究强调认知和判断机制,指出技术发展速度带来的不确定感、对技术透明度和可解释性的怀疑,以及对人工智能过度依赖所造成的信任悖论,都会强化替代担忧^[28-29]。还有研究从任务和技术特征切入,发现拟人化、个性化、任务相似性和技术威胁感会显著提高工作替代焦虑,而批判性思维、可解释性和技术掌控感则具有

一定缓解作用^[28,30]。这些研究共同说明,替代焦虑的形成并不是单一因素驱动的,而是个体特质、认知机制和技术情境共同作用的结果。

教育、知识和能力变量在近年来的影响因素研究中越来越重要。已有研究发现,知识和信心会共同塑造公众对人工智能的风险与收益判断^[26]。人工智能素养和技术自我效能不仅可以降低焦虑,也会影响个体对人工智能风险的理解方式。在职业场景中,人工智能素养越高,劳动者越可能将人工智能理解为一种可学习、可利用的工具,而不是完全外在的替代威胁^[31]。这说明,替代焦虑并不只是由技术冲击本身决定,也与劳动者是否具有理解和应对技术变化的能力密切相关。

但是,从现有证据的分布看,影响因素研究仍然呈现出明显的样本和场景局限。大量研究建立在特定群体之上,如办公室职员、服务业员工、大学生、护士等,研究视角则多聚焦于技术属性、任务类型和组织场景^[5,28,30-31]。这些研究为识别人工智能替代焦虑的形成条件提供了重要线索,但它们能否推广到更广泛的在业劳动者群体,仍然需要进一步检验。特别是在不同制度环境和劳动市场结构下,劳动者对人工智能替代的主观判断很可能并不完全遵循同一逻辑。

(三)评述与研究问题

综合已有研究可以看到,人工智能替代焦虑研究已经证明了这种焦虑可以作为一种主观职业风险感知被识别和测量,也积累了它影响工作态度、心理状态和社会心态的证据。相比之下,人工智能替代焦虑如何形成的研究仍明显滞后。现有证据多建立在特定群体之上,研究视角也集中于技术属性、任务类型和组织场景。国内研究主要讨论替代焦虑的后果,对其来源、触发条件和群体分布特征关注不足。基于上述文献脉络,本文将人工智能替代焦虑界定为劳动者对未来一段时期内自身现有工作被机器人、人工智能和自动化设备替代可能性的主观判断。据此,本文关注中国在业劳动者的替代焦虑水平及其群体差异,考察社会人口学、社会经济、健康、数字技能与就业特征的影响,并进一步比较机器学习方法识别出的

关键因素及其交互关系。

二、研究设计

(一)数据来源与研究样本

本文使用 2023 年中国社会状况综合调查(CSS 2023)数据^①。CSS 2023 由中国社会科学院社会学研究所组织实施,采用全国范围概率抽样,调查内容覆盖就业、收入、互联网使用、健康状况及社会态度等多个方面,能够为人工智能替代焦虑的影响因素分析提供较为丰富的微观信息。本文将研究对象限定为 18~65 岁、当前有工作且回答“未来 5 年,您目前这份工作被机器人、人工智能、自动化设备替代的可能性?”题项的受访者。剔除因变量无效值后,得到 5 729 个研究样本;在纳入全部解释变量并剔除缺失值后,实证分析部分均基于 5 285 个有效研究样本展开。

(二)变量定义

被解释变量:人工智能替代焦虑。该变量来自 CSS 2023 关于“未来 5 年当前工作被机器人、人工智能、自动化设备替代可能性”的题项,参照赵青矣等^[10]的研究,将其命名为“人工智能替代焦虑”。为便于解释,本文将原始题项反向编码为 1~5 分,数值越大表示替代焦虑越强。在机器学习分类任务中,进一步将原始题项中“完全有可能”和“有可能”界定为高人工智能替代焦虑,其余为低人工智能替代焦虑。需要说明的是,这里的“焦虑”并非临床诊断意义上的焦虑障碍,而是劳动者面对技术替代风险时形成的预期性不安。工作不安全感研究强调,此类不安源于个体对就业连续性和稳定性受到威胁的主观感知^[32];风险感知研究也指出,公众对技术风险的反应取决于其对潜在危害及可控性的评估^[33]。因此,本文所称“人工智能替代焦虑”旨在捕捉劳动者对未来技术替代可能性的感知性工作不安全感,而非临床心理症状。

解释变量:主要分为四类。第一类为社会人口学特征,包括年龄、性别、户口性质、受教育年限和婚姻状况。第二类为社会经济与健康特征,包

括个人收入、身体健康、心理健康和当前社会经济地位。第三类为数字能力特征,包括数字技能和互联网使用时长。第四类为工作特征,包括单位属性、平台工作、岗位技能要求和专业职称或技术等级。各变量的具体题项来源与赋值方式见表 1。

表 1 变量的描述性统计结果(N=5 285)

变量类别	变量名称	具体题项与赋值	均值	标准差
被解释变量	人工智能替代焦虑	基于“未来 5 年,当前工作被机器人、人工智能、自动化设备替代的可能性”题项反向编码,1~5 分,数值越大表示焦虑越强	2.307	1.282
	机器学习被解释变量	机器学习使用的二分类变量。原始题项取值为“完全有可能”或“有可能”记为 1,其余记为 0	0.252	0.434
社会人口学特征变量	年龄	18~65 岁	42.080	10.880
	性别	女性=0;男性=1	0.541	0.498
	户口性质	农业户口或登记在村委会的居民户口=0;非农业户口或登记在居委会的居民户口=1	0.355	0.479
	受教育年限	依据最高受教育程度折算为受教育年限,0~19 年	11.414	3.845
	婚姻状况	未婚、离婚、丧偶=0;已婚或同居=1	0.797	0.402
社会经济与健康特征变量	当前社会经济地位	当前社会经济地位自评反向编码,取值越大表示地位越高,1~5 分	2.254	0.891
	个人收入	2022 年个人总收入加 1 后取自然对数	10.356	2.184
	身体健康	身体健康自评,1~10 分,分值越高表示身体健康越好	7.555	1.866
	心理健康	心理健康自评,1~10 分,分值越高表示心理健康越好	8.187	1.849
数字能力特征变量	数字技能	依据信息获取、工作事业、社会交往、经济活动和生活服务五项互联网活动频率反向编码后取均值,0~5 分	2.407	1.269
	互联网使用时长	每日互联网使用时长,0~4 分	1.577	1.011
工作特征变量	单位属性	体制内(党政机关、国有企业及国有控股企业、国有或集体事业单位)=1;体制外(其余)=0	0.207	0.405
	平台工作	当前工作必须通过互联网平台完成=1;否则=0	0.110	0.313
	岗位技能要求	岗位技能要求反向编码,取值越大表示胜任该岗位的技能要求越高,1~4 分	1.948	0.908
	专业职称	有专业职称或技术等级=1;否则=0	0.235	0.424

(三)研究方法

在传统计量分析部分,本文首先采用分组 t 检验对样本中的人工智能替代焦虑分布特征进行描述与比较。在推断性分析部分,采用控制省份固定效应的多元线性回归(OLS)模型以识别影响人工智能替代焦虑的相关因素,模型为:

$$AIAnxiety_i = \alpha + \beta X_i + \mu_p + \varepsilon_i$$

其中,被解释变量 $AIAnxiety_i$ 表示第 i 个受访者的人工智能替代焦虑水平; X_i 为一组解释变量向量; α 为常数项; β 为待估计的偏回归系数向量; μ_p 为省份固定效应; ε_i 为随机误差项。此外,本文在所有回归中均采用省份层面的聚类稳健标准误。

在机器学习分析阶段,如表 1 所示,本文构建“人工智能替代焦虑”的二分类变量,旨在进一步

① 本论文使用数据全部来自中国社科院重大项目“2023 年中国社会状况综合调查”。该调查由中国社会科学院社会学研究所执行,感谢提供数据协助。

挖掘并识别关键影响因素的非线性特征。研究分别构建了逻辑回归(Logistic Regression)、支持向量机(SVM)、随机森林(Random Forest)、梯度提升机(GBM)与极端梯度提升(XGBoost)五类候选分类模型。为评估模型泛化能力,本文采用分层5折交叉验证(Stratified 5-fold Cross-Validation),并综合考察各模型的AUC值。

在机器学习模型选择上,本文以平均AUC值最高作为最优模型的遴选标准,并基于折外预测概率(Out-of-fold Predictions)实施成对AUC差异检验,以验证不同模型间性能差异。最后,为提升模型的可解释性,本文运用SHAP(Shapley Additive Explanations)方法解释最优模型的预测结果,通过SHAP重要性条形图与蜂群图刻画各关键因素的影响权重及其作用方向,并基于TreeSHAP算法提取特征交互值、绘制变量交互网络图,补充揭示核心影响因素之间的交互效应。

三、实证结果分析

(一)人工智能替代焦虑的分布及其特征

1. 人工智能替代焦虑的总体分布与空间特征

当前全体在业劳动者的人工智能替代焦虑均值为2.307,其中高焦虑群体占比为25.22%。这表明,目前劳动者对人工智能替代效应的整体焦虑水平处于中等偏低状态,但高焦虑现象已在超过四分之一的群体中显现,具备不容忽视的现实普遍性。

从空间分布来看,人工智能替代焦虑呈现出非均衡的省际异质性。按照省级均值分位划分,排名前25%的高焦虑省份既包括西北内陆的宁夏、青海,也涵盖中东部的安徽等地;而排名后25%的低焦虑省份则散布于西藏、吉林、海南等地。这一空间格局表明,主观上的人工智能替代焦虑并未表现出简单的“东—中—西”梯度递减规律,其地域差异可能更多地嵌套于各省份具体的产业结构、劳动力市场形态以及数字化转型进程之中。

2. 人工智能替代焦虑的结构性分化

为进一步考察人工智能替代焦虑在不同社会群体间的差异,本文通过分组t检验对核心人口学特征与工作特征进行了比较(见图1)。结果显示,人工智能替代焦虑在个体的社会人口学特征与劳

动力市场位置上均存在显著的结构分化:首先,在社会人口学特征方面,女性的替代焦虑水平显著高于男性($p < 0.01$);18~39岁的青年群体显著高于40~65岁的中老年群体($p < 0.01$);大专以下学历群体的焦虑水平显著高于大专及以上群体($p < 0.01$);农村户口劳动者显著高于城镇户口劳动者($p < 0.001$)。这一结果反映出,在面对技术冲击时,女性、处于职业早期的青年群体,以及人力资本或制度保障资源相对薄弱的弱势群体,表现出了更强的敏感度与脆弱性。最后,在劳动力市场位置方面,体制外劳动者的焦虑水平显著高于体制内劳动者($p < 0.001$);而在新就业形态中,平台工作者的焦虑程度亦显著高于非平台工作者($p < 0.05$)。这表明,职业稳定性的缺失以及高度算法化的工作环境,进一步加剧了个体在面对人工智能时的替代危机感。

(二)人工智能替代焦虑的影响因素回归分析

图2报告了控制省份固定效应后的多元线性回归估计结果。首先,在社会人口学特征方面,年龄与人工智能替代焦虑呈显著负相关($\beta = -0.007, p < 0.01$),表明随着年龄增长,劳动者对当前工作被技术替代的担忧逐渐减弱;男性的焦虑水平显著低于女性($\beta = -0.083, p < 0.05$),这与前文描述性统计的分化趋势保持一致。其次,在结构位置与劳动力市场特征方面,“体制庇护”与“阶层地位”均展现出缓冲效应。具体而言,就职于体制内单位能够大幅且显著地降低个体的替代焦虑($\beta = -0.184, p < 0.001$);当前主观社会经济地位越高的个体,其焦虑水平越低($\beta = -0.099, p < 0.001$)。

最后,心理健康水平越高的个体,其应对技术冲击时的焦虑感显著更低($\beta = -0.033, p = 0.007$)。值得注意的是,在纳入全部解释变量后,受教育年限、数字技能、专业职称、个人收入对数以及城乡户口等变量,均未达到统计显著水平。这说明,在当前中国情境下,公众抵御人工智能替代风险的核心机制更多依附于体制内的制度性保障与综合社会经济地位;单纯依赖“人力资本”(如教育和技能)的提升或绝对收入的增加,尚不足以有效缓解技术变革带来的深层职业不安全感。

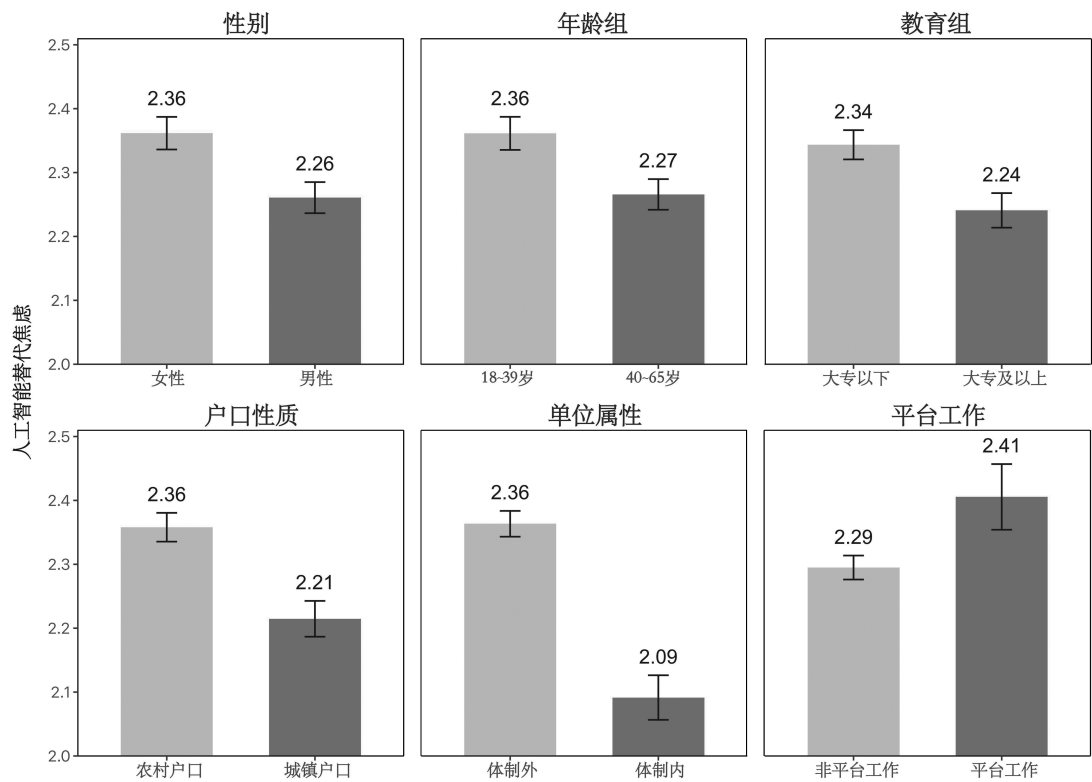


图 1 人工智能替代焦虑的分组差异

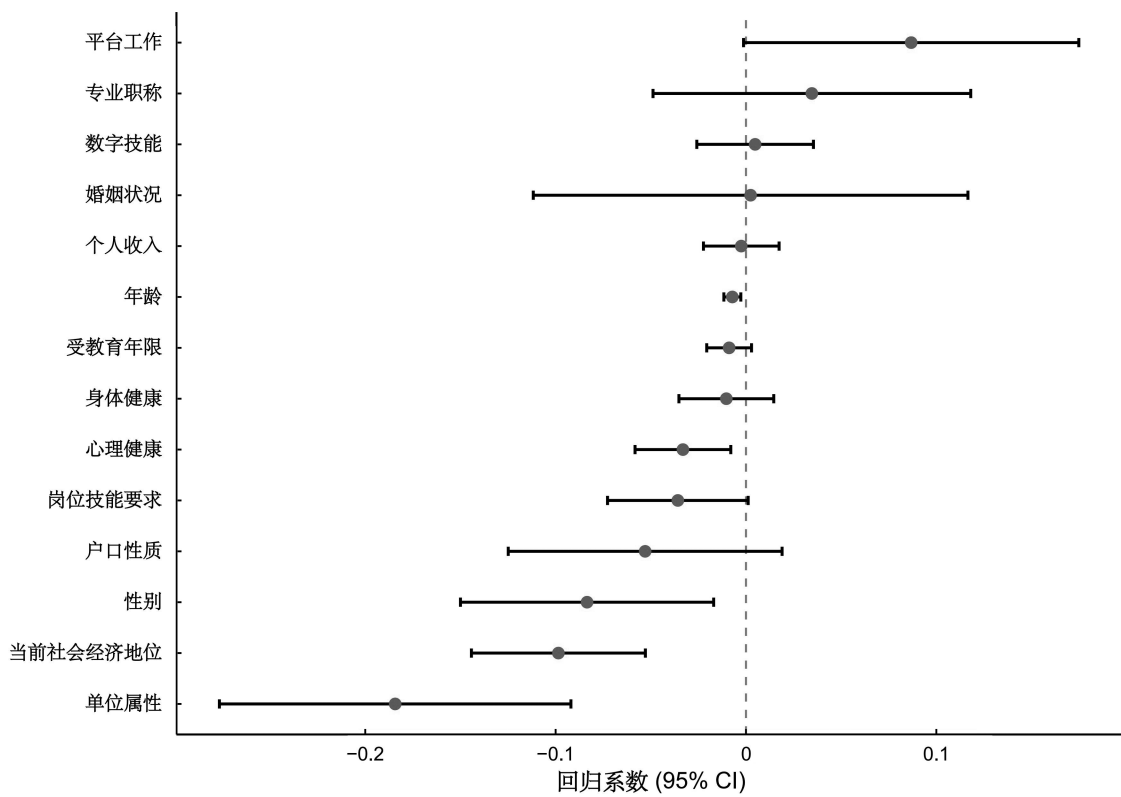


图 2 人工智能替代焦虑多元线性回归模型结果

注:圆点表示 OLS 回归系数,横线表示按省份聚类标准误计算的 95% 置信区间。置信区间不跨越 0,表示该变量在 5% 水平上显著。N = 5 285,控制省份固定效应。因变量为人工智能替代焦虑。

(三)基于机器学习模型的影响因素与非线性机制识别

为处理传统线性回归难以捕捉的非线性关系,本文进一步引入机器学习算法,以高人工智能替代焦虑为预测目标进行建模分析。表2汇报了五类候选分类模型在统一样本下的分层5折交叉验证性能。对比发现,梯度提升机模型的平均AUC值最高(0.5981),略优于极端梯度提升模型(0.5961)与逻辑回归(0.5909)。受限于社会科学问卷调查数据的固有信噪比,本研究中机器学习模型的AUC指标区分度整体有限。因此,本文应用机器学习的核心目的并非追求预测精度,而是利用其特征提取能力,对影响替代焦虑的关键因素进行相对重要性排序与非线性识别。表3给出了基于t检验的模型性能差异显著性结果,结果显示梯度提升机模型已显

表2 机器学习模型性能比较					
模型	准确率	敏感性	特异性	F1 值	AUC
梯度提升机	0.594 7	0.534 1	0.615 1	0.399 3	0.598 1
极端梯度提升	0.563 1	0.592 6	0.553 1	0.406 3	0.596 1
逻辑回归	0.601 9	0.516 9	0.630 6	0.395 7	0.590 9
随机森林	0.446 4	0.757 7	0.341 3	0.408 4	0.562 9
支持向量机	0.511 8	0.542 4	0.501 5	0.359 2	0.527 8

著优于随机森林模型($p < 0.001$)与支持向量机模型($p < 0.001$)。综合考量,本文确立梯度提升机为最优模型以进行后续的SHAP解释。

表3 不同模型 AUC 差异的成对显著性检验					
模型	梯度提升机	极端梯度提升	逻辑回归	随机森林	支持向量机
梯度提升机	—	0.492	0.193	< 0.001	< 0.001
极端梯度提升	—	—	0.414	< 0.001	< 0.001
逻辑回归	—	—	—	0.002	< 0.001
随机森林	—	—	—	—	< 0.001
支持向量机	—	—	—	—	—

图3和图4分别展示了梯度提升机模型的SHAP解释结果。特征重要性排序结果显示,影响个体高人工智能替代焦虑的关键变量依次为:个人收入、当前社会经济地位、受教育年限、年龄以及单位属性。其中,个人收入和当前社会经济地位的预测贡献最高,说明经济资源和社会位置是模型识别高人工智能替代焦虑个体时最重要的维度。

结合图4的特征分布可以直观看出:收入水平偏低、自评社会经济地位较低、受教育年限较短、年龄较小以及处于体制外就业的群体,其SHAP值多数位于零线右侧,即这些人群被预测具有更高的人工智能替代危机感。

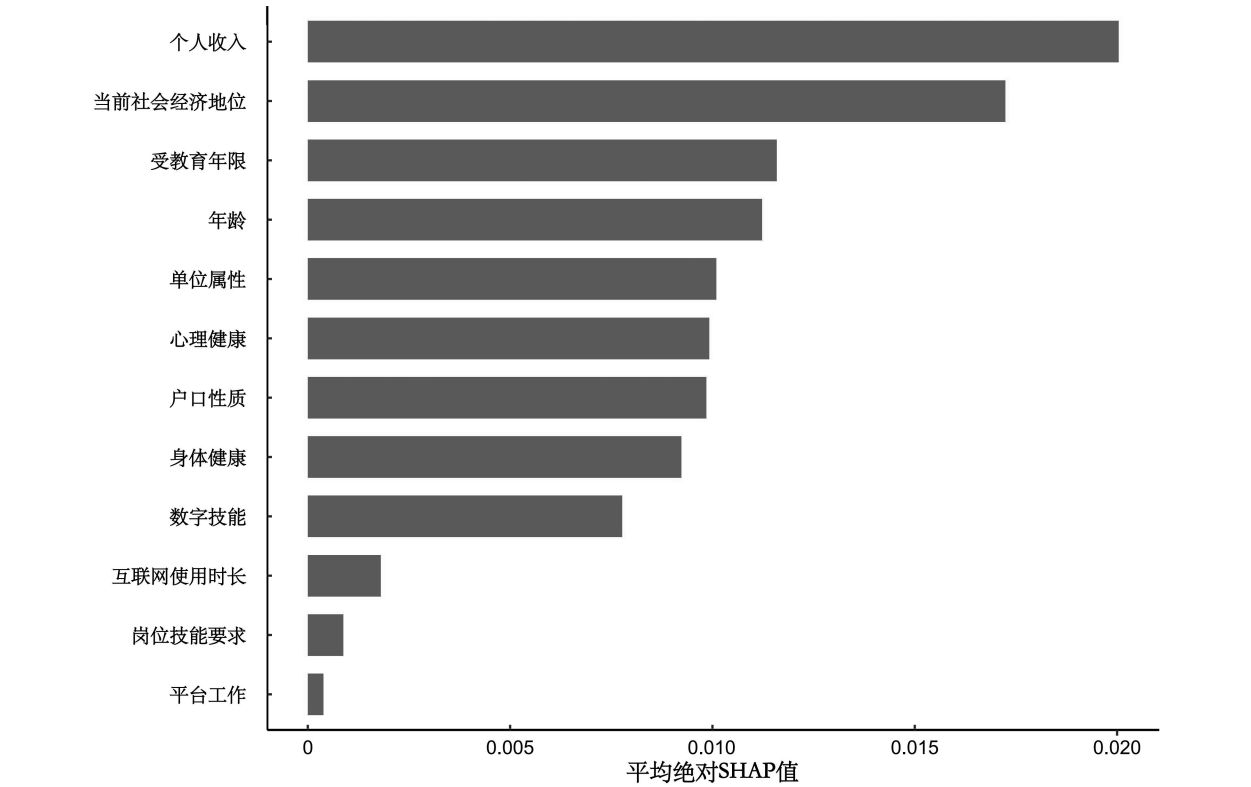


图3 高人工智能替代焦虑预测中各变量的重要性排序

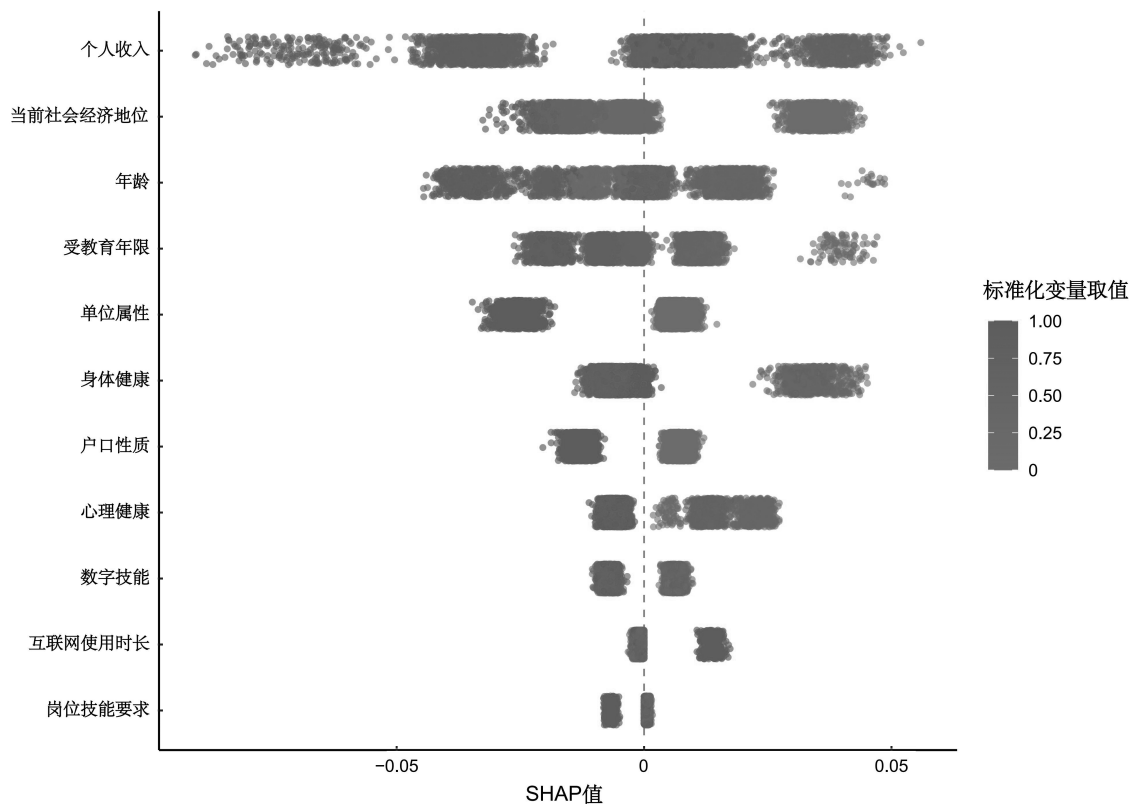


图4 高人工智能替代焦虑预测中各变量的预测方向与分布

个人收入与受教育年限在传统的多元 OLS 回归中未表现出显著线性主效应,但在机器学习模型中跃升为重要的预测特征。这揭示了经济资源与人力资本在抵御技术风险时具有复杂的非线性机制。收入和教育对焦虑的缓解作用不一定呈现全局线性递减;而机器学习的高解释权重则表明,二者可能存在“门槛效应”,也可能与体制内外等结构性位置发生交叉互动。换言之,绝对收入的匮乏与教育资源的短缺,可能在特定的脆弱群体内部急剧放大了技术替代焦虑,从而成为识别高焦虑风险人群的关键特征。

(四)人工智能替代焦虑的影响因素交互效应分析

为进一步揭示关键因素间的交互效应,本文基于 TreeSHAP 算法进行交互效应分析。图 5 展示了排名前 8 个核心变量的 SHAP 交互网络。图 5 中节点大小与颜色深浅表示变量的重要性大小,连线的粗细表示交互作用强度的相对大小,颜色表征交互方向:部分表示协同效应,即两变量的联合特征对识别高焦虑个体具有正向增益;部分代表拮

抗效应,即两变量联合时会部分抵消彼此的预测贡献。结果显示,户口性质与受教育年限之间展现出较强的协同关系;而收入与当前社会地位之间则呈现出较明显的拮抗关系。这表明,人工智能替代焦虑的生成受到复杂社会经济特征影响,并呈现结构性交互特征。结合前文分析可以发现,一些在线性主效应框架下作用不够突出的变量(如受教育年限),在与特定结构性因素结合时,可能会表现出更强的联合预测贡献。

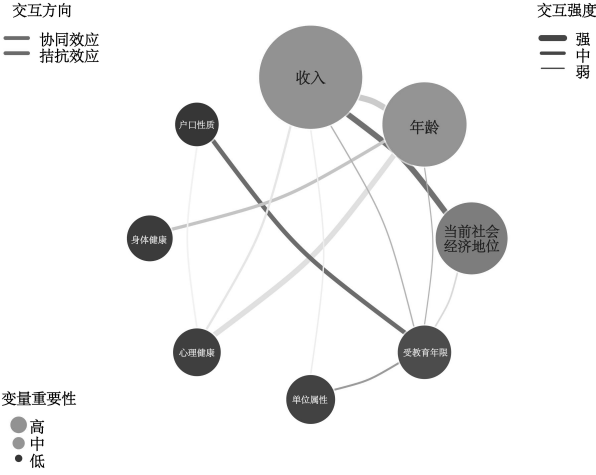


图5 关键变量的 SHAP 交互网络图

四、结论与讨论

本文基于 CSS2023 数据,结合省份固定效应回归、机器学习模型与 SHAP 解释,考察了中国在业劳动者人工智能替代焦虑的分布特征及其影响因素。研究表明,人工智能替代焦虑是一种嵌入劳动者的社会位置、就业条件和心理状态的分化性技术情绪。回归结果显示,年龄、性别、体制内外、当前社会经济地位和心理健康是影响人工智能替代焦虑的关键因素;机器学习结果进一步表明,收入、当前社会经济地位、教育、年龄和体制内外在识别高替代焦虑个体时具有更高的重要性,且这些因素存在着不同程度的交互作用。这意味着,人工智能替代焦虑的形成主要受劳动者所处的资源结构和不确定性处境影响,单一的技术接触或知识差异难以充分解释。

首先,社会经济位置和制度嵌入程度越弱,劳动者越容易产生较强的人工智能替代焦虑。当前社会经济地位越低和体制外就业,都与更高水平的替代焦虑相关。这一结果与近期关于人工智能主观风险后果的研究形成了清晰对话。已有研究发现,人工智能替代焦虑会通过失业预期削弱青年劳动者的社会经济地位感知,且这种心态剥夺效应在禀赋较弱群体中更为明显^[10];职业替代风险也会显著压低新社会阶层的地位认同,并在高自动化暴露行业中表现得更强^[11]。本文进一步把研究起点前移,说明社会位置较弱、制度保障较少和就业关系更不稳定的劳动者,更容易形成这种焦虑。人工智能替代焦虑由此表现为分层化的主观职业风险,并会继续外溢为心理健康、地位认同和社会心态问题。

其次,年龄、性别和教育差异表明,人工智能替代焦虑具有明确的群体异质性。本文发现,女性、年轻劳动者和受教育年限较低者的替代焦虑更强。这与部分近期研究的总体判断大体一致,说明人工智能焦虑的社会分布存在规律性,会受性别、教育和技术暴露经验等因素影响^[9,34]。但本文的发现也与部分国际研究存在差异。例如,基于一般性人工智能焦虑量表的研究发现,年龄差异在控制心理变量后并不显著,部分性别差异也仅存在于特定维度之中^[9]。这种差异说明,一

般性的人工智能态度或焦虑,与面向工作替代的职业性焦虑并不完全相同。对处于职业上升期、岗位积累尚不充分的劳动者而言,替代焦虑更容易与职业竞争、收入预期和家庭责任联结在一起,因此其社会分布更接近劳动市场中的脆弱性结构,而不是一般科技态度的社会分布。

另外,心理健康变量的结果表明,人工智能替代焦虑包含对技术的判断,也包含劳动者对不确定环境的情绪性回应。本文发现,心理健康水平越高,人工智能替代焦虑越低。已有研究表明,人工智能工作替代风险感知会通过工作不安全感、财务不安全感和地位不安全感损害劳动者心理健康^[12]。同时,组织层面的人工智能采纳也呈现出明显的双重效应:它既可能通过提升自我效能提升员工的工作旺盛感,也可能通过触发焦虑抑制这种兴盛,而且这种效应取决于员工更倾向于将组织引入人工智能理解为帮助人还是替代人^[35]。本文的结果进一步说明,这种关系很难被概括为单向后果链条,心理状态本身也构成人工智能替代焦虑的重要社会心理基础。正因如此,人工智能替代焦虑不宜被简单理解为需要消除的负面情绪。更准确地说,它是在技术快速扩散和规则尚未稳定的背景下,劳动者对未来风险所作出的正常评估和情绪反应。围绕科技争议和科技风险的研究也指出,公众态度并不能仅由知识水平解释,心理因素如幸福感和信任都是理解新兴技术态度的重要维度^[36]。据此,面对人工智能替代焦虑,治理方向应避免要求劳动者“不焦虑”,而要通过透明的信息沟通、可预期的制度安排和可获得的支持资源,使这种焦虑被引导为可讨论和可应对的风险意识。

在研究意义上,相比于既有研究对“后果如何”的侧重,本文将视角前移,深度剖析人工智能替代焦虑的生成机制。通过刻画高焦虑群体的社会画像,试图将人工智能替代焦虑锚定在社会分层、制度位置与心理特质的交互逻辑之中,从而建立起从结构位置到微观心理的系统解释方案。这种视角转变在实践层面同样具有警示意义:人工智能对劳动力市场的冲击呈现明显差异,技术能否实现岗位更迭只是问题的一面,替代预期在社

会阶层间的不均等分布同样值得关注。考虑到绝大多数劳动者并非身处人工智能核心产业,其对技术的风险感知往往受到其接触的社交媒体信息及叙事框架影响^[37],这种建构性的认知容易脱离算法现实而产生过度恐慌。因此,当前的治理重点应在于引导公众客观拆解“机器换人”的逻辑局限,明确简单替代在现实复杂场景中的不可能性^[38]。针对体制外劳动者、平台从业者等在资源储备与心理防御上相对脆弱的群体,应当通过提升数字素养等方式,提高其对技术变化的可理解性和可应对性,减少由不确定性累积造成的持续性焦虑^[39]。

当然,本文仍存在若干局限。第一,现有关于人工智能替代焦虑的研究,包括本文在内,主要依赖问卷调查数据;但人工智能技术本身处于快速演进之中,其应用范围、社会叙事和劳动市场影响都可能在较短时间内发生显著变化。因此,单次截面调查难以捕捉替代焦虑随技术扩散和政策调整而变化的动态过程。未来研究有必要结合追踪调查、实验研究和平台行为数据,进一步识别人工智能替代焦虑的演变机制。第二,尽管本文引入机器学习与 SHAP 解释以识别高维变量的重要性,但相关结果仍基于观测数据,主要提供关联解释和预测依据,不能据此作出严格的因果推断。第三,本文的被解释变量来自对未来五年工作被替代可能性的主观判断,测量的主要是职业风险感知,不属于临床心理学意义上的焦虑症状。一定程度上,该指标可能混淆“知道风险但不担心”与“不知道风险而不担心”两种状态。低焦虑既可能是个体在充分意识到替代风险后仍保持较低担忧,也可能源于人工智能素养不足所导致的“无知者无畏”。未来研究可在替代可能性判断之外,进一步引入人工智能素养和人工智能职业暴露等因素,以区分前述两种状态,从而更准确地解释人工智能替代焦虑的形成机制。本文的探索表明,人工智能替代焦虑很难被视为技术进步带来的无差别副产品,它带有鲜明的社会分层和结构位置特征。正视这种不安全感的现实根基,要求我们真正深入具体的制度环境与行业场景,去探寻兼顾技术创新与劳动者权益保护的良善治理之道。

参考文献:

- [1] ACEMOGLU D, LOEBBING J. Automation and polarization[J]. *Journal of political economy*, 2026, 134(3): 1017-1072.
- [2] ACEMOGLU D, RESTREPO P. Robots and jobs: evidence from US labor markets [J]. *Journal of political economy*, 2020, 128(6): 2188-2244.
- [3] FREY C B, OSBORNE M A. The future of employment: how susceptible are jobs to computerisation? [J]. *Technological forecasting and social change*, 2017, 114: 254-280.
- [4] LY B, LY R, MA S. Navigating AI-induced job displacement and skill demands: insights from an emerging market perspective[J]. *Human behavior and emerging technologies*, 2026(1): 3424335.
- [5] LIANG Y, ZHAI Y. The impact of artificial intelligence impact awareness on college students' employment risk perception: a moderated mediation model[J]. *Acta psychologica*, 2025, 261: 105808.
- [6] WANG Y, GAO Y. Exploring the relationship between AI acceptance and career decision-making anxiety among Chinese university students: a latent profile analysis[J]. *The Asia-Pacific education researcher*, 2025: 1-11.
- [7] 王硕, 张徐姗, 李正风. 多层视域下的 AI for Science [J]. *自然辩证法研究*, 2025, 41(9): 42-50.
- [8] 张彦, 李汉林. 国家治理视角下的总体性社会情绪 [J]. *社会*, 2024, 44(2): 1-35.
- [9] UĞUR N G, DURSUN F. The social anatomy of AI anxiety: gender, generations, and technological exposure[J]. *Frontiers in psychiatry*, 2025, 16: 1641546.
- [10] 赵青矣, 柳建坤. 人工智能替代焦虑、数字技能与青年劳动者的社会经济地位感知: 基于 CSS2023 调研数据 [J]. *中国青年研究*, 2025(12): 24-34.
- [11] 周骥腾. 智能时代的职业替代风险与新社会阶层的地位认同[J]. *延边大学学报(社会科学版)*, 2025, 58(6): 102-114.
- [12] 徐延辉, 刘昆智. 人工智能工作替代风险感知与劳动者的心理健康: 不确定性的作用机制研究[J]. *西北民族研究*, 2025(6): 110-123.
- [13] LIN L, PARKER K. On future AI use in workplace, US workers more worried than hopeful[EB/OL]. (2025-02-25) [2026-04-09]. <https://www.pewresearch.org/social-trends/2025/02/25/u-s-workers-are-more-worried-than-hopeful-about-future-ai-use-in-the-workplace/>.
- [14] KRIEGER J B, BOUDER F, WIBRAL M, et al. A

- systematic literature review on risk perception of artificial narrow intelligence[J]. *Journal of risk research*, 2025, 28(9-10): 1111-1129.
- [15] 李森林, 张乐. 公众对人工智能的接受度及其形成机制研究[J]. *科学学研究*, 2025, 43(12): 2484-2495.
- [16] 孙望书, 孙旭. 人工智能将会“抢走”谁的工作: 异质劳动者的就业替代风险[J]. *河北经贸大学学报*, 2024, 45(2): 66-76.
- [17] 沈可, 石笑峰, 张安妮. 人工智能发展何以赋能高质量就业[J]. *人口研究*, 2026, 50(1): 51-67.
- [18] 沈可, 石笑峰. 城市人工智能发展对个人劳动参与和工作时长的影响: 基于人工智能技术密度的发现[J]. *中国人口科学*, 2025, 39(2): 45-61.
- [19] WŁOCH R, ŚLEDZIEWSKA K, ROŻYNEK S. Who's afraid of automation? examining determinants of fear of automation in six European countries [J]. *Technology in society*, 2025, 81: 102782.
- [20] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[C]//*Advances in neural information processing systems* 30. Long Beach: Neural information processing systems foundation, 2017: 4765-4774.
- [21] WANG Y Y, WANG Y S. Development and validation of an artificial intelligence anxiety scale: an initial application in predicting motivated learning behavior[J]. *Interactive learning environments*, 2022, 30(4): 619-634.
- [22] LI J, HUANG J S. Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory [J]. *Technology in society*, 2020, 63: 101410.
- [23] BROUGHAM D, HAAR J. Smart technology, artificial intelligence, robotics, and algorithms (STARA): employees' perceptions of our future workplace [J]. *Journal of management & organization*, 2018, 24(2): 239-257.
- [24] SCHWABE H, CASTELLACCI F. Automation, workers' skills and job satisfaction [J]. *PLoS one*, 2020, 15(11): e 0242929.
- [25] BAEK S U, YOON J H, WON J U. Association between workers' anxiety over technological automation and sleep disturbance: results from a nationally representative survey [J]. *International journal of environmental research and public health*, 2022, 19(16): 10051.
- [26] SAID N, POTINTEU A E, BRICH I, et al. An artificial intelligence perspective: how knowledge and confidence shape risk and benefit perception [J]. *Computers in human behavior*, 2023, 149: 107855.
- [27] KAYA F, AYDIN F, SCHEPMAN A, et al. The roles of personality traits, AI anxiety, and demographic factors in attitudes toward artificial intelligence [J]. *International journal of human-computer interaction*, 2024, 40(2): 497-514.
- [28] JO H, PARK D H. The fear of being replaced by generative AI: an examination of influential factors among office workers [J]. *Technological forecasting and social change*, 2025, 220: 124326.
- [29] EREBAK S, TURGUT T. Anxiety about the speed of technological development: effects on job insecurity, time estimation, and automation level preference [J]. *The journal of high technology management research*, 2021, 32(2): 100419.
- [30] VOROBEOVA D, EL FASSI Y, COSTA PINTO D, et al. Thinking skills don't protect service workers from replacement by artificial intelligence [J]. *Journal of service research*, 2022, 25(4): 601-613.
- [31] LIU S, XIAO Y, NIE M, et al. Nurses' attitudes toward artificial intelligence: AI literacy as a predictor and the mediating effect of AI anxiety [J]. *BMC nursing*, 2025, 24(1): 1511.
- [32] SHOSS M K. Job insecurity: an integrative review and agenda for future research [J]. *Journal of management*, 2017, 43(6): 1911-1939.
- [33] SLOVIC P. Perception of risk [J]. *Science*, 1987, 236(4799): 280-285.
- [34] GERLICH M. Public anxieties about AI: implications for corporate strategy and societal impact [J]. *Administrative sciences*, 2024, 14(11): 288.
- [35] ZHANG Z, GAO Q. Organizational AI adoption: a bane or a boon for employee thriving at work? [J/OL]. *Asia pacific journal of management*, 2026: 1-28. <https://doi.org/10.1007/s10490-025-10104-7>.
- [36] 徐超凡, 李正风. 超越知识: 科技伦理争议态度的心理因素研究 [J]. *科学学研究*, 2025, 43(12): 2666-2675.
- [37] CUI D, WU F. The influence of media use on public perceptions of artificial intelligence in China: evidence from an online survey [J]. *Information development*, 2021, 37(1): 45-57.
- [38] 马尔登, 格雷厄姆, 坎特. 投喂 AI [M]. 贾青青, 牟一凡, 译. 北京: 中信出版社, 2025: 49-51.
- [39] ZHANG J, YAO Z, KANG D, et al. A SEM-ANN analysis to examine impact of AI overreliance through a social cognitive lens: the roles of dependency, FOMO, addiction, and anxiety [J]. *BMC psychology*, 2026, 14(1): 547.