

人工智能背景下员工亲组织非伦理行为的形成机制仿真研究

于文婷, 刘生敏*

上海理工大学管理学院, 上海

收稿日期: 2023年11月16日; 录用日期: 2023年12月5日; 发布日期: 2024年1月18日

摘要

人工智能技术(AI)对人类工作的替代刺激员工进行非伦理行为。本文探讨了AI背景下员工亲组织非伦理行为(UPB)的促发机制以及不同替代风险下的演化差异。UPB系统动态模型300天仿真显示: 员工感知到AI风险后的45天内UPB水平显著增长, 160天后保持相对稳定; AI替代风险越高, UPB整体越高。其中不同风险感知下UPB累计差异主要源于替代风险前45天内带来的UPB差异。该结论有助于组织优化人工智能和员工的替代关系, 改善AI的伦理困境。

关键词

亲组织非伦理行为, 人工智能替代, 系统仿真

System Dynamics Simulation on the Action Mechanism of Unethical Pro-Organizational Behavior

Wenting Yu, Shengmin Liu*

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Nov. 16th, 2023; accepted: Dec. 5th, 2023; published: Jan. 18th, 2024

Abstract

The applications of artificial intelligence bring a series of ethical problems. The paper explores the action mechanism of UPB under artificial intelligence substitution, the evolution differences of

*通讯作者。

UPB under different substitution risk. Based on the 300 day simulation, the following two conclusions are obtained: about 45 days before the evolution result, the level of UPB will increase significantly in a short time after employees perceive the risk of artificial intelligence, and remain relatively stable after about 160 days; The system operation found that the higher the risk perception of artificial intelligence substitution, the higher the overall level of UPB. Among them, the cumulative difference of UPB under different risk perceptions mainly comes from the difference of UPB within 45 days before the perceived replacement risk. This conclusion is helpful to prioritize the system and ethical dilemma of AI.

Keywords

Unethical Pro-Organizational Behavior, Artificial Intelligence Substitution, System Simulation

Copyright © 2024 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

当世界围棋冠军柯洁以 0 比 3 的比分完败于智能机器人阿尔法狗时, 人工智能的力量给人类智能形成了冲击, 围棋界人士甚至呼吁取消人类的围棋比赛。尽管人工智能可以帮助员工提升生产和工作效率, 但同时也给人类员工的管理伦理带来了挑战[1]。他们担心智能机器取代自己, 或害怕无法与机器人达到人机互动, 所以依附于组织利益而抵制人工智能。尽管人工智能技术发展已经相对成熟, 但人机互动伦理问题的研究相对匮乏。学界探索了人工智能对员工创造力、工作绩效和技术中心化的促进作用, 但缺乏对伦理行为的影响机制研究。

有鉴于此, 本文从资源保存视角出发, 以人工智能的替代风险为起点, 进行人机协作情境下人工智能风险感知的亲组织非伦理行为仿真预测, 完善人工智能与人类伦理决策的互动, 以深化人工智能伦理体系。

2. 系统构建

亲组织非伦理行为(unethical pro-organizational behavior, UPB), 也叫亲组织不道德行为, 由 Umphress 等(2010)提出, 是个体出于组织利益而主动采取的违反伦理道德甚至法律法规的行为[2]。此概念同时具有两重性质: 1) 非伦理性, 违反法律法规或普遍认同的伦理准则; 2) 亲组织性, 维护组织利益, 但并非来自于组织规定或领导指令。

2.1. 前提假设

根据系统动力学分析的基本过程, 在构建模型之前, 应当明确模型的边界与假设以更好的对主题进行研究, 主要提出以下假设:

假设 1: 亲组织非伦理行为形成机制系统主要研究的是感知到人工智能替代风险后, 工作不安全感、竞争氛围和自利型伦理氛围以及道德推脱、UPB 和领导成员交换关系之间的演化, 暂不考虑其他群体因素、个体特质以及组织因素带来的影响。

假设 2: 系统运行时间为 300 天, 无论是高、中或低替代风险感知, 员工在此阶段内仍在组织内工作, 并未失去工作岗位。

假设 3: 不考虑强人工智能的实现, 系统运行时间内不会出现技术突飞猛进式发展以及其他突发环境状况, 300 天内替代风险感知稳定为常数值。

2.2. 研究变量构成

前文曾提到系统动力学并非要求准确的穷尽系统运行要素, 而是基于特定视角进行动态系统的分析即可。因此, 本文延续第五章结构方程模型思路, 在 Trevino 伦理决策过程指导下对组织和个人因素进行动态的因果反馈分析。为了尽可能让系统分析更加科学和全面, 通过文献研究在系统增加了上下级关系因素。人工智能替代自然员工工作岗位的背景下, 个体、组织、领导交互进行物质和信息的交换, 形成了相互影响、相互反馈、相互支持的系统, 即员工亲组织非伦理行为形成机制系统。

本文研究的系统变量有 11 个。其中状态变量包括领导成员交换关系质量与亲组织非伦理行为; 对应速率变量为关系亲密度与关系疏远度, UPB 增加与 UPB 减少; 其他辅助变量工作不安全感、竞争氛围、自利型伦理氛围以及道德推脱。常量为人工智能替代风险感知。

系统动力学仿真更重要的是模型的构建, 数值绝对大小的差异对仿真结果影响不大, 对于一手或二手数据也没有要求。因此, 本文模型中变量之间的关系与取值基本采用了前人的路径系数, 工作不安全感与道德推脱的关系为 0.123, 竞争氛围与自利型伦理氛围的关系为 0.657, 自利型伦理氛围与道德推脱的关系为 0.637, 道德推脱与 UPB 增加的关系为 0.255。UPB 与领导成员关系之间的系数为 0.16, 领导成员交换关系与工作不安全感之间的关系为-0.12。人工智能替代风险感知为 0.2。

2.3. 因果关系

根据前文文献梳理以及结构方程模型的检验结果可得到以下三条工作不安全感与竞争氛围的因果关系, 并基于此进行下一步的流图绘制以及仿真模拟。

- 1) 工作不安全感↑ - 道德推脱↑ - 亲组织非伦理行为↑ - 领导成员交换关系↑ - 工作不安全感↓
- 2) 竞争氛围↑ - 自利型伦理氛围↑ - 道德推脱↑ - 亲组织非伦理行为↑ - 领导成员交换关系↑ - 工作不安全感↓ - 竞争氛围↓
- 3) 竞争氛围↑ - 自利型伦理氛围↑ - 亲组织非伦理行为↑ - 领导成员交换关系↑ - 工作不安全感↓ - 竞争氛围↓

2.4. 流量存量图建立

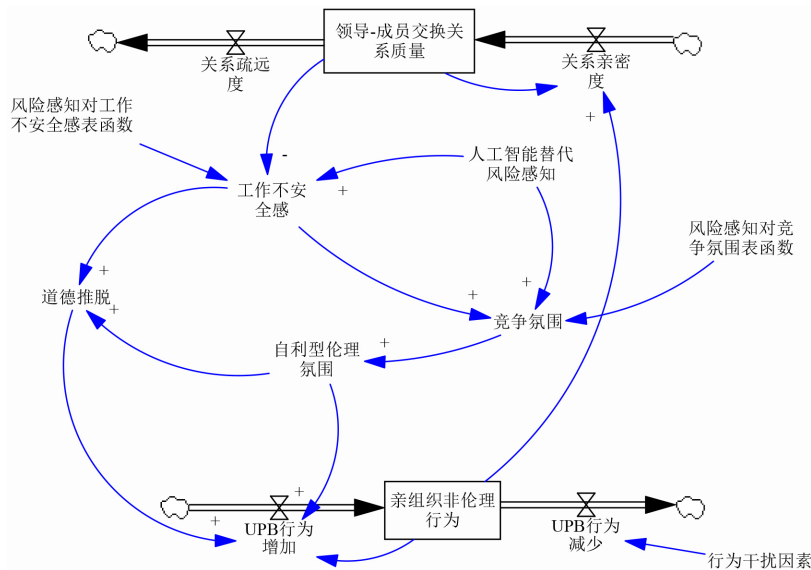


Figure 1. Model system flow
图 1. 模型系统流图

因果关系图是从定性的角度描述的变量之间的逻辑反馈。为了更好的研究变量间数量关系, 在因果关系的基础上增加了速率变量、水平变量以及系统干扰因素, 构建系统流图 1, 以具体阐释系统的动态变化。

2.5. 仿真方程构建

系统运行需要确定变量与变量之间数逻辑关系, 设计仿真方程。系统仿真时间为三百天, 步长为一天。亲组织非伦理行为以及领导成员交换关系质量为水平变量, 用积分函数 INTEG 表示; 速率变量为 UPB 行为增加、UPB 行为减少、关系疏远度和关系亲密度, 其中 UPB 行为减少与关系疏远度用干扰因素即随机函数表示, UPB 行为增加与关系亲密度用回归函数表示; 其他辅助变量之间的关系通过第五章结构方程模型以及文献分析得到, 道德推脱与自利型伦理氛围表示为回归函数, 竞争氛围与工作不安全感表示为指数函数以及对替代风险的表函数; 常量人工智能替代风险表示为低风险感知水平 0.2。

$$(01) \text{ FINAL TIME} = 300$$

$$(02) \text{ INITIAL TIME} = 0$$

$$(03) \text{ SAVEPER} = \text{TIME STEP}$$

$$(04) \text{ TIME STEP} = 1$$

$$(05) \text{ UPB 行为减少} = \text{行为干扰因素 1}$$

$$(06) \text{ UPB 行为增加} = (0.255 * \text{道德推脱} + 0.512 * \text{自利型伦理氛围}) * (50 - \text{亲组织非伦理行为}) * 0.115$$

$$(07) \text{ 亲组织非伦理行为} = \text{INTEG}(\text{UPB 行为增加} - \text{UPB 行为减少}, 4.36)$$

$$(08) \text{ 人工智能替代风险} = 0.2$$

$$(09) \text{ 关系亲密度} = 0.01047 * \text{亲组织非伦理行为} + (23 - \text{“领导 - 成员交换关系质量”}) * 0.015$$

$$(10) \text{ 关系疏远度} = \text{行为干扰因素 2}$$

(11) $\text{工作不安全感} = \text{EXP}(-0.126 * \text{“领导 - 成员交换关系质量”}) + \text{替代风险对工作不安全感表函数}(\text{人工智能替代风险})$

$$(12) \text{ 竞争氛围} = \text{EXP}(0.1 * \text{工作不安全感}) + \text{替代风险对竞争氛围表函数}(\text{人工智能替代风险})$$

$$(13) \text{ 自利型伦理氛围} = 0.657 * \text{竞争氛围}$$

$$(14) \text{ 行为干扰因素 1} = \text{RANDOM UNIFORM}(0, 1, 2)$$

$$(15) \text{ 道德推脱} = 0.123 * \text{工作不安全感} + 0.637 * \text{自利型伦理氛围}$$

$$(16) \text{“领导 - 成员交换关系质量”} = \text{INTEG}(\text{关系亲密度} - \text{关系疏远度}, 7.51)$$

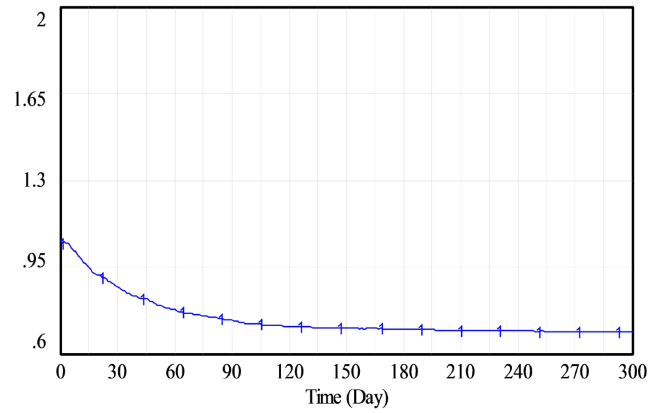
(17) $\text{替代风险对工作不安全感表函数}([(0, 0) - (1, 1)], (0, 0.46), (0.2, 0.67), (0.4, 0.81), (0.6, 0.89), (0.8, 0.95), (1, 1))$

(18) $\text{替代风险对竞争氛围表函数}([(0, 0) - (1, 1)], (0, 0.2), (0.2, 0.36), (0.4, 0.44), (0.6, 0.49), (0.8, 0.53), (1, 0.56))$

3. 仿真结果分析

初始替代风险为 0.2, 在这种风险水平下仿真结果如图所示, 观察亲组织非伦理行为变化趋势可以分成四个阶段: 在前 45 天左右, 员工的亲组织非伦理行为表现为迅速增长的态势; 从第 45 天开始至第 110 天左右, UPB 的增速放缓, 但仍然处于增加趋势; 从第 110 天至 160 天左右, 员工亲组织非伦理行为呈现一个略微减少的现象; 而从 160 天左右开始, 伴随着小幅波动, UPB 保持在一个稳定的范围内。

工作不安全感

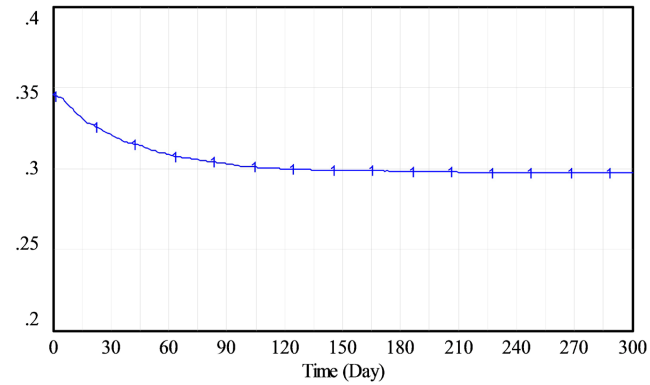


工作不安全感:基础

Figure 2. Job insecurity basement

图 2. 工作不安全感基础

道德推脱

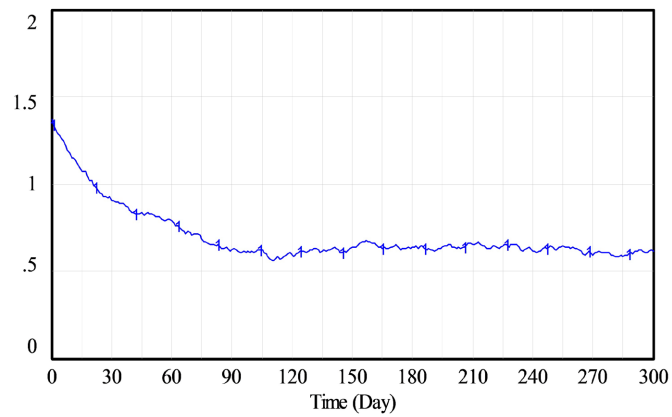


道德推脱:基础

Figure 3. Moral disengagement basement

图 3. 道德推脱基础

UPB行为增加

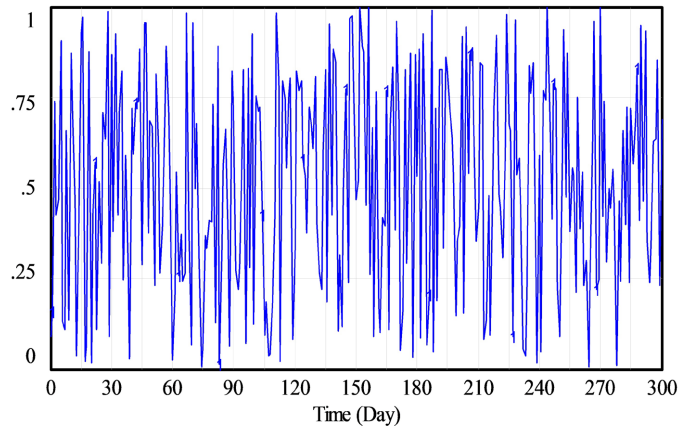


UPB行为增加:基础

Figure 4. UPB increment

图 4. UPB 增量

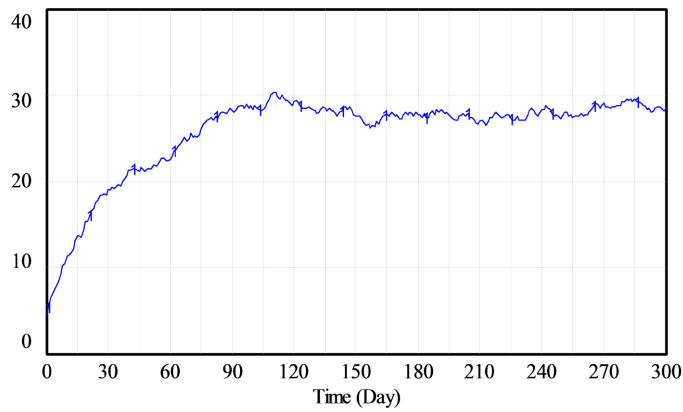
UPB行为减少



UPB行为减少:基础 ———— ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑

Figure 5. UPB decrement
图 5. UPB 减量

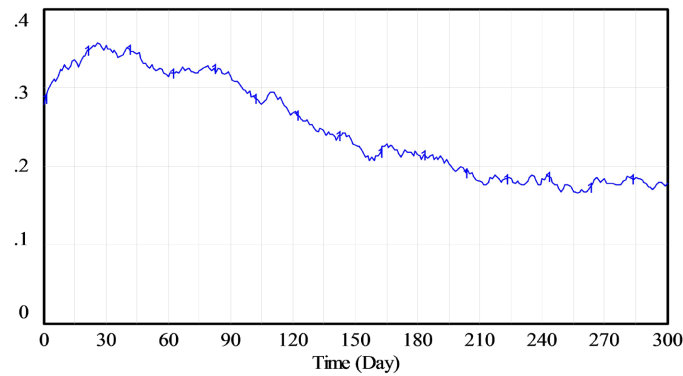
亲组织非伦理行为



亲组织非伦理行为:基础 ———— ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑

Figure 6. UPB
图 6. UPB

关系亲密度



关系亲密度:基础 ———— ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑

Figure 7. Relationship closeness
图 7. 关系亲密度

同时结合现实, 对仿真结果综合分析: 在人工智能进入组织初期, 员工开始感知到人工智能替代风险时, 工作不安全感是最高的(如图 2), 压力会诱发人的非理性决策, 即在未了解人工智能与自身的优势和劣势的情况下, 个体会在紧张情绪下更容易为了维持现有资源做出道德推脱, 不惜做出不道德或违法行为以给组织争取更多的利益, 从而增加组织存在感以获得认可与承诺, 做出更多的 UPB(如图 3)。对应的前 45 天, 速率变量 UPB 增加的数值在前期较高(如图 4), 亲组织非伦理行为随着基础波动(如图 5), 其积累量呈指数式增长(如图 6); 之后, 随着员工与人工智能的协作增加, 人工智能局限性也会逐渐清晰, 比如人工智能虽然能持续 24 小时工作, 但无法自主判断既定程序外应该终止运行的情况, 像上海频发的“列车夹人”事故就是智能失灵的典型。另外像宾馆内的“智能服务员”, 可以配送上门顾客需要的物品, 替代了服务员部分工作内容, 但是实时的监督控制、设备运维、满足住客个性化需求, 还有疫情期间对智能设备的消杀工作等工作, 仍离不开人工。因此, 随着对人工智能了解的客观性增加, 工作不安全感继续下降(如图 2), UPB 增加量持续减少(如图 4), 员工从事 UPB 的增速会有所放缓。对应的 45 至 110 天, 亲组织非伦理行为增速放缓(如图 6); 之后, 人机协同带来接触增加, 了解的程度越来越深入, 了解人数也增加, 意识到自身的竞争优势, 紧张程度减少并保持在稳定水平。由于 UPB 减少变量的存在(如图 5)以及人的心理变化呈现到行为变化存在时滞, 对应的亲组织非伦理行为在 110 天到 160 天呈现一个小幅下降(如图 6); 此后, 在人工智能并未出现突破性发展的前提下, 除了个体在日常中可能会感受到刺激而产生 UPB 波动, 长期人机协作中对人工智能替代的心理和行为反应会稳定在适应区间内, 即对应的 160 天开始亲组织非伦理行为水平基本维持在 28 左右的数值(如图 6)。

观察领导成员交换关系质量仿真结果可知, 领导成员交换关系质量整体呈现对数型增长的趋势, 大致可以分为两个阶段: 在前 60 天增速较快, 之后则增速放缓并趋于稳定值。

同时结合实际对仿真结果综合分析: 领导成员交换关系质量与关系亲密度的变化直接相关。在前 45 天, 亲组织非伦理行为水平激增(如图 6), 引起了与领导亲密度的明显增加(如图 7), 增至其最大限度。因此, 领导成员交换关系质量曲线在前期增速较快, 由于同时受关系疏远度(如图 8)以及水平变量时滞的影响, 在 90 天左右仍保持较快的增速; 之后, 随着员工亲组织非伦理行为增加量减少(如图 4), UPB 带来的边际亲密度会下降, 并且领导可能会开始考虑到 UPB 长期带来的风险, 关系亲密度开始下降(如图 7)并持续此下降趋势。因此, 自 90 天之后领导成员交换关系增速放缓, 平稳上升, 并在 270 天左右对应的关系亲密度的 240 天左右趋于稳态。这与亲组织非伦理行为水平的图像变化大致相符, 也印证了亲组织非伦理行为与领导成员交换关系的正相关性。

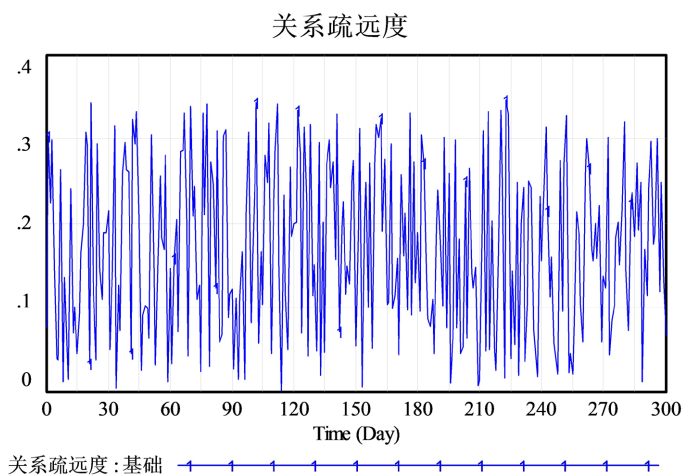


Figure 8. Relationship estrangement
图 8. 关系疏远度

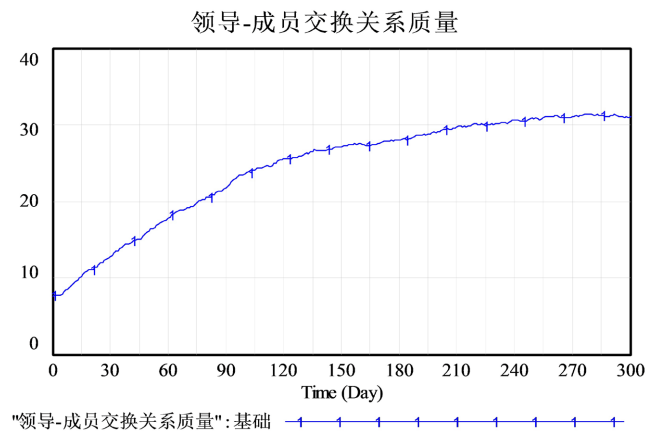


Figure 9. Leader-member relationship

图 9. 领导成员交换关系质量

此外, 领导成员交换关系的变化又直接影响了工作不安全感。在前 90 天内, 领导成员交换关系质量迅速提高(如图 9), 在我国关系文化社会中, 为员工注入了镇静剂, 同时可能带来更多的工作培训与能力培养资源, 促进职业发展, 换言之降低了工作不安全感(如图 1); 之后可能员工对关系质量感知敏感度下降, 但由于关系质量仍在增长(如图 9), 因此虽然增速放缓, 工作不安全感会基本稳定在 0.95 左右。工作不安全感会再通过影响道德推脱(如图 3)作用于 UPB 增加(如图 6), 因此, 三者曲线的变化基本一致。这些因果关系层层作用塑造了这一系统。

4. 情境分析

陈文晶等(2022)在测量人工智能潜在替代风险时, 将[0, 0.3]设为低替代风险, [0.3, 0.7]设为中等替代风险, [0.7, 1]设为高替代风险[3]。基于此, 本文拟探索在低中高三种替代风险情境下, 工作不安全感、竞争氛围、自利型伦理氛围、道德推脱、亲组织非伦理行为以及领导成员交换关系会发生何种变化。工作不安全感、竞争氛围、自利型伦理氛围以及道德推脱四个变量在三种替代风险情境下总体变化趋势基本不变, 并且替代风险越高, 对应的工作不安全感、竞争氛围、自利型伦理氛围以及道德推脱的值也越高。因此也验证了替代风险对四者的正向作用。

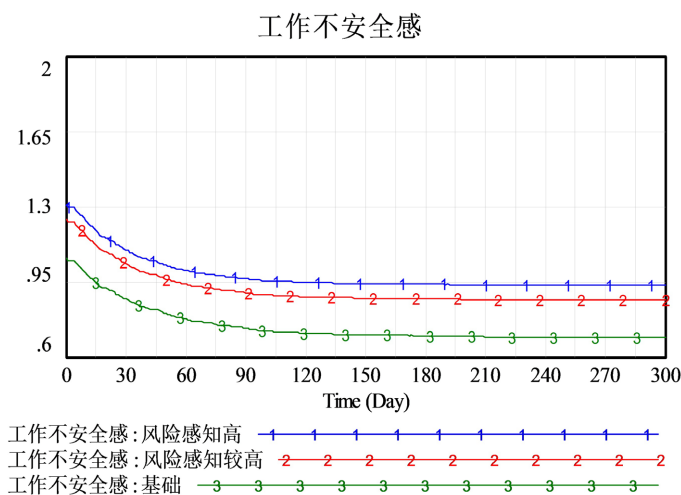


Figure 10. Job insecurity

图 10. 工作不安全

竞争氛围

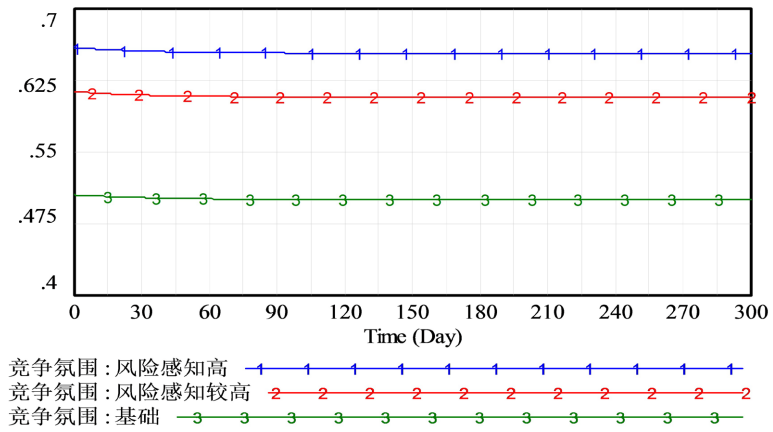


Figure 11. Competitive climate
图 11. 竞争氛围

自利型伦理氛围

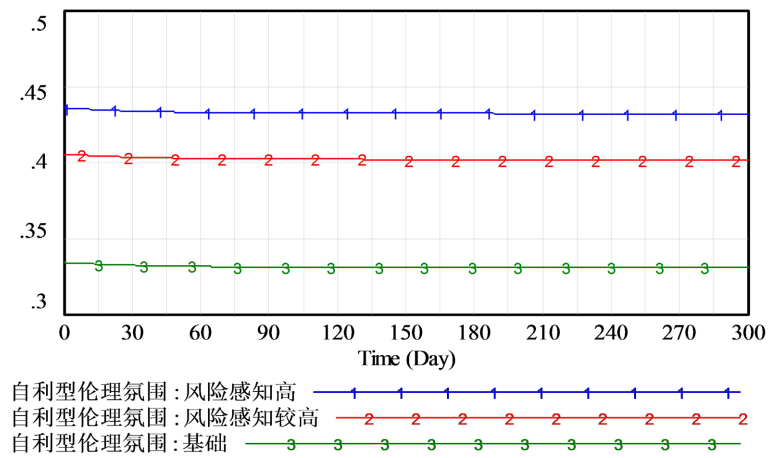


Figure 12. Self-interest climate
图 12. 自利型伦理氛围

道德推脱

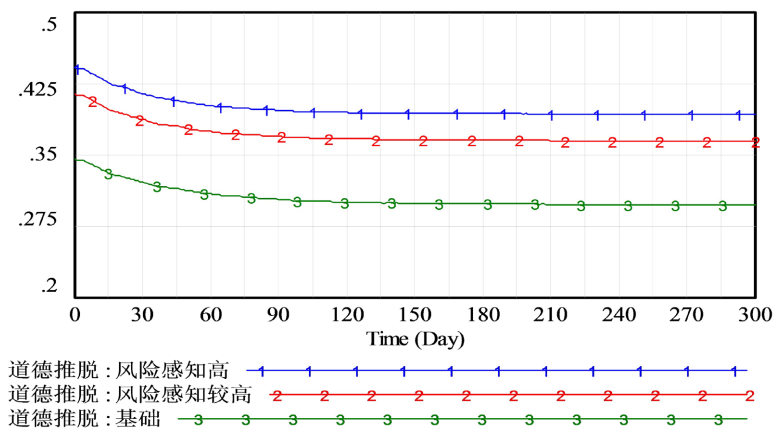


Figure 13. Moral disengagement
图 13. 道德推脱

UPB行为增加

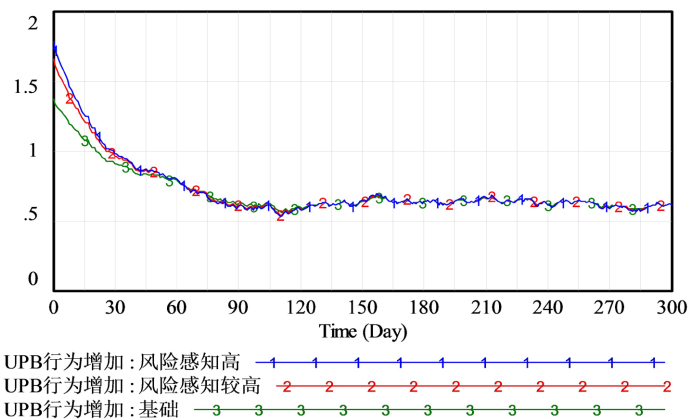


Figure 14. UPB increment

图 14. UPB 增量

亲组织非伦理行为

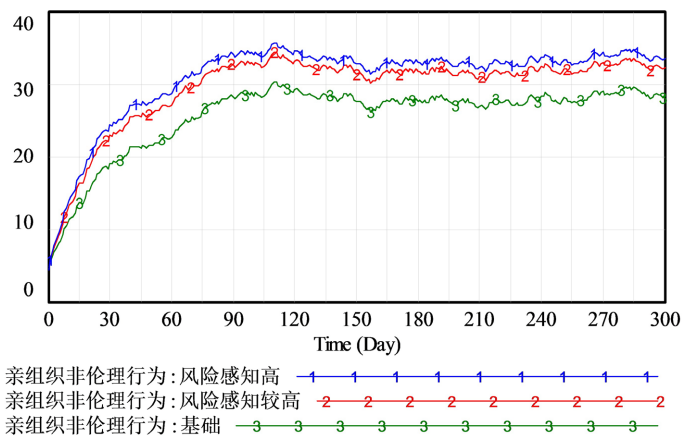


Figure 15. UPB

图 15. UPB

关系亲密度

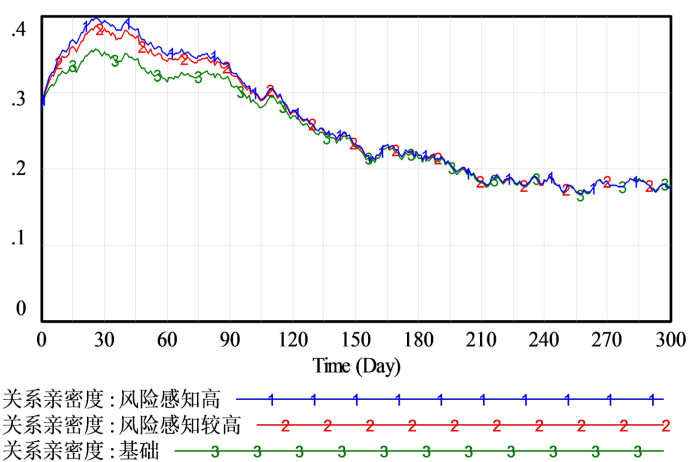


Figure 16. Relationship closeness

图 16. 关系亲密度

领导-成员交换关系质量

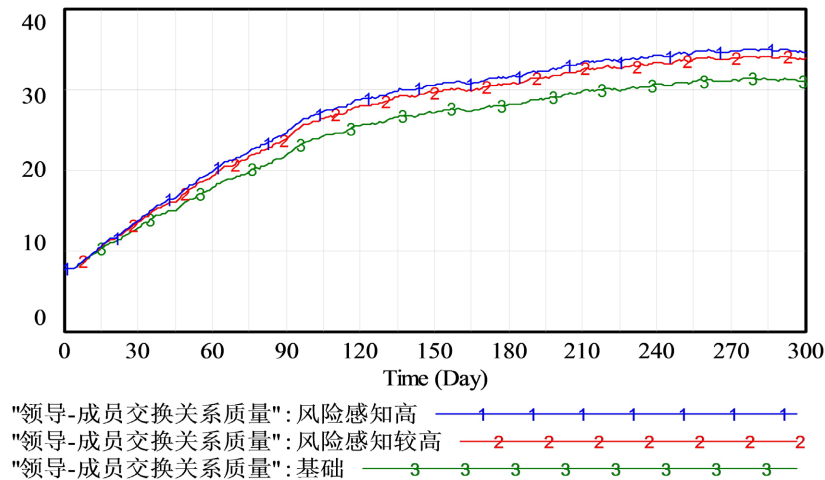


Figure 17. Leader-member relationship

图 17. 领导成员交换质量

观察三种风险下亲组织非伦理行为变化情况可知,三种情境下,亲组织非伦理行为的趋势并未变化,差异在于亲组织非伦理行为的绝对值。风险越高,工作不安全感(如图 10)、竞争氛围会随之增加(如图 11),自利型伦理氛围值(如图 12)也越高,从而导致道德推脱值(如图 13)也越高,因此,UPB 增加也呈现与替代风险正相关的变化。在前 45 天,高替代风险下 UPB 增加大于中等风险大于低风险。但是随着时间推进,无论起初感受到的替代风险高低,逐渐都会形成对人工智能替代风险相对客观的判断,之后基于风险感知从事 UPB 的程度则与个体对自身竞争力以及劣势的评估相关,与风险感知高低程度无关,因此,UPB 增加值在三种状态下会趋于一致(如图 14),而亲组织非伦理行为的水平差异也是由前 45 天的 UPB 增加值不同所造成的。

观察三种风险下领导成员交换关系质量变化情况可知,领导成员交换关系质量的趋势并未变化,差异在于领导成员交换关系质量的绝对值。领导成员交换关系质量之间的差值又来源于于关系亲密度的不同。在前 45 天,由于风险水平越高,亲组织非伦理行为水平越高(如图 15),从而与上级的亲密度也越高(如图 16)。因此,在前期,替代风险感知越高,会积累较多的关系亲密度,领导成员交换关系质量水平也越高(如图 17)。此外,可能由于员工间博弈一直存在以及组织内氛围在较长时间内会维持稳定状态,既定替代风险感知下,竞争氛围与自利型伦理氛围在 300 天内基本没有变化(如图 11、图 12),而员工工作不安全感下降并维持稳定(如图 10),因此通过道德推脱的作用,UPB 增加也减少并保持稳定(如图 14),再通过 UPB 水平影响关系亲密度,从而使领导成员交换关系质量二百天后的趋势保持一致,平行运行(如图 17)。

5. 结论

在一定时空范围内,亲组织非伦理行为是一个多因素影响的动态系统。工作不安全感、竞争氛围、自利型伦理氛围、道德推脱、领导成员交换关系以及亲组织非伦理行为之间存在因果关系。员工感知人工智能风险后的较短时间内(模型中前 45 天左右),亲组织非伦理行为会发生显著的增长,随后(模型中 160 天左右)会以后保持相对稳定的状态。在不同的人工智能替代风险感知程度下,员工的亲组织非伦理水平不同,且替代风险感知越高,亲组织非伦理水平整体越高。具体而言,差异主要来源于感知到替代风险前期(模型中前 45 天内)亲组织非伦理行为增加程度的不同。

参考文献

- [1] 徐鹏, 徐向艺. 人工智能时代企业管理变革的逻辑与分析框架[J]. 管理世界, 2020, 36(1): 122-129+238.
- [2] Umphress, E.E., Bingham, J.B. and Mitchell, M.S. (2010) Unethical Behavior in the Name of the Company: The Moderating Effect of Organizational Identification and Positive Reciprocity Beliefs on Unethical Pro-Organizational Behavior. *Journal of Applied Psychology*, **95**, 769-780. <https://doi.org/10.1037/a0019214>
- [3] 陈文晶, 康彩璐, 杨玥, 万岩. 人工智能潜在替代风险与员工职业能力发展: 基于员工不安全感视角[J]. 中国人力资源开发, 2022, 39(1): 84-97.