

员工-人工智能合作 如何影响领导感知员工努力?

金杨华 朱汕怡 吴珍珍 谢江佩

摘要:在人工智能(AI)广泛应用于组织的背景下,员工与AI合作日益普遍,但对于领导如何解读员工此类行为并形成相应的管理判断尚缺乏较为系统的理论探讨和实践检验。本文通过两项研究探讨了员工-AI合作影响领导感知员工努力的机制及其管理后果。情境实验发现,员工-AI合作对领导感知努力的影响受到领导印象管理归因及领导AI信任的调节:当领导将员工行为归因为印象管理的倾向更低时,员工-AI合作更易被感知为努力;当领导AI信任程度更高时,员工-AI合作更容易被领导感知为努力。领导-下属配对问卷调查进一步验证了其中的调节效应:领导印象管理归因通过调节员工-AI合作与领导感知员工努力的关系,间接影响员工工作绩效和领导授权;领导AI信任通过调节员工-AI合作与领导感知员工努力的关系,间接影响员工工作绩效和领导授权。本文揭示了AI时代领导者社会认知的内在机制,为推动人与AI协作的组织管理实践提供了理论参考。

关键词:员工-AI合作 领导感知员工努力 印象管理归因 AI信任

中图分类号:F272.7

文献标识码:A

文章编号:1000-7636(2026)09-0146-15

一、问题提出

在数字经济与人工智能(AI)技术快速发展的背景下,越来越多的企业推动员工在工作中使用AI,并进一步将更深度的员工-AI合作嵌入工作流程,以构建更高效、更智能的组织管理体系^[1-3]。员工-AI合作(employee-AI collaboration)是员工使用AI的一种深度协同形式,强调员工围绕特定工作目标,通过与AI持续互动、反馈修正和结果整合,共同完成问题分析与成果产出^[4-6]。事实上,这一协作模式的组织意义不仅在于它是否提高了效率^[3],还在于它如何被领导理解和评价^[6-7]。在组织评价中,努力是领导判断员工工作价值、责任感和可托付性的重要线索^[8]。领导通常难以完整观察员工完成任务的过程,他们往往会依据员工表现出的投入、坚持和付出来推断其努力程度,形成绩效评价和授权判断^[9-10]。然而,AI的介入使这一判断过程变得更加复杂:员工与AI合作可能模糊员工投入与技术贡献之间的边界^[11],使领导难以判断高质量结果究竟是否源于员工努力^[7]。

收稿日期:2026-01-13;修回日期:2026-08-26

基金项目:国家社会科学基金重点项目“人工智能算法劳动的权力平衡研究”(22AGL014)

作者简介:金杨华 浙江工商大学工商管理学院教授、博士生导师,杭州,310018;浙江金融职业学院教授,杭州,310018;

朱汕怡 浙江工商大学工商管理学院博士研究生;

吴珍珍 浙江工商大学工商管理学院硕士研究生;

谢江佩 浙江工商大学工商管理学院副教授,通信作者。

作者感谢匿名审稿人的评审意见。

现有研究为理解员工-AI 合作提供了重要基础,但对其如何影响领导评价过程的解释仍不充分。已有研究多从技术效能视角关注员工-AI 合作如何影响任务结果和使用者自身体验^[4,6,12-15],较少解释当员工-AI 合作成为一种可被领导观察的工作行为时,领导如何据此形成管理判断。近年来,学术界开始关注 AI 相关行为的第三方评价后果。例如,领导可能将机器人带来的绩效提升归因于员工,从而高估员工绩效^[7];员工使用 AI 可能会被误认为懒惰,进而引发社会评价惩罚^[16]。上述发现表明,AI 相关行为的评价后果并不统一,其关键在于评价者如何理解员工与技术,在结果形成中的相对作用。忽视这一问题将难以理解员工-AI 合作在组织中的复杂后果:同样是使用 AI 完成工作,有的员工可能被视为主动学习、积极投入,有的员工却可能被视为依赖工具、减少付出^[9,17-19]。

依据归因理论,观察者在解释他人行为时,会综合行为者自身因素与外部情境因素,对行为原因作出判断^[20-21]。在员工-AI 合作情境中,领导对员工是否努力的判断不是由 AI 合作行为本身直接决定,而是取决于领导如何解释这一行为的原因^[20,22]。具体而言,领导需要同时处理两类关键信息:一是员工为何使用 AI,即对行为动机的判断;二是 AI 是否值得信任,即对外部工具的判断。前者对应领导的印象管理归因。当领导认为员工与 AI 合作主要是为了塑造积极形象或迎合评价期待时,该行为所传递的投入、主动性和责任感会被削弱,因而较难被解读为真实的努力^[22-23]。后者对应领导 AI 信任。当领导认为 AI 是可靠、适配且可控的工作工具时,员工-AI 合作更容易被视为员工有效利用技术完成工作的表现;相反,当领导缺乏对 AI 的信任时,AI 合作更可能被解读为依赖工具或规避个人投入,从而削弱其努力信号的诊断价值^[20,24]。因此,员工-AI 合作能否提升领导对员工努力的感知,取决于领导对员工动机和 AI 工具的双重归因^[20,22]。进一步地,这种由隐性认知投入引发的近端努力知觉,将深度渗入管理者的决策过程中。努力感知在不同决策语境下具有非对称的信号价值,它不仅影响领导对员工过去绩效成果的回溯性认定^[25-26],更深刻影响领导对未来控制权让渡的前瞻性判断,从而驱动组织资源配置的深层分化^[27-28]。

综上,本文可能的边际贡献在于三个方面。一是将第三方视角下 AI 行为研究的近端机制从“结果评价”前移至“努力感知”层面。既有研究多聚焦远端的道德评判或绩效结果,较少解释这些评价是如何形成的^[25]。本文关注领导如何依据员工在任务界定、输出检验与结果整合中的投入形成努力判断,并进一步将员工-AI 合作与后续管理评价相联系^[3],有助于揭示员工-AI 合作由行为线索转化为社会评价的认知过程。二是整合员工行为动机与 AI 工具属性两类解释线索,拓展归因理论在员工-AI 合作情境中的应用。本文将领导印象管理归因与领导 AI 信任纳入同一框架^[22],有助于说明员工-AI 合作能否被识别为努力,既取决于领导对行为动机真实性的判断^[9],也取决于其对 AI 作为工作资源的能力与可靠性的认可^[24]。三是将绩效评价与领导授权纳入同一分析框架,拓展员工-AI 合作组织后果的研究边界。本文区分绩效评价与领导授权的决策属性:前者侧重对既有工作贡献的认定,后者涉及对未来责任承担与自主权使用的判断^[27-28]。通过考察努力感知在两类结果中的作用,有助于说明员工-AI 合作不仅可能影响领导对其既有贡献的认可,也可能影响其后续获得自主权与发展机会的可能性。

二、理论基础与研究假设

(一) 归因理论

归因理论最早由海德(Heider)^[29]在 1958 年提出,由凯利(Kelley)^[22]和韦纳(Weiner)^[20-21]等学者发展完善,成为解释如何推断个体行为原因的重要社会认知理论。内部归因与外部归因构成了其核心分析维度^[20]。当领导将员工的高绩效归因于内部因素时,通常认为该行为源自员工的能力、努力或责任感;反之,当高绩效被解释为外部因素所致时,则该行为更可能被视为由情境、工具或运气等外部条件驱动^[30]。这种

内外因的权重分配决定了观察者如何理解和评价他人的行为^[20]。同时,依据归因理论的折扣原则,内外因的权重分配取决于线索的显著性与诊断性^[22]。当可见的外部原因足以解释结果时,观察者对内部原因的推断会减弱^[22,31]。

在组织情境中,领导作为关键的观察者与评价者,往往会基于归因过程来理解员工的工作行为^[32],并据此形成对员工潜力和绩效的评估^[7]。归因理论为理解员工-AI合作等复杂场景中的评价提供了解释:当行为结果可以由人与技术共同解释时,评价者会根据内外因线索的显著性与诊断性进行条件性的因果权重分配^[22]。在此过程中,领导不仅需要判断“结果在多大程度上归因于员工本人”,还需要判断“员工采取这一行为的动机是什么”,这两种判断共同影响其对员工努力程度的感知^[33]。

(二) 员工-AI合作与领导感知员工努力

在组织实践中,领导通常无法直接观察到员工完成任务的全过程,因此需要通过有限的外部信息^[29,32],对员工的行为原因进行归因推断^[22]。在这一过程中,努力启发式是领导进行社会判断的重要心理机制^[10]——观察者倾向于根据行为中可见的努力程度来评估其价值、质量或难度^[10,34]。当个体展现出更高的投入、付出或坚持时,观察者更容易将其评价为勤勉、可靠和值得信任的员工。这种“努力即价值”的判断逻辑广泛存在于绩效评估和任务分配等管理决策过程中^[10]。

然而,员工与AI的合作改变了领导对员工努力的判断方式。一方面,AI的自动化与认知替代使得员工完成工作的路径变得隐蔽,尤其是削弱了加班、高强度投入等传统努力线索的可见性,模糊了“谁在努力”这一核心归因信息^[11,13,34]。特别地,当技术这一外部因素足以解释高产出时,领导可能依据折扣原则降低对员工内部因素(如能力、努力)的归因权重,将AI视为绩效结果的竞争解释源^[22,35]。另一方面,员工主动将AI纳入任务过程,可能体现其技术适应能力、学习投入,以及在人机互动中进行的任务定义、输出检验与结果整合等“隐性劳动”^[7,11,36]。

员工-AI合作本身并不具有固定的努力意义,它既可能被解读为“技术依赖”而削弱努力感知,也可能被解读为“技术赋能”而增强努力感知^[7,16]。员工-AI合作的努力信号是模糊且可多重解读的,领导如何“解码”该行为就成为决定其效应方向的关键^[22]。换言之,员工-AI合作对领导感知员工努力的影响,依赖于领导采用何种认知框架来解释该行为背后的因果归属与动机意图^[22,32]。因此,本文从归因理论出发,聚焦两类影响领导解读的认知框架:一是将印象管理归因作为对员工行为动机的归因^[9,23],二是将AI信任作为对AI工具能力的判断^[24]。这两类框架分别提供动机归因与工具判断线索,从而影响领导对员工-AI合作中所包含的员工投入的解读,并形成对员工努力程度的差异化判断。

(三) 员工-AI合作与领导印象管理归因的交互影响

印象管理归因通过领导对员工行为动机的解读,影响员工-AI合作被感知为努力信号的强度。从领导视角看,员工-AI合作并非单纯的技术使用行为,也是一种包含动机信息、可被观察和解读的社会信号,并据此影响领导对员工努力程度的判断。根据归因理论,当领导无法直接观察员工的真实动机时,通常会依据行为特征、情境信息及既有认知推断其行为原因^[20,29]。领导印象管理归因是指领导认为员工将AI纳入工作过程,部分或主要是为了塑造有利形象、影响他人评价或避免负面评价,而非完全出于改进工作和完成任务的内在需要^[18,36-37]。基于博利诺等(Bolino et al.)^[9]的研究,印象管理涵盖积极争取形象的进取性策略与避免负面评价的防御性策略两种类型;员工-AI合作本身具有双重的形象信号价值:它既可能被解读为主动适应数字化转型的表现,也可能被视为借助外部工具减少个人投入的策略性行为。这种信号的多义性,使得领导在观察员工-AI合作行为时,可能会将其解读为带有形象维护意图,这也是一种认知上合理且常见的

判断。

当领导作出较高的印象管理归因时,其关注点会由员工-AI 合作所体现的工作投入,转向员工实施该行为可能获得的形象收益^[9]。此时,领导更可能认为员工使用 AI 并非主要为了改进工作,而是为了向领导展示自身具备数字化能力或者掩饰其实际投入不足^[23]。根据折扣原则,当某一行为可以由外在的形象管理动机加以解释时,观察者便会降低对内在努力动机的归因权重^[22]。相反,当领导印象管理归因较低时,领导较少怀疑员工使用 AI 是在进行形象塑造或评价管理,而更可能将其理解为完成任务、解决问题和改进工作方式的真实需要^[18]。

基于以上分析,本文提出假设 1:员工-AI 合作和领导印象管理归因的交互作用负向调节领导感知员工努力;当领导印象管理归因较低时,员工-AI 合作更容易被领导感知为努力。

(四) 员工-AI 合作与领导 AI 信任的交互影响

除了对主体行为动机的解读外,归因理论也强调评价者会利用与“工具”相关的线索在人与技术之间分配因果权重^[20-21]。在员工-AI 合作情境中,这意味着领导对 AI 的认知判断会改变其将绩效结果归因于员工努力还是归因于技术本身,从而调节员工-AI 合作在多大程度上被解读为“努力工作”的信号。基于此,本文引入领导 AI 信任,并将其界定为领导相信 AI 能够稳定、可靠地支持任务完成,以及由此形成的对 AI 应用于工作场景的总体接受程度^[3-6,24]。在此基础上,本文进一步考察领导 AI 信任对员工-AI 合作与领导感知员工努力之间关系的调节作用。

领导 AI 信任会影响其对员工-AI 合作行为的解释及其努力信号的识别。当领导 AI 信任程度较低时,领导容易将 AI 视为不稳定、不准确或难以控制的工具,因而对员工将 AI 纳入工作过程的合理性持怀疑态度,并可能将其解读为对技术的过度依赖或减少个人投入。由此,员工在学习 AI、判断输出和协调人机任务过程中所付出的努力难以被充分识别,领导对员工-AI 合作在表征“努力工作”方面的感知也随之降低^[3,18]。相反,当领导 AI 信任程度较高时,领导更认可 AI 作为工作资源的有效性与适用性,因而更可能将员工-AI 合作理解为主动学习新技术和优化工作方式的表現,并注意到员工在和 AI 合作过程中所投入的认知与协调努力。因此,较高的领导 AI 信任并非使领导简单地将工作成果归因于 AI,而是提升了员工使用和整合 AI 行为的合理性,使员工-AI 合作更容易被识别为主动适应、资源整合和能力拓展的努力,进而强化领导对员工努力程度的积极判断^[5,24]。

基于以上分析,本文提出假设 2:员工-AI 合作和领导 AI 信任的交互作用正向调节领导感知员工努力;当领导 AI 信任较高时,员工-AI 合作更容易被领导感知为努力。

(五) 有调节的中介模型

前述分析表明,领导对员工行为动机和 AI 工具属性的判断会改变员工-AI 合作所传递的努力含义。进一步而言,这种具有情境依赖性的判断并不会停留在认知层面,而会影响领导对员工既有贡献和未来可托付性的判断,进而转化为绩效评价与授权决策。基于此,本文进一步提出有调节的中介模型,即员工-AI 合作与领导印象管理归因、领导 AI 信任的交互作用,通过领导感知员工努力这一中介机制,对员工工作绩效与领导授权产生条件性间接影响。

领导感知员工努力是两类交互作用影响员工工作绩效的关键中介机制。工作努力体现员工为实现任务目标而投入行为的方向、强度与持续性,是推动工作绩效形成的重要近端因素^[3,26]。然而,领导对员工绩效的评价并非对客观产出的简单记录,还包含对绩效来源和员工个人贡献的归因判断^[25]。在员工-AI 合作情境中,领导感知员工努力程度越高,越可能将 AI 支持下形成的成果归因于员工的问题界定、专业判断、输

出检验与结果整合,而非完全归因于技术本身,从而给予员工更积极的绩效评价^[3,6,16]。

员工-AI合作与两类领导解释线索的交互作用会通过领导感知员工努力形成差异化的绩效评价。尽管领导印象管理归因与领导AI信任分别指向员工的行为动机和AI的工具属性,但二者最终都会影响领导是否将员工-AI合作识别为员工真实投入。具体而言,当领导印象管理归因较低时,领导更可能将员工-AI合作理解为完成任务、改善工作和适应技术变革的真实需要,因而更容易识别员工在任务设计、AI输出检验和结果整合中的努力,形成更积极的绩效评价。相反,当领导印象管理归因较高时,员工-AI合作更容易被理解为塑造积极形象或迎合评价期待的策略性行为,努力信息因对行为动机的真实性的质疑而减弱,进而削弱员工-AI合作通过努力感知对工作绩效产生的积极影响^[6,37]。

当领导AI信任程度较高时,领导更可能将AI视为具备任务能力、运行可靠性和结果可预测性的适宜工作资源,并对采用AI持更积极的态度。在这一认知基础上,员工选择、检验和整合AI输出更容易被理解为主动学习与有效配置技术资源,领导也更容易识别其中包含的认知投入。相反,当领导AI信任程度较低时,员工-AI合作更容易被理解为工具依赖、投入规避或责任转移。员工在AI合作过程中的实际投入较难被识别为有价值的努力,其通过努力感知获得积极绩效评价的可能性也相应降低^[3,7,24]。

基于以上分析,本文提出以下假设:

假设3a:领导印象管理归因负向调节员工-AI合作通过领导感知员工努力对员工工作绩效的间接效应;当领导印象管理归因较低时,该条件性间接效应更为正向。

假设3b:领导AI信任正向调节员工-AI合作通过领导感知员工努力对员工工作绩效的间接效应;当领导AI信任较高时,该条件性间接效应更为正向。

领导感知员工努力也是两类交互作用影响领导授权的关键中介机制。与主要面向员工既有表现的绩效评价不同,授权意味着领导向员工让渡未来的任务自主权和决策影响力,并承担由此产生的管理风险。因此,领导在作出授权决策时,不仅关注员工是否具备完成任务的能力,还会判断其是否愿意持续投入、主动承担责任并审慎运用所获得的自主权。已有研究指出,授权并非越多越好,因此领导需要综合评估下属的能力与工作态度^[38];同时,领导是否赋予下属决策影响力,取决于其对下属特征以及领导-下属关系的判断^[27-28]。这说明,授权并非对既有绩效的简单奖励,而是建立在领导对员工未来可靠性与可托付性的综合评估之上。较高的努力感知能够向领导传递员工持续投入、责任承担与审慎使用技术的信号,从而增强其授权意愿。

员工-AI合作与两类领导解释线索的交互作用,会通过领导感知员工努力形成差异化的授权决策。当领导印象管理归因较低时,员工-AI合作更容易被视为源于真实任务需要,较高的努力感知表明员工愿意持续投入、主动承担任务责任并审慎运用技术资源,因而更可能被视为值得托付的授权对象。相反,当领导印象管理归因较高时,努力信号的真实性受到质疑,领导对员工责任承担意愿和未来可靠性的判断也趋于保守,从而降低授权意愿^[4,6,16]。

当领导AI信任程度较高时,其会认为AI是可靠且适宜的任务资源,员工与AI合作也更容易被理解为主动获取技术支持并整合多种资源的表现。员工在这一过程中的投入越容易被识别为努力,领导越可能提高授权水平。相反,当领导AI信任程度较低时,员工-AI合作更容易被解读为依赖不稳定工具、减少个人投入或转移责任。即使员工进行了信息甄别与结果修正,其投入也可能因领导对AI本身的质疑而难以获得认可。较低的努力感知会强化领导对员工能否可靠使用自主权的担忧,进而削弱其授权意愿^[6,16,37]。

基于以上分析,本文提出以下假设:

假设4a:领导印象管理归因负向调节员工-AI合作通过领导感知员工努力对领导授权的间接效应;当

领导印象管理归因较低时,该条件性间接效应更为正向。

假设 4b:领导 AI 信任正向调节员工-AI 合作通过领导感知员工努力对领导授权的间接效应;当领导 AI 信任程度较高时,该条件性间接效应更为正向。

三、研究设计

(一) 情境实验

1. 被试与程序

本文通过软件 G Power 3.1 计算出所需样本量。当设定中等效应量 0.25、显著性水平 0.05 及统计检验力达到 0.80 时,至少需要 128 名被试。通过见数(Credamo)平台招募被试,设置筛选标准为:(1)全职就业人员;(2)担任企业领导职务(将“企业领导职务”操作化为在组织中承担正式管理职责,直接管理 1 名及以上下属员工的人员);(3)所在组织已经采用 AI 技术。其中,前两项由见数平台基于用户信息进行预筛选;最后一项通过问卷设置的额外筛选题进行确认:“您所在组织是否已应用以下任一 AI 技术?”,并在题目前提供 AI 技术的操作性定义及示例(如聊天机器人、智能推荐系统、自动化决策系统等)。被试需要在“是/否/不确定”中进行选择。本文仅保留回答“是”的被试,剔除回答“否”和“不确定”的样本。最终共有 271 名符合条件的被试自愿参与正式实验。

为确保数据质量与研究结论的可靠性,本文对原始样本进行了严格筛选,剔除了作答时间过短、作答极端以及未通过注意力检测的无效样本。最终保留了 200 份有效样本,有效回收率为 73.80%。其中,男女比例基本合理,女性 109 人(54.50%),男性 91 人(45.50%);年龄分布主要集中在青年群体,25 岁及以下 11 人(5.50%),26~35 岁 131 人(65.50%),36~45 岁 42 人(21.00%),46~55 岁 11 人(5.50%),56 岁及以上 5 人(2.50%);教育水平分布主要集中在本科和硕士阶段,本科及以上学历 190 人(95.00%)。

在研究设计方面,本文采用 2×2 被试间实验设计。具体程序如下:被试首先填写了包括性别、年龄、教育水平在内的人口统计学信息及 AI 信任;其次,将被试随机分配至四种实验情境之一(高/低员工-AI 合作水平×高/低领导印象管理归因),并要求其认真阅读相应的情境操纵材料;最后,被试依据所阅读的材料完成相关变量的测量。

2. 实验材料

借鉴以往关于员工-AI 合作和印象管理归因的情境研究设计,并基于本文的具体研究情境,设计了领导观察员工与 AI 合作行为的情境及后续判断的描述^[7,35,39]。被试阅读以下实验情境描述:“请设想您目前在一家科技公司担任部门领导。公司最近上线了一批先进的 AI 工具,这些工具可协助员工完成信息提取、趋势识别、报告撰写等工作。AI 工具对所有员工开放,但是否使用完全由个人决定。您的下属小王已经在您手下工作一年有余。最近,您要求他/她准备一份关于‘市场趋势分析与产品优化建议’的专项报告,该报告用于即将召开的高层会议。项目接近尾声时,您开始回顾小王在准备这份报告过程中的表现。”

在高员工-AI 合作组,被试会阅读以下内容:“您注意到小王在工作日频繁与 AI 合作,即大约 60%的工作时间用于与 AI 合作。他/她在信息识别与评估、机会与风险识别、预测、决策及问题解决等多个环节广泛使用 AI 工具。”在低员工-AI 合作组,被试会阅读以下内容:“您注意到小王在工作日很少与 AI 合作,即大约 5%的工作时间用于与 AI 合作。他/她在信息识别与评估、机会与风险识别、预测、决策及问题解决等部分环节偶尔使用 AI 工具。”

在高印象管理归因组,被试会阅读以下内容:“研究表明,员工与 AI 合作是指在工作场所中,员工与 AI 共同完成工作任务的行为过程。关于下属为何会在工作中与 AI 进行合作一直存在争议。一项在美国和中

国开展、涉及多个组织中数千名参与者的新研究,为以下结论提供了支持:员工与 AI 合作的行为,往往是因为员工在刻意维护和塑造积极形象。根据这项研究,员工与 AI 合作的行为是一种带有印象管理意图的策略性行为。”在低印象管理归因组,被试会阅读以下内容:“研究表明,员工与 AI 合作是指在工作场所中,员工与 AI 共同完成工作任务的行为过程。关于下属为何会在工作中与 AI 进行合作,长期以来一直存在争议。一项在美国和中国开展、涉及多个组织中数千名参与者的新研究,排除了‘员工与 AI 合作的行为是下属在刻意维护和塑造积极形象’这一观点。根据这项研究,员工与 AI 合作的行为不带有印象管理意图,而是出于实际工作需要。”

3. 测量工具

本文采用的量表均来自国内外较为成熟的测量方式。为保证其在中国情境下的效度,采用标准的翻译-回译程序对英文量表进行处理,以保证量表题项在语义上的完整性。所有量表均通过李克特(Likert)7点评分法进行评价,1表示非常不同意,7表示非常同意。

员工-AI 合作:采用孔等人(Kong et al.)^[5]开发的 5 题项量表,样题为“AI 参与小王的问题解决过程”,克隆巴赫系数 Cronbach's $\alpha=0.96$ 。

感知员工努力:采用莫尔和比特纳(Mohr & Bitner)^[34]开发的 5 题项量表,样题为“小王为完成工作任务付出了很多精力”,Cronbach's $\alpha=0.96$ 。

印象管理归因:采用里乌和彭纳(Rioux & Penner)^[17]开发的,经过李等人(Li et al.)^[18]验证的 5 题项量表,样题为“小王与 AI 合作是因为他/她试图避免看起来懒惰”,Cronbach's $\alpha=0.95$ 。

AI 信任:采用乔杜里等(Chowdhury et al.)^[3]的 11 题项量表,样题为“我对 AI 在组织内部业务运营方面所产生的影响持积极态度”,Cronbach's $\alpha=0.81$ 。

控制变量:本文控制了被试者的基本信息,包括性别、年龄和教育水平。

4. 实验结果

(1) 操纵检验

为检验实验中员工-AI 合作水平与印象管理归因的操纵是否有效,本文采用独立样本 t 检验对相关变量进行检验($N=200$)。结果显示,高员工-AI 合作条件下的参与者所感知的员工-AI 合作水平($M=6.1460$, $SD=0.4253$)明显高于低员工-AI 合作条件下的参与者($M=3.4780$, $SD=1.4658$), $t(198)=17.4805$, $P<0.001$, Cohen's $d=2.4721$ 。与此同时,高印象管理归因条件下的参与者所报告的印象管理归因水平($M=5.3160$, $SD=1.1848$)亦高于低印象管理归因条件下的参与者($M=2.3720$, $SD=1.0146$), $t(198)=18.8730$, $P<0.001$, Cohen's $d=2.6690$ 。上述结果表明,本文对员工-AI 合作水平与印象管理归因的操纵均是有效的。

(2) 假设检验

本文将印象管理归因作为实验操纵变量(高与低),采用 2×2 协方差分析检验其与员工-AI 合作的交互作用。结果显示,在控制了性别、年龄和教育水平后,员工-AI 合作对领导感知员工努力具有正向影响, $F(1,193)=20.8639$, $P<0.001$, $\eta_p^2=0.0976$ 。员工-AI 合作与领导印象管理归因的交互项对领导感知员工努力具有负向影响, $F(1,193)=7.7591$, $P<0.01$, $\eta_p^2=0.0386$ 。如图 1 所示,当领导印象管理归因较低时,高员工-AI 合作中领导对员工努力的感知高于低员工-AI 合作($M_{\text{高}}=5.8875$, $M_{\text{低}}=4.3218$);而当领导印象管理归因较高时,高、低员工-AI 合作条件下领导感知员工努力的差异减弱($M_{\text{高}}=4.6165$, $M_{\text{低}}=4.2342$)。上述结果表明,高领导印象管理归因削弱了员工-AI 合作向努力信号的转化作用。由此,假设 1 得到支持。

AI 信任为量表测量的连续个体差异变量。为避免信息损失与人为分组偏差,本文进一步利用软件 Mplus 7.4 进行回归分析,以检验员工-AI 合作与 AI 信任的交互项对领导感知员工努力的影响。在控制性别、年龄和教育水平后,表 1 结果显示,员工-AI 合作与领导 AI 信任的交互项对领导感知员工努力具有较强

的预测作用($\beta = 0.866, SE = 0.391, P < 0.05$)。该结果表明,领导 AI 信任能够调节员工-AI 合作与领导感知员工努力之间的关系。进一步结合交互效应方向可以发现,当领导 AI 信任水平较高时,员工-AI 合作水平越高,领导越倾向于将员工的工作表现解读为真实努力;而当领导 AI 信任水平较低时,员工-AI 合作对领导努力感知的正向作用明显减弱。由此,假设 2 得到支持。

情境实验通过操纵员工-AI 合作强度与印象管理归因,检验了领导视角下员工-AI 合作分别与印象管理归因、领导 AI 信任的调节作用,增强了研究的内部效度。考虑到模拟的实验情境可能与现实工作场景存在一定差异,研究结论的外部效度还有待扩展。接下来,基于下属自评与领导评价相结合的配对问卷数据,本文使用多时点的实地问卷调查对整体模型进行验证。

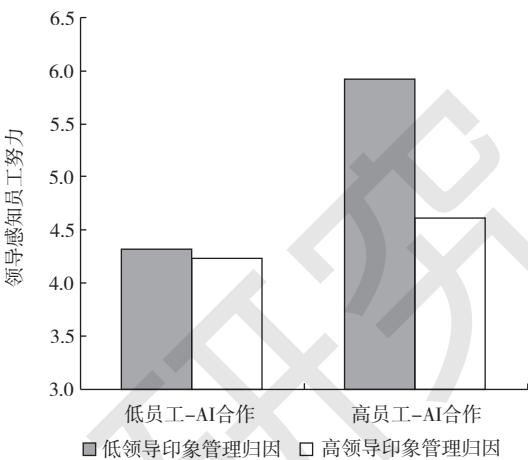


图1 员工-AI 合作和领导印象管理归因的交互对领导感知员工努力的影响结果

表 1 交互效应回归结果

变量	未考虑交互效应		考虑交互效应	
	β	SE	β	SE
截距	4.516 ***	0.697	4.305 ***	0.244
性别	-0.384	0.221	-0.376	0.215
年龄	0.024 *	0.012	0.029 *	0.012
教育水平	-0.267	0.186	-0.153	0.170
员工-AI 合作	0.980 ***	0.221	1.561 ***	0.258
领导印象管理归因		-0.042	0.351	
领导 AI 信任		-0.425	0.362	
员工-AI 合作×领导印象管理归因			-1.169 **	0.426
员工-AI 合作×领导 AI 信任			0.866 *	0.391
R^2	0.128			0.220
ΔR^2	0.087			0.092
F	7.160 ***			6.721 ***

注: N=200。回归系数 β 均为非标准化系数; * 表示 $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; 后表同。

(二) 实地问卷调查

1. 研究对象与数据收集

调研前期,项目组与样本企业负责人进行了沟通,以确认企业的 AI 应用情况及员工-AI 合作模式。5 家样本企业均已将智能客户关系管理、客户分析和业务辅助等 AI 系统嵌入日常工作流程。为避免同源偏差,本文采用多时点多来源的样本数据收集方法。问卷发放前由合作企业人力资源部门提供组织架构图,

研究团队为每对“领导-下属”分配唯一识别编码,下属报告其在过去一个月与 AI 进行实质性合作的频率与深度,而直属领导则报告对该下属努力程度的感知。数据分三个时间点收集:第一阶段(T1),228 名员工填写了人口统计学变量、员工-AI 合作的问卷。第二阶段(T2)直属领导($N=63$)提供人口统计学变量、感知员工努力、印象管理归因及 AI 信任信息。通过员工识别码确保员工-领导回答精准匹配。第三阶段(T3)要求领导使用含特定员工识别码的测量量表,评估对员工的授权行为以及员工工作绩效。剔除不匹配配对、逻辑矛盾回答及规律性答案后,最终获得 168 份有效领导-员工配对数据(回复率 73.68%),嵌套于 49 名领导(回复率 77.78%),每位领导平均评估 3 名员工。其中,年龄方面,平均年龄为 35.72 岁($SD=9.34$);性别方面,以女性为主,占比 67.26%($N=113$);教育背景方面,高中及以下 17.26%、大专 26.19%、本科 51.19%、硕士 5.36%。

2. 变量测量

本文采用的量表均来自国内外较为成熟的测量方式。为保证其在中国情境下的效度,本文采用标准的翻译-回译程序对英文量表进行处理,以保证量表题项在语义上的完整性。所有量表均通过李克特 7 点评分法进行评价,1 表示非常不同意,7 表示非常同意。

员工-AI 合作(T1)、领导感知员工努力(T2)、领导印象管理归因(T2)、领导 AI 信任(T2)均采用与上文情境研究相同的量表,只是指导语有所更改。量表的 Cronbach's α 依次为 0.94、0.90、0.99、0.97。

领导授权(T3):采用赵锴等^[40]的 7 题项量表,样题为“我给该员工相应的权限,让他/她在工作中能自主决策”,Cronbach's $\alpha=0.99$ 。

员工工作绩效(T3):采用卡拉基塔波卢-艾京等(Karakitapoğlu-Aygün et al.)^[41]的 3 题项角色内绩效量表,由领导进行评价,样题为“该员工的工作表现达到了公司的预期要求”,Cronbach's $\alpha=0.97$ 。

控制变量:本文将人口统计学因素列为控制变量,包括性别、年龄、教育水平,其中员工性别作虚拟变量处理,女性为 0,男性为 1。

3. 数据分析结果

(1) 验证性因子分析

本文采用软件 Mplus 进行多层次验证性因子分析(MCFA)。领导 AI 信任由领导自评,且同一领导对应的不同下属共享同一 AI 信任测量值,因此将其作为领导层(Level 2)变量处理;对员工-AI 合作、领导感知员工努力、领导印象管理归因、员工工作绩效和领导授权则作为个体层(Level 1)构念进行区分效度检验。基于此,本文设定了个体层五因子、领导层 AI 信任单因子的两层测量模型。结果显示,理论模型拟合达到可接受水平: $\chi^2(80)=143.708, P<0.001, CFI=0.948, TLI=0.930, RMSEA=0.071, SRMR_{within}=0.038$,且优于能够正常估计的备选模型^①。因此,结果支持本文主要研究构念的区分效度。

(2) 描述性统计分析与相关性分析

本文使用软件 SPSS 26.0 对样本的人口统计学变量和主要研究变量进行描述性统计分析和皮尔逊(Pearson)相关性分析。人口统计学变量包括性别、年龄和教育水平,主要研究变量包括员工-AI 合作、领导感知员工努力、领导印象管理归因、领导 AI 信任、领导授权及员工工作绩效。各变量的均值、标准差以及变量之间的相关系数见表 2。初步分析结果显示,员工-AI 合作与领导感知员工努力不相关($r=-0.0704, P>0.05$);领导感知员工努力与领导授权呈显著正相关($r=0.6932, P<0.001$),与员工工作绩效呈显著正相关($r=0.6901, P<0.001$)。以上结果为后续分析提供了依据。

① 限于篇幅,具体结果留存备案。

表 2 描述性统计结果与相关性分析

变量	M	SD	1	2	3	4	5	6	7	8
1 性别	0.327 4	0.470 7								
2 年龄	35.720 2	9.343 7	0.128 5							
3 教育水平	3.553 6	0.838 9	0.144 9	0.518 0***						
4 员工-AI 合作	2.809 6	1.427 7	0.134 4	0.261 9***	0.183 2*					
5 领导感知员工努力	5.688 1	1.259 5	0.009 6	-0.124 2	-0.257 3***	-0.070 4				
6 领导 AI 信任	4.530 9	1.305 6	0.008 8	-0.165 6*	-0.291 1***	-0.063 2	0.180 8*			
7 领导印象管理归因	3.702 4	1.986 4	0.224 0**	-0.053 9	-0.059 4	-0.022 3	0.148 7	0.017 1		
8 领导授权	5.523 0	1.366 3	0.043 4	-0.186 8*	-0.331 0***	-0.083 1	0.693 2***	0.325 4***	-0.056 0	
9 员工工作绩效	5.658 7	1.161 0	0.001 1	-0.069 2	-0.169 7*	-0.156 0*	0.690 1***	0.132 6	-0.062 7	0.604 3***

注：N=168。

(3)假设检验

为了检验前文提出的有调节的中介模型,本文使用了软件 Mplus 7.4 进行多层次回归分析和蒙特卡洛中介效应检验。由于部分样本在模型外生变量上存在缺失,最终进入多层次分析模型检验的样本为 160 名员工和 47 名领导。具体结果如表 3 所示。

首先,员工-AI 合作和领导印象管理归因的交互项负向影响领导感知员工努力($\beta=-0.387,P<0.001$)。由图 2 可知,当领导印象管理归因较低时,员工-AI 合作相较于高领导印象管理归因更容易被领导感知为努力。因此,假设 1 得到了支持。表 3 结果显示,员工-AI 合作与领导 AI 信任的交互项对领导感知员工努力有正向影响($\beta=0.058,P<0.05$)。由图 3 可知,当领导 AI 信任较高时,员工-AI 合作更容易被领导感知为努力。因此,假设 2 也得到了支持。

其次,领导感知员工努力正向预测工作绩效($\beta=0.438,P<0.001$)。蒙特卡洛抽样 20 000 次得到的分析结果(表 4)显示,当领导印象管理归因较低时,员工-AI 合作通过努力感知对工作绩效的间接效应为正($IE=0.068\ 7,95\%CI=[0.003\ 9,0.133\ 4]$);当领导印象管理归因较高时,该间接效应为负($IE=-0.141\ 4,95\%CI=[-0.240\ 8,-0.059\ 6]$),高低条件之差显著($diff=-0.210\ 1,95\%CI=[-0.352\ 3,-0.086\ 1]$)。当领导 AI 信任较低时,间接效应显著为负($IE=-0.069\ 5,95\%CI=[-0.142\ 4,-0.016\ 7]$);当领导 AI 信任较高时,该间接效应不显著($IE=-0.003\ 2,95\%CI=[-0.044\ 6,0.032\ 0]$),但高低条件之差显著为正($diff=0.066\ 3,95\%CI=[0.012\ 7,0.136\ 4]$)。因此,假设 3a 和 3b 均得到支持。

最后,领导感知员工努力正向预测领导授权($\beta=0.757,P<0.001$)。蒙特卡洛抽样检验结果表明(表 4),当领导印象管理归因较低时,员工-AI 合作通过努力感知对领导授权的间接效应为正($IE=0.118\ 7,95\%CI=[0.006\ 9,0.226\ 7]$);当领导印象管理归因较高时,该间接效应为负($IE=-0.244\ 4,95\%CI=[-0.430\ 8,-0.099\ 6]$),高低条件之差显著($diff=-0.363\ 1,95\%CI=[-0.619\ 7,-0.145\ 1]$)。当领导 AI 信任较低时,间接效应显著为负($IE=-0.120\ 1,95\%CI=[-0.266\ 9,-0.025\ 2]$);当领导 AI 信任较高时,该间接效应不显著($IE=-0.005\ 5,95\%CI=[-0.070\ 1,0.059\ 6]$),高低条件之差显著为正($diff=0.114\ 6,95\%CI=[0.017\ 4,0.265\ 1]$)。因此,假设 4a 和 4b 得到了支持。

表 3 多层次路径分析结果

变量	领导感知员工努力		员工工作绩效		领导授权	
	β	SE	β	SE	β	SE
截距	5.758***	0.171	5.753***	0.156	5.572***	0.171
性别	0.040	0.058	0.018	0.097	-0.004	0.102
年龄	0.022*	0.010	0.008	0.005	-0.006	0.008
教育水平	-0.308**	0.111	-0.034	0.098	-0.036	0.084
员工-AI 合作	-0.083*	0.040	-0.021	0.035	-0.022	0.029
领导感知员工努力			0.438***	0.103	0.757***	0.165
领导印象管理归因	-0.008	0.038	0.058	0.038	-0.151	0.202
领导 AI 信任	0.061	0.150	0.012	0.143	0.253	0.141
员工-AI 合作×领导印象管理归因	-0.387***	0.114				
员工-AI 合作×领导 AI 信任	0.058*	0.024				

注：个体层面 N=160，领导层面 N=47。

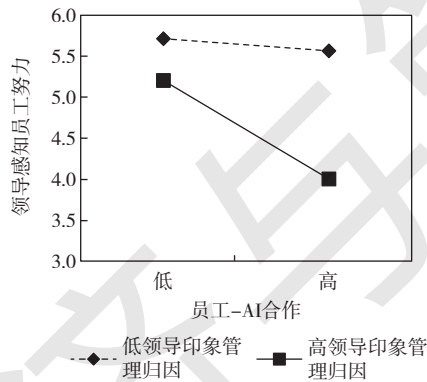


图 2 领导印象管理归因的调节效应

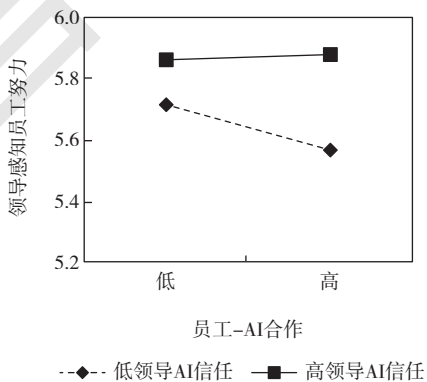


图 3 领导 AI 信任的调节效应

表 4 有调节的中介效应分析结果

调节路径	调节变量	间接效应	SE	95%置信区间	
				下限	上限
员工-AI 合作—领导感知员工努力—员工工作绩效	领导印象管理归因(+1 SD)	-0.141 4	0.046 7	-0.240 8	-0.059 6
	领导印象管理归因(-1 SD)	0.068 7	0.032 6	0.003 9	0.133 4
	领导印象管理归因差异	-0.210 1	0.068 5	-0.352 3	-0.086 1
	领导 AI 信任(+1 SD)	-0.003 2	0.019 0	-0.044 6	0.032 0
	领导 AI 信任(-1 SD)	-0.069 5	0.032 3	-0.142 4	-0.016 7
	领导 AI 信任差异	0.066 3	0.032 0	0.012 7	0.136 4

表4(续)

调节路径	调节变量	间接效应	SE	95%置信区间	
				下限	上限
员工-AI 合作—领导感知员工努力—领导授权	领导印象管理归因(+1 SD)	-0.244 4	0.084 9	-0.430 8	-0.099 6
	领导印象管理归因(-1 SD)	0.118 7	0.055 3	0.006 9	0.226 7
	领导印象管理归因差异	-0.363 1	0.121 3	-0.619 7	-0.145 1
	领导 AI 信任(+1 SD)	-0.005 5	0.032 4	-0.070 1	0.059 6
	领导 AI 信任(-1 SD)	-0.120 1	0.062 7	-0.266 9	-0.025 2
	领导 AI 信任差异	0.114 6	0.064 3	0.017 4	0.265 1

四、结论与启示

(一) 研究结论

基于归因理论,本文设定了一个以“被看见的努力”为近端机制的整合模型,考察员工-AI 合作如何经由领导对员工努力的感知,影响领导对员工的绩效评价与授权决策,并引入领导的印象管理归因与领导 AI 信任作为关键边界条件。研究结果表明:(1)当领导的印象管理归因较高时,员工-AI 合作较不易被解读为努力;(2)当领导 AI 信任较高时,员工-AI 合作更容易被解读为努力;(3)领导印象管理归因和领导 AI 信任会通过调节员工-AI 合作与领导感知员工努力的关系,进而间接影响员工工作绩效及领导授权。总体而言,本文揭示了 AI 时代领导者在“人-工具”因果权重分配下的社会认知机制:员工-AI 合作并非天然传达努力,其被如何解读,取决于领导对动机线索与工具正当性的并行加工,这为组织在推进人-AI 合作中放大被看见的努力、减少不当形象线索、提升工具信任与可控性提供了理论依据与实践路径。

(二) 理论贡献

第一,本文将第三方视角下员工-AI 合作行为的研究由远端评价结果推至领导感知员工努力这一近端认知机制。既有研究主要关注 AI 使用引发的绩效归因偏差、偷懒归因或道德评价等结果^[6-7,16],但较少解释这些评价如何形成^[25,34]。本文通过引入领导感知员工努力,揭示了领导如何依据员工在任务界定、输出检验和结果整合中的投入理解员工-AI 合作,完善了员工 AI 行为从行为观察到管理决策的解释过程。

第二,本文构建了整合“行为动机线索”与“工具正当性线索”的双路径调节框架。以往研究往往将对个体动机与对工具属性的判断割裂开来,或侧重于考察评价者对技术工具属性的知觉^[7],或侧重于分析观察者对个体动机的推断^[35],本文将动机归因路径(印象管理归因)与工具归因路径(领导 AI 信任)同时纳入模型^[22],系统考察了领导对“人”与“工具”两类线索的并行加工机制。这一框架拓展了归因理论在技术介入情境下的适用边界,也解释了为何同一 AI 合作行为在不同领导眼中会产生截然不同的努力解读。

第三,本文同时考察了员工-AI 合作对绩效评价和领导授权的影响,拓展了既有研究对单一评价结果的关注。以领导感知员工努力为关键机制,本文发现这种具有情境依赖性的判断既会影响领导在人与 AI 之间分配工作成果的功劳,进而塑造对员工既有绩效的评价^[10,25],也会影响领导对员工责任意识、持续投入意愿和可托付性的判断,进而影响授权决策^[27-28]。由此,将员工-AI 合作的评价后果由绩效认定

拓展至自主权配置,表明努力感知不仅关系到员工既有贡献能否得到认可,也关系到其能否获得更多决策空间与发展机会。这一发现不仅深化了对员工-AI 合作组织后果的认识,也为组织协调绩效评价与授权管理提供了理论依据。

(三) 实践启示

首先,组织需要更新 AI 情境下的努力评价标准。随着 AI 承担越来越多工作任务,传统以工作时长、过程可见性来判断努力的方式可能难以准确反映员工的真实贡献。管理者应更加关注员工在任务界定、AI 输出判断、修正与结果整合等方面的实际投入,并在绩效评价和授权决策中合理识别员工在人机协作中的独立贡献。其次,组织应提高员工-AI 合作过程的透明度,减少领导的印象管理误归因。组织可明确 AI 使用规范与边界,鼓励员工说明 AI 承担的任务及个人在判断、修正和整合中的贡献,降低领导对员工-AI 合作动机的不确定归因。最后,组织应促进管理者形成与 AI 实际能力相匹配的校准信任,通过 AI 培训等帮助管理者准确认识 AI 的能力与局限,避免因过度怀疑 AI 而误判员工投入。同时,在绩效评价与授权决策中,应区分“对 AI 的态度”与“对员工贡献的判断”,综合员工实际投入、任务责任与结果质量进行评价。

参考文献:

- [1] 戚聿东,沈天洋. 人工智能赋能新质生产力:逻辑、模式及路径[J]. 经济与管理研究,2024,45(7):3-17.
- [2] Einola K, Khoreva V. Best friend or broken tool?: Exploring the co-existence of humans and artificial intelligence in the workplace ecosystem[J]. Human Resource Management, 2023, 62(1): 117-135.
- [3] Chowdhury S, Budhwar P, Dey P K, et al. AI-employee collaboration and business performance: integrating knowledge-based view, socio-technical systems and organisational socialisation framework[J]. Journal of Business Research, 2022, 144: 31-49.
- [4] Tang P M, Koopman J, Yam K C, et al. The self-regulatory consequences of dependence on intelligent machines at work: evidence from field and experimental studies[J]. Human Resource Management, 2023, 62(5): 721-744.
- [5] Kong H Y, Yin Z H, Baruch Y, et al. The impact of trust in AI on career sustainability: the role of employee-AI collaboration and protean career orientation[J]. Journal of Vocational Behavior, 2023, 146: 103928.
- [6] Bankins S, Ocampo A C, Marrone M, et al. A multilevel review of artificial intelligence in organizations: implications for organizational behavior research and practice[J]. Journal of Organizational Behavior, 2024, 45(2): 159-182.
- [7] He G H, Yam K C, Zhao P C, et al. Leaders inflate performance ratings for employees who use robots to augment their performance[J]. Human Resource Management, 2025, 64(2): 543-563.
- [8] Cole M S, Walter F, Bruch H. Affective mechanisms linking dysfunctional behavior to performance in work teams: a moderated mediation study[J]. Journal of Applied Psychology, 2008, 93(5): 945-958.
- [9] Bolino M C, Kacmar K M, Turnley W H, et al. A multi-level review of impression management motives and behaviors[J]. Journal of Management, 2008, 34(6): 1080-1109.
- [10] Kruger J, Wirtz D, Van Boven L, et al. The effort heuristic[J]. Journal of Experimental Social Psychology, 2004, 40(1): 91-98.
- [11] Wilson H J, Daugherty P R. Collaborative intelligence: humans and AI are joining forces[J]. Harvard Business Review, 2018, 96(4): 114-123.
- [12] 蒋建武, 龙晗寰, 胡洁宇. 工作场所人工智能应用对员工影响的元分析[J]. 心理科学进展, 2024, 32(10): 1621-1639.
- [13] Jia N, Luo X M, Fang Z, et al. When and how artificial intelligence augments employee creativity[J]. Academy of Management Journal, 2024, 67(1): 5-32.
- [14] 孟祥婧, 徐鹏. 人工智能警觉对员工工作倦怠的影响研究: 基于科技公司的调查数据[J]. 经济与管理研究, 2024, 45(10): 99-110.
- [15] Shao Y D, Huang C Q, Song Y F, et al. Using augmentation-based AI tool at work: a daily investigation of learning-based benefit and challenge[J]. Journal of Management, 2025, 51(8): 3352-3390.
- [16] Reif J A, Larrick R P, Soll J B. Evidence of a social evaluation penalty for using AI[J]. Proceedings of the National Academy of Sciences of the

United States of America,2025,122(19):e2426766122.

- [17] Rioux S M, Penner L A. The causes of organizational citizenship behavior; a motivational analysis[J]. *Journal of Applied Psychology*, 2001, 86(6): 1306-1314.
- [18] Li W L, Qin X, Yam K C, et al. Embracing artificial intelligence (AI) with job crafting; exploring trickle-down effect and employees' outcomes[J]. *Tourism Management*, 2024, 104: 104935.
- [19] 王文涛, 修博文. 人工智能如何重塑劳动力市场: 创造还是替代? [J]. *首都经济贸易大学学报*, 2025, 27(6): 67-81.
- [20] Weiner B. An attributional theory of achievement motivation and emotion[J]. *Psychological Review*, 1985, 92(4): 548-573.
- [21] Weiner B. Attribution, emotion, and action[M]// Sorrentino R M, Higgins E T. *Handbook of motivation and cognition: foundations of social behavior*. New York: Guilford Press, 1986: 281-312.
- [22] Kelley H H. The processes of causal attribution[J]. *American Psychologist*, 1973, 28(2): 107-128.
- [23] Cheng J W, Lu K M, Chang Y Y, et al. Voice behavior and work engagement; the moderating role of supervisor-attributed motives[J]. *Asia Pacific Journal of Human Resources*, 2013, 51(1): 81-102.
- [24] Glikson E, Woolley A W. Human trust in artificial intelligence; review of empirical research[J]. *Academy of Management Annals*, 2020, 14(2): 627-660.
- [25] Reb J, Greguras G J. Understanding performance ratings: dynamic performance, attributions, and rating purpose[J]. *Journal of Applied Psychology*, 2010, 95(1): 213-220.
- [26] Van Iddekinge C H, Arnold J D, Aguinis H, et al. Work effort; a conceptual and meta-analytic review[J]. *Journal of Management*, 2023, 49(1): 125-157.
- [27] Yukl G, Fu P P. Determinants of delegation and consultation by managers[J]. *Journal of Organizational Behavior*, 1999, 20(2): 219-232.
- [28] Scandura T A, Graen G B, Novak M A. When managers decide not to decide autocratically; an investigation of leader-member exchange and decision influence[J]. *Journal of Applied Psychology*, 1986, 71(4): 579-584.
- [29] Heider F. The naive analysis of action[M]// Heider F. *The psychology of interpersonal relations*. Hoboken, NJ: John Wiley & Sons, 1958: 79-124.
- [30] Chen Z G, Huo Y Y, Lam W, et al. How perceptions of others' work and impression management motives affect leader-member exchange development; a six-wave latent change score model[J]. *Journal of Occupational and Organizational Psychology*, 2021, 94(3): 645-671.
- [31] Eberly M B, Holley E C, Johnson M D, et al. Beyond internal and external; a dyadic theory of relational attributions[J]. *Academy of Management Review*, 2011, 36(4): 731-753.
- [32] Martinko M J, Harvey P, Douglas S C. The role, function, and contribution of attribution theory to leadership; a review[J]. *The Leadership Quarterly*, 2007, 18(6): 561-585.
- [33] Inzlicht M, Shenav A, Olivola C Y. The effort paradox; effort is both costly and valued[J]. *Trends in Cognitive Sciences*, 2018, 22(4): 337-349.
- [34] Mohr L A, Bitner M J. The role of employee effort in satisfaction with service transactions[J]. *Journal of Business Research*, 1995, 32(3): 239-252.
- [35] Zhou X, Chen C, Li W L, et al. How do coworkers interpret employee AI usage; coworkers' perceived morality and helping as responses to employee AI usage[J]. *Human Resource Management*, 2025, 64(4): 1077-1097.
- [36] Wrzesniewski A, Dutton J E. Crafting a job; revisioning employees as active crafters of their work[J]. *Academy of Management Review*, 2001, 26(2): 179-201.
- [37] Lam W, Huang X, Snape E. Feedback-seeking behavior and leader-member exchange; do supervisor-attributed motives matter? [J]. *Academy of Management Journal*, 2007, 50(2): 348-363.
- [38] 林美珍, 罗忠恒. 授权型领导负面效应的形成机制: 基于“过犹不及”理论[J]. *首都经济贸易大学学报*, 2017, 19(5): 104-112.
- [39] Kang Y J, Peng J, Nie Q. Peer reaction to manager stewardship behavior; crediting or stigmatizing the behavior? [J]. *Journal of Business Ethics*, 2023, 183(2): 453-474.
- [40] 赵锴, 俞溪, 张山杉. 委以重任还是排斥打压?: 明星员工与团队领导的人际互动后果[J]. *心理学报*, 2024, 56(5): 630-649.
- [41] Karakitapoğlu-Aygün Z, Gumusluoglu L, Erturk A, et al. Two to Tango?: A cross-cultural investigation of the leader-follower agreement on authoritarian leadership[J]. *Journal of Business Research*, 2021, 128: 473-485.

How Does Employee-AI Collaboration Affect Leaders' Perceptions of Employee Effort?

JIN Yanghua^{1,2}, ZHU Shanyi¹, WU Zhenzhen¹, XIE Jiangpei¹

(1. Zhejiang Gongshang University, Hangzhou 310018;

2. Zhejiang Financial College, Hangzhou 310018)

Abstract: Artificial intelligence (AI) integration into organizational workflows has transformed employee-AI collaboration into a routine work practice. Existing research focuses primarily on the performance benefits and psychological outcomes of AI adoption from the employee perspective. However, less attention is paid to how leaders, as key observers, interpret this collaboration and form managerial judgments. This paper investigates when and why leaders perceive employee-AI collaboration as genuine effort, and how this perception shapes performance evaluations and empowerment decisions.

Attribution theory indicates that employee-AI collaboration carries no fixed meaning of effort. AI reduces the visibility of traditional effort cues while providing a strong external explanation for outputs, requiring leaders to decode the motivational meaning of employee behavior. Two leader-side cognitive frameworks govern this decoding process: impression-management attribution reflects the extent to which leaders attribute the collaboration to impression-enhancing motives. Trust in AI reflects the extent to which leaders view AI as a capable and reliable work resource. These frameworks shape leaders' perceptions of employee effort, which then serves as a proximal mechanism linking employee-AI collaboration to performance evaluation and empowerment.

Two complementary studies tested the proposed model. Study 1 employed a scenario-based experiment among practicing managers to test the moderating roles of impression-management attribution and trust in AI. Study 2 used a multi-wave, multi-source field survey of leader-employee dyads to replicate these interactive effects and examine downstream managerial consequences. Results supported the moderated mediation hypotheses: the indirect effects were more positive under low impression-management attribution and less negative under high trust in AI.

This paper makes three theoretical contributions. First, it advances third-party evaluations of AI-related behavior from distal outcomes such as performance inflation or laziness attributions to leaders' perceptions of employee effort as a critical proximal cognitive process. Second, it develops an integrative dual-pathway framework that incorporates both motivational cues of impression-management attribution and instrumental cues of trust in AI, thereby extending attribution theory to the emerging context of employee-AI collaboration. Third, by simultaneously examining performance evaluation and empowerment, it demonstrates that effort perceptions influence not only the recognition of employees' past contributions but also the allocation of future autonomy and developmental opportunities. Practically, the findings suggest that organizations should help leaders update traditional effort metrics and cultivate greater trust in AI technologies, so as to foster more accurate recognition of employee contributions in employee-AI collaboration.

Keywords: employee-AI collaboration; leaders' perceptions of employee effort; impression-management attribution; trust in AI

编校:李 叶;蒋 琰