

类人智能机器人信任对员工工作贡献意愿的影响：基于人机关系信任的归维考察*

唐小飞^{1†} 王昌梅^{1†} 孙晓东¹ 张恩忠²

(¹ 西南财经大学工商管理学院, 成都 611130) (² 中国人民大学商学院, 北京 100872)

摘要 类人智能机器人(Humanoid Intelligent Robots, HIR)不仅带给人类情感体验, 而且正在重塑人机关系信任。本文试图从独特代理假说视角探究组织关系情境中的新型人机关系信任的形成与影响。1 个定性访谈研究和 3 个定量研究发现, 组织关系情境中的新型人机关系信任可以分为 HIR 情感修缮信任和 HIR 功能辅助信任两个维度。自我修缮和友情修缮两种人机协作类型更能显著激发 HIR 情感修缮信任; 智力辅助和体力辅助两种人机协作类型更能显著激发 HIR 功能辅助信任。相比 HIR 功能辅助信任, HIR 情感修缮信任更有利于提升组织温暖感知, 进而更有效促进员工工作贡献意愿; 反之, 相比 HIR 情感修缮信任, HIR 功能辅助信任更有利于提升组织能力感知, 进而更有效促进员工工作贡献意愿。最后, 将 HIR 代理需求导向作为情境变量引入研究, 验证了其在人机关系信任与组织温暖-能力感知之间的调节作用。

关键词 类人智能机器人, 人机关系信任, 情感修缮, 功能辅助, 员工工作贡献意愿

分类号 B849: C93

1 前言

伴随着劳动力市场竞争加剧, 人工智能对职业场景、组织形态和工作方式的重塑产生了深刻影响(Kellogg et al., 2020)。组织关系中的社交恐惧(侯娟等, 2021)、职业焦虑(谢宝国, 龙立荣, 2008; Przybylski et al., 2013)、工作倦怠(谢敏等, 2022)也愈发普遍, 进而削弱了员工工作贡献意愿。工作贡献意愿源于 van Scotter 和 Motowidlo (1996)的研究, 指员工通过自律性与内驱性行为支持组织目标的心理倾向。例如, 倾向在工作中投入额外的时间和精力, 发挥主观能动性, 坚持完成困难任务(Grant, 2008; van Scotter & Motowidlo, 1996)。Molina 等(2014)认为工作贡献意愿反映了员工与组织之间的交换关系, 即员工通过贡献换取所需的组织资源。积极的员工工作贡献意愿有利于提升工作绩效(Tisu et al., 2020)、忠诚度(Memon et al., 2018)以及员工

对组织的承诺(Obschonka et al., 2023)。因此如何激发员工工作贡献意愿, 历来是组织行为学关注的焦点问题(韦慧民, 龙立荣, 2009), 而在人机融合发展的组织关系情境中更是一个亟待探讨的领域。

人工智能技术的突破加速了机器人在人类组织中的嵌入进程(Lu et al., 2019)。HIR 作为一种新兴的机器人形态, 由于拥有人类外貌、心智感知和社交功能等特性(Brengman et al., 2021), 正在悄然融入人类组织(Tang et al., 2023), 如 HIR 代理对员工进行心理辅导咨询、自动化工作辅助等, 并影响着员工职业态度(Gati & Kulcsar, 2021)。比如有研究表明, 人机协作有助于员工提升工作热情、生产力(Kong et al., 2023), 并具有更积极的工作状态(Tang et al., 2023)。然而, 也有研究表明组织引进 HIR 代理可能会造成员工的工作不安全感、职场物化(许丽颖等, 2024; Yam et al., 2023)、孤独感(Tang et al., 2023)和离职倾向(Li et al., 2019)。对此, 有学

收稿日期: 2023-09-04

* 国家社会科学基金西部项目(22XGL017)阶段性成果。

† 唐小飞和王昌梅为共同第一作者

通信作者: 孙晓东, E-mail: sunxd@swufe.edu.cn

者指出人工智能技术能否成功融入到组织关系中取决于员工对人工智能技术的信任(Chowdhury et al., 2022; Glikson & Woolley, 2020)。这表明组织如何引导人机协作以激发人机关系信任,对重塑员工工作贡献意愿至关重要。综上,本研究将着重探讨以下问题:在新型组织关系情境中的人机关系信任维度是什么?不同类型的人机协作能否激发人机关系信任?以及人机关系信任如何重塑员工工作贡献意愿?其边界条件是什么?

1.1 人机关系信任的维度与形成

人机关系信任通常指在不确定或脆弱感知情境下,人类对机器人代理帮助其实现目标的积极态度(王晨 等, 2024; de Visser et al., 2016)。独特代理假说(Unique Agent Hypothesis)认为,人-人信任和人-机信任本质上是不同的(de Visser et al., 2016; Hoff & Bashir, 2015)。例如,在表现形态上,与人-人之间的双向信任不同,人-机信任是单向的(Hoff & Bashir, 2015);在建构基础上,人-人之间的信任主要基于互动关系形成,人-机信任主要基于算法和数据形成(何贵兵 等, 2022);在信任感知上,尽管人-人和人-机信任均涉及能力、善意和诚信三个维度,但由于人类更强的社会属性,使得两者在这三个维度上存在显著差异。即便提高机器人的拟人化程度,也不能完全消除这些差异(Alarcon et al., 2023)。这也表明人机关系信任可能比以往认为的更为复杂和特殊。

进一步,不同学者基于协作情境对人机关系信任的分类并不完全相同。例如, Lewis 等(2018)将人-机协作分为基于绩效的协作(人类如何影响/控制机器人执行任务)和基于社交的协作(机器人行为如何影响人类情感)。前者引发 HIR 认知信任,即对机器人能力等理性属性的依赖意愿;后者促进 HIR 情感信任,即依赖机器人带来的安全舒适感(Chien et al., 2024)。这种信任分类也适用于智能推荐代理(Komiak & Benbasat, 2006)。最近,唐小飞等(2021)将人机关系分为社会存在价值(自我、亲情和友情)替代和社会功能价值(智力和体力)替代,并探讨其对信任的影响。上述研究为人机协作和信任的分类提供了广泛框架。基于此,本文在新型组织关系情境下提出自我修缮、友情修缮、智力辅助和体力辅助 4 类人机协作,有可能激发 HIR 情感修缮信任和 HIR 功能辅助信任。

研究表明,员工常常通过自我情感调节及社交互动情绪调节,以适应内外部情境的需要(Goldenberg

et al., 2021),但其作用有限。组织可引进 HIR 代理协助员工完成自我修缮和友情修缮,从而激发 HIR 情感修缮信任。具体而言,自我修缮指 HIR 代理修缮员工因低成就动机和自我效能感等个体内在心理因素引发的情绪缺陷,如自我怀疑或否定。友情修缮则指 HIR 代理缓解员工因上下级或同事等外在因素产生关系冲突,进而导致的负面情绪,如对上级的不满和抱怨,对同事的厌恶和冷漠(Lewis et al., 2019)。上述研究为 HIR 情感修缮信任的概念形成提供了支持。由此,HIR 情感修缮信任是 HIR 对员工在组织中展现出的情感缺失和关系冲突进行修缮或满足而带来的信任。类似地,HIR 代理可通过智力辅助和体力辅助促进 HIR 功能辅助信任(陈明生, 2019)。具体来说,智力辅助指 HIR 代理提供计算、分析等更具创造性的脑力支持;体力辅助指 HIR 代理提供打印、收发快递、搬运等更具重复性的体力支持。由此,HIR 功能辅助信任是组织依托 HIR 代理辅助员工完成功能性任务而形成的人机关系信任。基于上述文献,本文认为自我修缮和友情修缮更显著地激发 HIR 情感修缮信任;同理,智力辅助和体力辅助更显著地促进 HIR 功能辅助信任。从而推测,4 种人机协作类型对情感修缮信任和功能辅助信任的激发强度存在显著差异。综上,提出如下假设:

假设 1: 组织关系情境下,HIR 情感修缮信任和 HIR 功能辅助信任是人机关系信任的两个不同维度,其中自我修缮和友情修缮激发的 HIR 情感修缮信任均显著强于 HIR 功能辅助信任,而智力辅助和体力辅助激发的 HIR 功能辅助信任均显著强于 HIR 情感修缮信任。

1.2 人机关系信任对员工工作贡献意愿的影响

员工倾向将组织代理人所采取的行动视为组织意图的表现(Rhoades & Eisenberger, 2002)。HIR 代理的 4 种人机协作活动,也可能被员工视为组织意图的体现。即为了表达对员工福祉的关心和利益保护,组织期望通过 HIR 代理向员工提供情感修缮支持和功能辅助支持。根据组织支持理论,虽然组织给予帮助能使员工产生组织支持感,但其强度取决于员工对帮助行为的认同度(Rhoades & Eisenberger, 2002)。比如,研究发现,引入 HIR 代理可能引发员工的职业危机感,加剧工作不安全感、工作倦怠和不道德行为等消极影响(许丽颖 等, 2024; Yam et al., 2023)。然而,当员工意识到 HIR 代理能够为其提供帮助时,类似职业危机的风险感知会被削弱,从

而增进对 HIR 的信任(唐小飞 等, 2021; Omrani et al., 2022; Siau & Wang, 2018)。因此, 在组织关系情境中, 人机关系信任意味着员工认同组织通过 HIR 代理为其提供支持。逐渐地, HIR 不再只是工具, 真正成为员工的同事, 甚至扮演管理者的角色(何贵兵 等, 2022; Chien et al., 2024; Wynne & Lyons, 2018)。

组织中的人际信任能够激发信任方的积极情绪、态度和行为, 如增强工作动机以及投入额外的时间精力(贡喆 等, 2017; 李宁 等, 2006)。那么, 组织中的人机关系信任也可能激发员工以积极饱满的活力投入工作, 表现出更强的工作贡献意愿。Bakker 等(2012)和胡巧婷等(2020)提出类似观点, 组织为员工提供更多的工作资源(如人际互动和学习机会), 能有效促进员工实现自我价值。为了回报组织, 员工则表现出更高的工作贡献意愿。另一方面, 信任 HIR 有利于员工更清晰地理解 HIR 对其工作的影响, 以及双方的角色和任务。这极大减少对未来预期的模糊性和负面认知, 从而提高与工作贡献意愿相关的积极心理结果(Chowdhury et al., 2022; Wang et al., 2023)。总之, 人机关系信任有助于提升员工工作贡献意愿(Chowdhury et al., 2023; Kong et al., 2023)。然而, 更关键的是, 亟需厘清 HIR 情感修缮信任和 HIR 功能辅助信任作为人机关系信任的两个不同维度, 究竟如何提升员工工作贡献意愿?

组织支持理论认为, 员工倾向赋予组织人格化特征(Rhoades & Eisenberger, 2002)。在组织拟人化的基础上, 源于人际和群际互动的温暖-能力感知框架(Fiske et al., 2002)同样适用于对组织的综合评价(Aaker et al., 2010)。当员工认为组织为其提供的情感修缮和功能辅助支持是值得信任的, 就可能引发对组织温暖-能力感知的积极评价(Fiske et al., 2002; Yu et al., 2022)。即 HIR 情感修缮信任和 HIR 功能辅助信任能够提升员工的组织温暖感知和组织能力感知。一方面, 组织温暖感知反映了组织在满足个体建立和维持社会联系需求中体现的友好、乐于助人等意图(Fiske, 2018)。当员工相信组织能通过 HIR 代理满足其社会联系方面的需求, 如提供情感安慰(Gaudiello et al., 2016), 弥补情感缺失, 克服社交恐惧(侯娟 等, 2021), 员工会有更强的组织温暖感知。同时 HIR 情感修缮信任意味着员工感受到来自组织的关怀。一个能让员工感受到关怀的组织往往也具有更高的温暖感知(Aaker et al., 2010)。另一方面, 组织能力感知反映的是对组织实

现目标的程度与可能性的感知(Fiske, 2018)。当员工相信组织能通过 HIR 代理提供功能辅助支持, 帮助其实现工作目标, 例如辅助管理仓库、接待客户(Willems et al., 2023), 员工更倾向评价组织是有能力的。此外, HIR 代理提供算法等功能辅助支持(丁晓东, 2020), 能提升员工的自我效能感与完成任务时的自我评价(徐樟, 刘艺璇, 2021)。这有助于促进更高效的人机协作实现组织目标(Harris-Watson et al., 2023), 从而增强员工对组织能力的认可。还有研究证明, 提供人性关怀(vs. 注重绩效和效率)的组织有更强的组织温暖感知; 相反, 注重绩效和效率(vs. 提供人性关怀)的组织有更强的组织能力感知(Aaker et al., 2010)。因此, 员工相信组织通过 HIR 代理提供情感修缮支持(vs. 功能辅助支持), 带来更强的组织温暖感知; 反之, 组织能力感知更强。综上, 提出以下假设:

假设 2a: 相比 HIR 功能辅助信任, HIR 情感修缮信任能更明显提升员工的组织温暖感知;

假设 2b: 相比 HIR 情感修缮信任, HIR 功能辅助信任能更明显提升员工的组织能力感知。

对组织的温暖-能力感知评价可以塑造个体在组织中的情感和行为反应(Cuddy et al., 2011)。在社会认知中, 温暖感知维度用于评估他人的意愿是积极的还是消极的; 而能力感知维度用于评估他人根据意愿采取行动的能力(Cuddy et al., 2007; Fiske et al., 2002; Yu et al., 2022)。由 HIR 功能辅助信任(HIR 情感修缮信任)触发的组织温暖-能力感知评价, 使员工意识到组织不仅有积极的意愿去提供帮助, 还有能力实现这些意愿。组织支持理论指出, 当组织以友好的方式支持帮助员工时, 员工也会以互惠规范回报组织(Kurtessis et al., 2017; Rhoades & Eisenberger, 2002)。因此组织温暖-能力感知越强, 就越能有效地促进员工产生积极的工作贡献意愿回报组织。综上, 提出如下假设:

假设 3a: 相比 HIR 功能辅助信任, HIR 情感修缮信任通过组织温暖感知而正向作用于员工工作贡献意愿的效应更大;

假设 3b: 相比 HIR 情感修缮信任, HIR 功能辅助信任通过组织能力感知而正向作用于员工工作贡献意愿的效应更大。

1.3 HIR 代理需求导向的调节作用

对 HIR 代理的评价不仅取决于 HIR 代理的特性, 还取决于人类的需求和偏好, 并且不同的需求和偏好会对 HIR 代理的信任与价值感知等结果产

生不同的影响(Blut & Wang, 2020)。人类的需求偏好可分为互动需求导向和任务需求导向(Köhler et al., 2011)。在特定的组织关系情境下,员工对HIR代理的互动需求导向是指员工与HIR代理交互愿望的一般倾向(Dabholkar, 1996; Yoganathan et al., 2021);而任务需求导向则是指员工与HIR代理协同实现任务目标的一般倾向(Köhler et al., 2011; Shamim et al., 2023)。进一步有研究表明,HIR互动需求导向高的员工更注重HIR带来的社会存在价值(Chandra et al., 2022)。尤其是缺乏社交联系特征的个体,如存在长期孤独、社会脱节、依恋焦虑等个体,更倾向HIR能够提供情感修缮支持弥补自身的情感缺陷(Bartz et al., 2016)。如果HIR互动需求导向高(vs. 低)的员工认为HIR代理提供的情感修缮支持值得信任,将会有更强的组织温暖感知。这可能增强HIR情感修缮信任(vs. HIR功能辅助信任)对组织温暖感知的积极效应。同时,依据任务技术适配理论(Crolic et al., 2022),由于员工对技术的认知心理和认知行为不够契合,互动需求导向可能会削弱HIR功能辅助信任(vs. HIR情感修缮信任)对组织能力感知的积极影响。另一方面,HIR任务需求导向高的员工更关注个体的胜任需要,人工智能对程序化、事务化工作的快速处理(Köhler et al., 2011)。这类员工更倾向HIR能够提供功能辅助支持,从而更加专注于核心的工作任务并提升工作绩效(Shamim et al., 2023)。如果HIR任务需求导向高(vs. 低)的员工认为HIR代理提供的功能辅助支持值得信任,将会有更强的组织能力感知。这可能导致HIR功能辅助信任(vs. HIR情感修缮信任)对组织能力感知的积极效应增强。但HIR代理任务需求导向不能满足人际互动准则,一定程度上会削弱HIR情感修缮信任(vs. HIR功能辅助信任)对组织温

暖感知的积极影响。综上,提出如下假设:

假设 4: HIR 代理需求导向在人机关系信任与组织温暖-能力感知之间起调节作用。

假设 4a: 随着员工的 HIR 代理互动需求导向增强,HIR 情感修缮信任(vs. HIR 功能辅助信任)对组织温暖感知的影响效应增强,HIR 情感修缮信任(vs. HIR 功能辅助信任)对组织能力感知的影响效应减弱。

假设 4b: 随着员工的 HIR 代理任务需求导向增强,HIR 功能辅助信任(vs. HIR 情感修缮信任)对组织能力感知的影响效应增强,HIR 功能辅助信任(vs. HIR 情感修缮信任)对组织温暖感知的影响效应减弱。

1.4 研究概述

本研究由 1 个质性研究和 3 个定量研究组成,主要研究目的是厘清组织关系情境下人机关系信任的维度与形成,探究人机关系信任如何影响员工工作贡献意愿及边界条件。具体而言,研究 1 拟运用质性研究法归纳出 2 个人机关系信任维度和 4 种人机协作类型,并运用情境实验法验证 4 种人机协作类型能不同程度地激发HIR情感修缮信任和HIR功能辅助信任。研究 2 拟采用情境实验法验证组织温暖感知和组织能力感知的中介作用。研究 3 拟采用情境实验法探讨人机关系信任提升组织温暖-能力感知的边界条件,即验证两种HIR代理需求导向的调节作用。研究概念模型如图 1。

2 研究 1: 人机关系信任的维度与形成

2.1 质性研究

2.1.1 访谈对象与内容

为了初步探索组织关系情境中人机关系信任

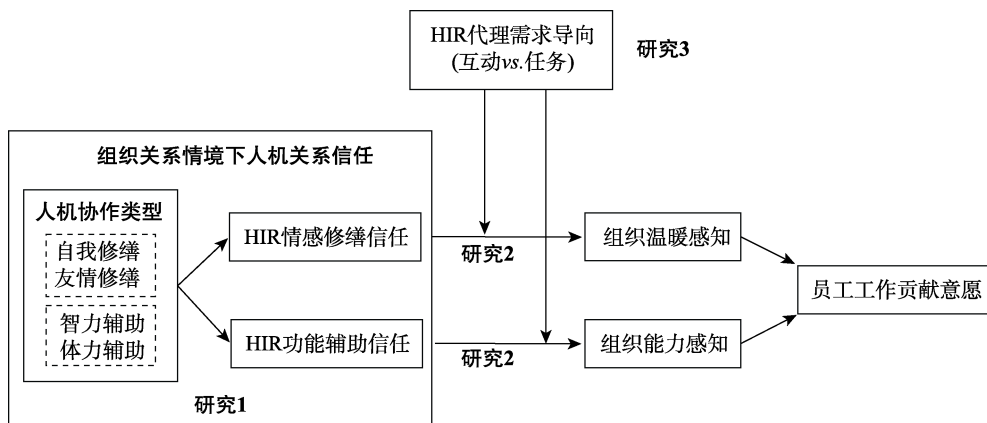


图 1 研究概念模型

的维度和形成, 研究 1 首先采用质性研究法邀请访谈对象参与半结构深度访谈。现有研究指出, 一项半结构深度访谈的样本可以小, 但必须多样化; 8 个访谈对象每人 30~60 分钟的访谈内容可以成为充分的样本依据(邓新明, 2014; 曹剑琴 等, 2023)。遵循上述建议, 本研究受邀访谈的对象共 9 名(4 名男性, 5 名女性), 分别来自金融、美容整形、通信制造、房地产、工程投资建设和教育行业。他们中既有中高层管理人员, 也有基层工作人员和新入职员工, 平均从业 8.56 ($SD = 6.13$) 年, 能够充分表达有关“HIR 嵌入组织关系情境中人机关系信任”的看法。访谈内容主要分为三个部分。第一部分主要询问访谈对象在过去的工作经历中面临的问题以及经历过的负面情绪、情感, 并详细描述。第二部分询问访谈对象对 HIR 等人工智能的了解程度、偏好, 以及其对 HIR 在工作中的应用期望和信任程度。第三部分收集访谈对象的基本信息。

2.1.2 数据分析与结果

通过协商一致的质的研究方法(Consensual Qualitative Research, CQR)对数据进行分析(朱旭, 江光荣, 2011; Hill et al., 2005)。CQR 核心思想是通过邀请多名分析员组成研究小组, 引入多元视角, 对访谈中收集的文本数据进行定性分析(Hill et al., 2005)。CQR 主要包括“划分域”、“提炼核心观点”、“交叉分析”和“外部审核”四个分析步骤。

CQR 结果发现, 与人机关系信任相关的内容可分为 HIR 情感修缮信任和 HIR 功能辅助信任 2 个维度(域), 具体为自我修缮、友情修缮、智力辅助、体力辅助 4 种人机协作(类)。其中, HIR 情感修缮信任维度(域)涉及自我修缮和友情修缮 2 种人机协作(类)。在访谈中, 有访谈对象相信 HIR 能缓解长期工作导致的麻木倦怠: “每天兼顾运营相关的各种各样的杂事让我觉得自己像是大内总管、维修工, 丧失工作热情、感觉麻木, 通过机器人的心理辅助治疗, 我相信能有效提升我的工作热情和活力”。这不仅与自我修缮的定义较为一致, 还间接揭示了自我修缮可以激发 HIR 情感修缮信任。还有访谈对象相信 HIR 能缓解部门、领导冲突导致的人际负面情绪: “平级两个部门之间或者两个单位之间因责任不明晰, 流程不明确产生推诿, 让我对推诿部门的人员产生对立情绪, 采用机器人作为沟通的桥梁, 能够有效缓解这种冲突带来的负面情绪”。这类消极情绪、情感往往伴随着组织中的人际关系破裂, 对“职场友情”带来严峻考验。访谈对象也从

不同程度表明, 职场中的 HIR 或许比人类同事更适合作为朋友, 愿意依赖 HIR 而不是人类。这些观点不仅与友情修缮的定义较为契合, 还间接说明友情修缮能够促进 HIR 情感修缮信任。

HIR 功能辅助信任维度(域)涉及智力辅助和体力辅助 2 种人机协作(类)。在访谈中, 有访谈对象相信 HIR 可以提供文案写作方面的智力辅助: “将来, 机器人可能陪同我一起完成投资项目的研判, 甚至它还能帮我进行投资陈述, 所以我认为他可以成为像人类一样的工作伙伴”。也有访谈对象相信 HIR 可以提供高强度、高重复甚至有安全隐患的体力辅助: “将来一些工作强度大、危险性高, 或者一些特殊的工作环境可以依赖机器人, 比如桩基础的浇灌, 人工操作会存在一定的危险性”。这些观点与智力辅助和体力辅助的定义高度吻合的同时, 也间接论证了智力辅助和体力辅助能够带来功能辅助信任。完整的质性研究内容和分析结果详见网络版附录 1。

2.1.3 讨论

通过访谈, 我们发现员工在工作中面临的问题可以分为自我修缮、友情修缮、智力辅助和体力辅助 4 类。建立在未来 HIR 兼具智商、情商和物理能力的前提下, 访谈对象相信在组织中这 4 类问题都可以通过 HIR 的情感修缮和功能辅助支持得到不同程度的改善。但是也有部分访谈对象在访谈中表达了对 HIR 消极或中立的观点。例如有访谈对象表示: “我感觉让 HIR 提供情绪价值好像有点难, 因为我始终会觉得它是一个机器人、程序化的电脑、没有感情, 除非它真的拥有了人类的情感, 变成了一个真正的人, 才能够真正的去理解我, 并与我共情”。事实上, 随着人机交互中情感计算概念的提出(Picard, 2003), 有关探索人工情感(Novikova & Watts, 2015)和服务机器人情感交互(Röning et al., 2014)的研究日益增多。这说明 HIR 能否具备情感交互能力已经从技术问题演变成心理问题, 即愿不愿意相信 HIR 能够实现对员工的情感修缮。因此, 本文进一步设计了情境实验, 验证自我修缮、友情修缮、智力辅助和体力辅助 4 种人机协作能否激发员工对 HIR 产生情感修缮信任和功能辅助信任。

2.2 刺激物设计

为进一步探究人机关系信任的维度与形成, 本文根据人机协作的定义设计了 4 段文字材料作为刺激材料, 分别以 HIR 作为第一人称来描述组织关系情境中 HIR 与员工之间的 4 种人机协作。为了检验

4 段文字材料的代表性, 预测试以线下便利抽样的形式招募了 192 名 A 银行员工, 参与者被随机分配阅读 4 段刺激材料中的 1 段。然后让参与者评估以下 4 个题项作为人机协作类型的操纵检验(1 = “完全不同意”, 7 = “完全同意”): “在上述特定情境中, (1)你在多大程度上同意 HIR 能实现员工的自我修缮, 即在多大程度上相信 HIR 能缓解成就动机低、自我效能感不足等员工个体的心理因素导致的情感缺陷; (2)你在多大程度上同意 HIR 能实现员工的友情修缮, 即在多大程度上相信 HIR 能缓解同事间关系冲突等外在因素带来的员工个体负面情绪/情感; (3)你在多大程度上同意 HIR 能实现员工的智力辅助, 即在多大程度上相信 HIR 能提供计算、统计、分析等方面的智力支持; (4)你在多大程度上同意 HIR 能实现员工的体力辅助, 即在多大程度上相信 HIR 能提供诸如打印、送文件、收发快递、搬运物品等体力辅助”。数据分析结果表明, 4 段文字刺激材料具有较好的代表性, 具体分析结果详见网络版附录 2。

刺激材料增加 HIR 的外形拟人化能够在一定程度上增进信任(Waytz et al., 2014), 但过度拟人化也会诱发恐怖谷效应(Mori, 1970), 使积极态度急转直下。为了避免潜在的混淆因素, 在接下来的预实验中, 本文选取拟人化外形不同的两种 HIR——Nao V6 和 Sophia 作为备选项¹ (如图 2), 并检验哪一种 HIR 更具代表性。将 Nao V6 和 Sophia 的图片分别与上述 4 段文字材料进行匹配, 最终形成 8 张刺激图片(详见网络版附录 3)。其中 4 张为 Nao V6 以第一人称阐述 4 种人机协作, 另外 4 张以 Sophia 为第一人称进行阐述。虽然, 两者外形存在差异, 但 Nao V6 和 Sophia 所阐述的内容没有差异。

预实验通过线下便利抽样的方式一共招募 205 名 B 银行员工, 剔除掉未通过注意力检查、极端值和回答过于一致的参与者后, 最终获得 182 份有效数据。参与者被随机分配到 Nao V6 组或 Sophia 组, 然后按照随机顺序向参与者展示 4 张 Nao V6 (或 Sophia)以第一人称描述的组织中人机协作的刺激图片。随后, 参与者被要求回答与预测试相同的人

机协作类型操纵检验题项以及对 HIR 的恐慌不适感(Mende et al., 2019)。数据分析结果表明, 首先 Nao V6 组和 Sophia 组的参与者对自我修缮、友情修缮、智力辅助和体力辅助 4 张刺激图片的评价均不存在显著差异。其次, 虽然本研究选取的偏卡通 HIR (Nao V6)和偏真人 HIR (Sophia)均不会引发恐怖谷效应, 但是相较于 Nao V6, Sophia 的 HIR 外形让参与者有更高的恐慌不适感。进一步发现 Sophia 组的方差波动更大, 19.8%的参与者对 Sophia 外形的恐慌不适感评分高于中值, 而 Nao V6 组中仅有 4.4%。随着样本量的扩大, 这种波动可能会使 Sophia 组整体均值上升。此外, Nao V6 为中性性别, 可以避免性别特征偏差影响测试的准确性(Desideri et al., 2018)。综上, 为了降低 Sophia 的 HIR 外形导致的人机关系信任偏差, 对研究结果稳健性产生影响, 研究 1 情境实验选取 Nao V6 组刺激图片作为刺激物, 以便验证自我修缮和友情修缮主要激发 HIR 情感修缮信任, 而智力辅助和体力辅助则主要激发 HIR 功能辅助信任。预实验研究内容和分析结果详见网络版附录 4。



图 2 拟人化外形不同的两种 HIR——Nao V6 (左)和 Sophia (右)

2.3 情境实验

2.3.1 参与者

采用 G*Power 3.1 软件(Faul et al., 2007)计算本实验所需样本量, 对于本实验适用的单因素方差分析, 取中等效应量 $f = 0.25$, 显著性水平 $\alpha = 0.05$, 组数为 4, 要达到 90%的统计检验力至少需要 232 名参与者。实验通过线下便利抽样的方式, 以 C 银行下辖的三家支行员工为对象, 招募了 400 名在职人员参与实验, 筛除未通过注意力检查、极端值和回答过于一致的参与者后, 确认了 357 名在职人员, 其中自我修缮组 88 人、友情修缮组 89 人、智力辅

¹ Nao V6 是由软银机器人品牌(Soft Bank Robotics)研发的一款具有类人外形的机器人, 能够听、看、说, 以及与人互动。Sophia 是由汉森机器人技术公司(Hanson Robotics)开发的类人机器人, 是历史上首个获得公民身份的一台机器人, 拥有橡胶皮肤以及面部表情, 能够理解人类语言与人交流互动。Nao V6 和 Sophia 均是市场上颇具代表性的 HIR。

助组和体力辅助组各 90 人。包括男性 182 名(51%), 女性 175 名(49%), 平均年龄 $M = 30.61$ ($SD = 8.45$) 岁。参与者的学历包括, 专科及以下 60 名(16.8%), 本科 244 名(68.3%), 硕士 49 名(13.7%), 博士 4 名(1.1%)。所有参与者均自愿参加实验且知情同意, 并在实验结束之后获得相应实验礼品。

2.3.2 实验设计与程序

参与者被随机分配到 4 个人机协作组(自我修缮组 vs. 友情修缮组 vs. 智力辅助组 vs. 体力辅助组)中的任意一组。首先告知所有参与者: 有朝一日, 当 HIR 融入你的工作情境时, 人机协作将成为组织管理工作的一部分。然后参与者被要求观看预实验确定的刺激图片。自我修缮组观看以 Nao V6 为第一人称呈现的关于员工通过 HIR 完成自我修缮的图片。其余三组的情境操纵同理。

在观看完刺激图片后, 参与者被要求回答与预实验相同的人机协作类型操纵检验题项。随后参与者被要求填写关于 HIR 情感修缮信任和 HIR 功能辅助信任的测量题项(1 = “完全不同意”, 7 = “完全同意”), 所有的英文题项采用回翻程序翻译为中文题项。HIR 情感修缮信任($\alpha = 0.93$)的 5 个测量题项根据 McAllister (1995)以及 Levin 和 Cross (2004)量表改编而成。McAllister (1995)采取了 4 个题项测量基于情感的信任, 强调了“受托人的情感元素和社交技能”。本研究情境中的 HIR 除了富有情感和社交技能外, 还让员工相信 HIR 为其提供的情感修缮支持充满善意, 不是出于自利动机。这与 Levin 和 Cross (2004)在组织关系情境中采用的善意信任一致。因此, 本研究对基于情感的信任的 4 个测量题项进行了修正, 在原 4 个测量题项上增加了 1 个题项测量“员工对 HIR 善意的信任”。综上, HIR 情感修缮信任的测量题项前 4 个题项改编自 McAllister (1995), 最后一个题项改编自 Levin 和 Cross (2004), 具体包括: (1) “我和 HIR 是一种分享关系, 我们之间可以相互分享想法和感受”; (2) “当我情感上需要安抚时, 我可以自由倾诉, 并且知道 HIR 愿意倾听”; (3) “如果 HIR 从我的工作场景中离开, 我会有一种失落感”; (4) “如果我向 HIR 倾诉工作中遇到的问题, 我相信 HIR 会给予积极和有建设性的回应”; (5) “我相信 HIR 会像亲人或朋友一样关心我在工作中的遭遇”。

HIR 功能辅助信任($\alpha = 0.92$)的 4 个题项改编自 Levin 和 Cross (2004)以及 Komiak 和 Benbasat (2006)的研究。Levin 和 Cross (2004)从 McAllister (1995)

认知信任量表中选取了载荷因子最高的两个题项, 从工作态度和在工作表现两方面评估组织关系情境中员工对同事能力的认知信任程度。Komiak 和 Benbasat (2006)开发了两个新题项, 从工作经验和专业性两方面测量消费者对智能推荐代理能力的认知信任程度。在此基础上, 本研究从 HIR 工作态度、工作表现、工作经验和工作专业性四个方面综合评估组织关系情境中员工对 HIR 功能辅助能力的信任。具体包括: (1) “我相信 HIR 以专业和敬业的态度辅助我开展工作”; (2) “鉴于 HIR 以往的工作表现, 我没有理由怀疑 HIR 的辅助性工作能力”; (3) “HIR 在辅助方面有丰富的经验”; (4) “HIR 在辅助方面是一个专家”。最后, 参与者报告了性别、年龄和受教育程度三项人口统计学信息。

2.3.3 数据分析与结果

操纵检验。数据结果表明, (1)单样本 t 检验分析发现自我修缮组参与者在较高程度上同意 HIR 能实现自我修缮($M_{\text{自我修缮}} = 5.63$, $SD = 0.84$, $t(87) = 18.26$, $p < 0.001$, $d = 1.94$, 与中值 4 比较)。单因素方差分析的事后多重比较(Tukey)表明, 自我修缮组报告的自我修缮同意程度显著高于其他三组($M_{\text{自我修缮}} = 5.63$ vs. $M_{\text{友情修缮}} = 4.99$, Tukey-adjusted $p < 0.001$; $M_{\text{自我修缮}} = 5.63$ vs. $M_{\text{智力辅助}} = 4.49$, Tukey-adjusted $p < 0.001$; $M_{\text{自我修缮}} = 5.63$ vs. $M_{\text{体力辅助}} = 3.88$, Tukey-adjusted $p < 0.001$)。 (2)同理, 友情修缮组参与者在较高程度上同意 HIR 能实现友情修缮($M_{\text{友情修缮}} = 5.64$, $SD = 1.13$, $t(88) = 13.69$, $p < 0.001$, $d = 1.45$, 与中值 4 比较), 且显著高于其他三组报告的友情修缮同意程度($M_{\text{友情修缮}} = 5.64$ vs. $M_{\text{自我修缮}} = 4.93$, Tukey-adjusted $p = 0.002$; $M_{\text{友情修缮}} = 5.64$ vs. $M_{\text{智力辅助}} = 4.66$, Tukey-adjusted $p < 0.001$; $M_{\text{友情修缮}} = 5.64$ vs. $M_{\text{体力辅助}} = 3.80$, Tukey-adjusted $p < 0.001$)。 (3)智力辅助组参与者在较高程度上同意 HIR 能实现智力辅助($M_{\text{智力辅助}} = 6.01$, $SD = 0.76$, $t(89) = 25.21$, $p < 0.001$, $d = 2.64$, 与中值 4 比较), 且显著高于其他三组报告的智力辅助同意程度($M_{\text{智力辅助}} = 6.01$ vs. $M_{\text{自我修缮}} = 5.00$, Tukey-adjusted $p < 0.001$; $M_{\text{智力辅助}} = 6.01$ vs. $M_{\text{友情修缮}} = 4.60$, Tukey-adjusted $p < 0.001$; $M_{\text{智力辅助}} = 6.01$ vs. $M_{\text{体力辅助}} = 3.82$, Tukey-adjusted $p < 0.001$)。 (4)体力辅助组参与者在较高程度上同意 HIR 能实现体力辅助($M_{\text{体力辅助}} = 6.19$, $SD = 0.73$, $t(89) = 28.34$, $p < 0.001$, $d = 3.00$), 且显著高于其他三组报告的体力辅助同意程度($M_{\text{体力辅助}} = 6.19$ vs. $M_{\text{自我修缮}} = 3.51$, Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助}} = 6.19$ vs. $M_{\text{友情修缮}} = 3.38$,

Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助}} = 6.19$ vs. $M_{\text{智力辅助}} = 3.80$, Tukey-adjusted $p < 0.001$)。这说明实验操纵有效。

假设检验。配对 t 检验分析结果表明, 自我修缮组的 HIR 情感修缮信任显著高于 HIR 功能辅助信任($t(87) = 7.26, p < 0.001, d = 0.78$)。友情修缮组的 HIR 情感修缮信任显著高于 HIR 功能辅助信任($t(88) = 9.65, p < 0.001, d = 1.02$)。智力辅助组的 HIR 功能辅助信任显著高于 HIR 情感修缮信任($t(89) = -10.12, p < 0.001, d = 1.07$)。最后, 体力辅助组参与者报告的 HIR 功能辅助信任显著高于 HIR 情感修缮信任($t(89) = -14.43, p < 0.001, d = 1.52$)。综上, 假设 1 得到验证。此外, 单因素方差分析结果显示自我修缮、友情修缮、智力辅助、体力辅助 4 种协作类型的 HIR 情感修缮信任(HIR 功能辅助信任)存在显著差异。具体来说, 首先以组别(自我修缮 = 1, 友情修缮 = 2, 智力辅助 = 3, 体力辅助 = 4)作为因子, HIR 情感修缮信任作为因变量进行单因素方差分析。结果表明组间有显著差异, $F(3, 353) = 81.60, p < 0.001, \eta^2 = 0.41$ 。然后以组别(自我修缮 = 1, 友情修缮 = 2, 智力辅助 = 3, 体力辅助 = 4)作为因子, HIR 功能辅助信任作为因变量进行单因素方差分析。结果表明组间有显著差异, $F(3, 353) = 44.96, p < 0.001, \eta^2 = 0.28$ 。变量的描述性统计结果, 具体见表 1, 4 种人机协作类型的评价分布图如下(见图 3)。

表 1 研究变量的描述性统计结果(研究 1)

组别	均值		标准差	
	HIR 情感修缮信任	HIR 功能辅助信任	HIR 情感修缮信任	HIR 功能辅助信任
自我修缮组	5.82	4.97	1.01	1.03
友情修缮组	5.71	4.58	1.01	1.02
智力辅助组	4.53	6.11	1.16	1.02
体力辅助组	3.66	5.79	1.11	0.93

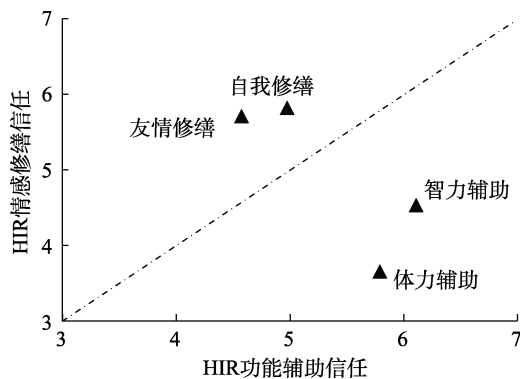


图 3 人机关系信任维度的评价分布图

2.3.4 讨论

研究 1 通过情境材料构建了未来组织关系情境中 4 种人机协作类型, 并通过配对 t 检验发现不同人机协作类型能够激发不同强度的 HIR 情感修缮信任和 HIR 功能辅助信任, 说明 4 种人机协作类型是人工智能时代组织关系发展的有益补充。这在一定程度上削弱了“技术替代论”和“技术威胁论”制造的社会焦虑。虽然研究 1 论证了 HIR 情感修缮信任和 HIR 功能辅助信任是人机关系信任的 2 个不同维度, 以及人机协作类型对 2 个信任维度的差异化影响, 但是并没有验证人机关系信任与员工工作贡献意愿的内在关系。因此研究 2 拟验证, HIR 情感修缮信任和 HIR 功能辅助信任如何通过组织温暖-能力感知影响员工工作贡献意愿。

3 研究 2: 组织温暖感知和组织能力感知的双中介作用

3.1 参与者

采用 G*Power 3.1 软件(Faul et al., 2007)计算本实验所需样本量, 对于本实验适用的独立样本 t 检验分析, 显著性水平 $\alpha = 0.05$ 且中等效应量($d = 0.5$), 为了达到 90%统计检验力至少需要 172 名参与者。研究 2 通过线下便利抽样的方式, 以 D 银行员工为对象, 招募了 200 名在职人员参与实验, 剔除掉未通过注意力检查、极端值和回答过于一致的参与者后, 最终获得 184 个有效样本, 其中 HIR 情感修缮信任组 92 名, HIR 功能辅助信任组 92 名。包括男性 94 名(51.1%), 女性 90 名(48.9%), 平均年龄 $M = 30.49$ ($SD = 7.55$)岁。参与者的学历包括, 专科及以下 15 名(8.2%), 本科 134 名(72.8%), 硕士 27 名(14.7%), 博士 8 名(4.3%)。所有参与者均自愿参加实验且知情同意, 并在实验结束之后获得相应实验礼品。

3.2 实验程序

研究 2 为单因素两水平组间实验设计(HIR 情感修缮信任 vs. HIR 功能辅助信任), 参与者被随机分配到其中一组。首先按照要求阅读人机关系信任的情境材料(见网络版附录 5)。在阅读材料时要求参与者想象: 如果这就是自己未来的工作场景, 自己对 HIR 会持有什么样的态度。此外, 为了加强操控效果, 参与者在阅读完情境材料之后还需要尽可能详细地描述, 在这样的工作场景中采用 HIR, 关于 HIR 情感修缮信任和 HIR 功能辅助信任方面的感受, 以及自己的工作贡献意愿, 时间和字数不限。随后, 参与者被要求回答与研究 1 相同的 HIR

情感修缮信任($\alpha = 0.93$)和 HIR 功能辅助信任($\alpha = 0.88$)测量题项, 作为研究 2 的操纵检验。然后要求参与者分别报告员工工作贡献意愿, 组织温暖感知和组织能力感知(1 = “完全不同意”, 7 = “完全同意”)。员工工作贡献意愿($\alpha = 0.77$)的 3 个题项是在 Madden 等(1992)以及 Song 和 Kim (2022)的量表基础上改编而成: “我相信在未来有更强的工作贡献意愿”; “我有可能在未来有更强的工作贡献意愿”; “我预测我会在未来有更强的工作贡献意愿”。组织温暖感知($\alpha = 0.91$)测量题项根据韦庆旺等(2018)量表改编而成, 其中 6 个词是典型的组织温暖特质词, 即通过组织提供 HIR 的行为使参与者在多大程度上认为组织是“乐于助人的”、“善解人意的”、“友善的”、“讨人喜欢的”、“可靠的”和“富有同情心的”。组织能力感知($\alpha = 0.80$)的 3 个双极测量题项根据 Scott 等(2013)改编, 即通过组织提供 HIR 的行为使参与者在多大程度上认为组织是“没有能力的/有能力的”、“不智慧的/智慧的”、“无适应性的/有适应性的”。

考虑到不同参与者对 HIR 的看法和知识可能存在差异, 从而影响员工工作贡献意愿, 为了排除混淆因素, 参与者还被要求报告他们对 HIR 的熟悉程度(“你对类人智能机器人有多熟悉?”, 从 1 = “完全不熟悉”到 7 = “完全熟悉”)、了解程度(“与普通入相比, 你认为你对类人智能机器人有多了解?”, 从 1 = “完全不了解”到 7 = “完全了解”)、喜爱程度(“你有多喜欢类人智能机器人?”, 从 1 = “完全不喜欢”到 7 = “完全喜欢”)和恐慌不适度。其中熟悉程度、了解程度和喜爱程度的题项改编自许丽颖等(2022)的研究, 恐慌不适度的题项与研究 1 的预实验相同。最后, 参与者报告了性别、年龄和受教育程度三项人口统计学信息。

3.3 数据分析与结果

操纵检验。独立样本 t 检验结果显示, HIR 情感修缮信任组报告的 HIR 情感修缮信任显著高于 HIR 功能辅助信任组($M_{\text{功能辅助组}} = 4.31, SD = 1.40, M_{\text{情感修缮组}} = 5.80, SD = 0.71, t(135.31) = 9.09, p < 0.001, d = 1.34$)。HIR 功能辅助信任组对 HIR 产生的 HIR 功能辅助信任显著高于 HIR 情感修缮信任组($M_{\text{功能辅助组}} = 5.81, SD = 0.80, M_{\text{情感修缮组}} = 4.83, SD = 0.98, t(182) = -7.48, p < 0.001, d = 1.10$)。因此, HIR 情感修缮信任和 HIR 功能辅助信任情境操纵成功。

组织温暖感知。假设 2a 提出相比 HIR 功能辅助信任, HIR 情感修缮信任能显著提升员工的组织

温暖感知。独立样本 t 检验结果显示, HIR 情感修缮信任组的组织温暖感知($M_{\text{组织温暖}} = 5.70, SD = 0.97$)显著高于 HIR 功能辅助信任组($M_{\text{组织温暖}} = 5.00, SD = 1.15; t(182) = 4.46, p < 0.001, d = 0.66$)。假设 2a 得到验证。

组织能力感知。假设 2b 提出相比 HIR 情感修缮信任, HIR 功能辅助信任能显著提升员工的组织能力感知。独立样本 t 检验分析显示, HIR 功能辅助信任组的组织能力感知($M_{\text{组织能力}} = 5.95, SD = 0.81$)显著高于 HIR 情感修缮信任组($M_{\text{组织能力}} = 4.99, SD = 0.93; t(182) = -7.51, p < 0.001, d = 1.10$)。假设 2b 得到验证。

双重中介假设检验。假设 3a (假设 3b)提出, 组织温暖感知(组织能力感知)在人机关系信任与员工工作贡献意愿之间起中介作用。对人机关系信任进行虚拟变量编码作为自变量(1 = HIR 功能辅助信任, 0 = HIR 情感修缮信任)。在 95% 的置信区间下, 采用 PROCESS 程序自主抽样 5000 次检验组织温暖感知和组织能力感知的中介效应(Model 4; Hayes, 2013)。Bootstrap 分析结果显示, 组织温暖感知的间接中介效应是显著负向的(indirect effect = $-0.30, SE = 0.08, 95\% CI [-0.45, -0.15]$), 组织能力感知的间接中介效应是显著正向的(indirect effect = $0.34, SE = 0.09, 95\% CI [0.19, 0.55]$)。为了进一步验证中介效应的稳健性, 把喜欢度、熟悉程度、了解程度和恐慌不适度纳入协变量再次进行 PROCESS (Model 4; Hayes, 2013)分析。Bootstrap 分析结果(见图 4)显示组织温暖感知(indirect effect = $-0.23, SE = 0.06, 95\% CI [-0.35, -0.12]$)和组织能力感知(indirect effect = $0.33, SE = 0.08, 95\% CI [0.19, 0.50]$)间接中介效应仍然显著。由此假设 3a 和假设 3b 得到验证。

3.4 讨论

首先, 研究 2 的结果验证了假设 2a 和 2b。即相比 HIR 功能辅助信任, HIR 情感修缮信任能更明显提升员工的组织温暖感知; 反之能更明显提升员工的组织能力感知。其次, 研究 2 的结果也验证了假设 3a 和 3b, 证明组织温暖感知和组织能力感知是解释人机关系信任影响员工工作贡献意愿的重要心理机制。再次, 由于 4 种人机协作类型与人机关系信任的内在联系, 本文设计了一个额外的情境实验检验 4 种人机协作类型对中介变量和因变量的影响, 详见网络版附录 6。

此外, 在研究 1 的质性研究中, 由于 HIR 更理性公平、效率更高的优势, 部分参与者表示更愿意

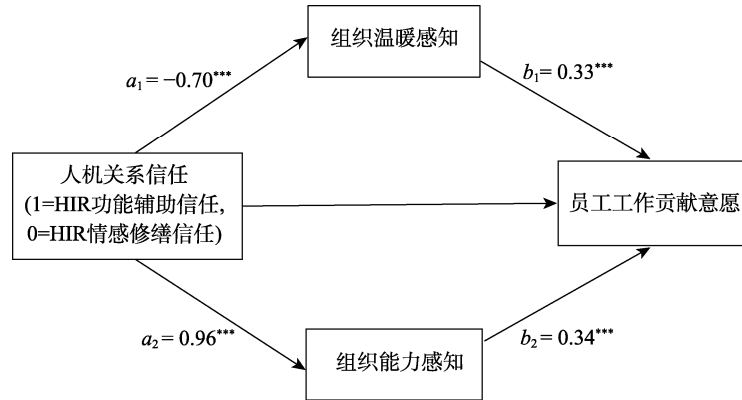


图 4 组织温暖感知和组织能力感知的中介作用(* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

与 HIR 沟通协作而不是人类。这间接揭示出员工的 HIR 代理需求导向是重要的情境变量, 并且与人机关系信任与组织温暖-能力感知之间的关系有着内在的逻辑联系。因此, 研究 3 将验证 HIR 代理需求导向的调节效应。

4 研究 3: HIR 代理需求导向的调节效应

4.1 参与者

采用 G*Power 3.1 软件(Faul et al., 2007)计算本实验所需样本量, 对于本实验适用的独立样本 t 检验分析, 显著性水平 $\alpha = 0.05$ 且中等效应量($d = 0.5$), 为了达到 90%统计检验力, 至少需要 172 名参与者。研究 3 通过线下便利抽样的方式, 以某医美整形机构的员工为对象, 招募了 200 名在职人员参与实验, 剔除掉未通过注意力检查、极端值和回答过于一致的参与者后, 最终获得 180 名有效参与者, 其中 HIR 情感修缮信任组 90 名, HIR 功能辅助信任组 90 名。包括男性 69 名(38.3%), 女性 111 名(61.7%), 平均年龄 $M = 30.83$ ($SD = 8.14$)岁。参与者的学历包括专科及以下 21 名(11.7%), 本科 125 名(69.4%), 硕士 28 名(15.6%), 博士 6 名(3.3%)。所有参与者均自愿参加实验且知情同意, 并在实验结束之后获得相应礼品。

4.2 实验程序

研究 3 为单因素两水平组间实验设计(HIR 情感修缮信任 vs. HIR 功能辅助信任), 参与者被随机分配到其中一组。首先对所有参与者进行 HIR 代理互动(任务)需求导向测量, 评估参与者的 HIR 代理需求导向高低程度。对 HIR 代理互动需求导向的测量题项改编自 Lee 和 Lyu (2016)以及 Dabholkar (1996), 包含 2 个题项(1 = “完全不同意”, 7 = “完全

同意”, $\alpha = 0.74$): “在特定的组织关系情境中, 我有与 HIR 进行沟通和交流的需要”; “在特定的组织关系情境中, 我有与 HIR 进行互动的需要”。对 HIR 代理任务需求导向的测量题项改编自 Mcfarland 等 (2006), 包含 4 个题项(1 = “完全不同意”, 7 = “完全同意”, $\alpha = 0.94$): “我想, 我需要与 HIR 代理协同工作”; “我想, 我需要 HIR 为我提供程序性事务的支持”; “我想, 我需要 HIR 为我提供算法方面的支持”; “我想, 我需要 HIR 协助我实现任务/工作的目标”。

然后要求参与者阅读人机关系信任的情境材料, 研究 3 仍然使用与研究 2 类似的情境材料, 但是有两处不同。一是, 为了增加参与者的代入感, 研究 2 的情境材料为第三人称视角(即参与者阅读另一个人经历的情境), 研究 3 则为第一人称视角, 要求参与者将自己代入情境(即参与者想象自己正在经历描述的情境)。二是, 为了增加 HIR 的真实感, 研究 3 增加了对 HIR 的形象描述。情境材料详见网络版附录 7。在阅读材料时要求被试想象: 如果这就是自己未来的工作场景, 自己对 HIR 持有的态度。此外, 为了加强操控效果, 参与者在阅读完实验材料之后还需要尽可能详细地描述, 在这样的工作场景中采用 HIR 时的感受, 时间和字数不限。随后参与者被要求填写 HIR 情感修缮信任($\alpha = 0.86$)、HIR 功能辅助信任($\alpha = 0.86$)、组织温暖感知($\alpha = 0.86$)、组织能力感知($\alpha = 0.87$)、员工工作贡献意愿($\alpha = 0.77$)等测量题项。最后收集了性别、年龄和教育水平三项人口统计信息。

4.3 数据分析与结果

操纵检验。独立样本 t 检验结果显示, HIR 情感修缮信任组的 HIR 情感修缮信任显著高于 HIR 功能辅助信任组($M_{\text{功能辅助组}} = 4.83$, $SD = 1.08$, $M_{\text{情感修缮组}} = 5.79$, $SD = 0.73$, $t(155.78) = 7.02$, $p < 0.001$, $d = 1.04$)。

HIR 功能辅助信任组的 HIR 功能辅助信任显著高于 HIR 情感修缮信任组($M_{\text{功能辅助组}} = 5.81, SD = 0.75, M_{\text{情感修缮组}} = 5.16, SD = 1.05, t(178) = -4.79, p < 0.001, d = 0.71$)。因此, HIR 情感修缮信任和 HIR 功能辅助信任情境操纵成功。

再次验证双中介效应。对人机关系信任进行虚拟变量编码作为自变量(1 = HIR 功能辅助信任, 0 = HIR 情感修缮信任), 组织温暖感知和组织能力感知作为双中介变量。在 95% 的置信区间下, 采用 PROCESS 程序自主抽样 5000 次检验组织温暖感知和组织能力感知的中介效应(Model 4; Hayes, 2013)。Bootstrap 分析结果显示, 组织温暖感知的间接中介效应是显著负向的(indirect effect = -0.38, $SE = 0.07, 95\% CI [-0.53, -0.24]$), 人机关系信任显著负向影响组织温暖感知($b = -0.82, SE = 0.13, p < 0.001$), 组织温暖感知显著正向影响员工工作贡献意愿($b = 0.46, SE = 0.04, p < 0.001$)。组织能力感知的间接中介效应是显著正向的(indirect effect = 0.27, $SE = 0.05, 95\% CI [0.17, 0.38]$), 人机关系信任正向影响组织能力感知($b = 0.87, SE = 0.17, p < 0.001$), 组织能力感知正向影响员工工作贡献意愿($b = 0.32, SE = 0.03, p < 0.001$)。即, 假设 3a 和假设 3b 得到验证。

HIR 代理互动需求导向的调节效应检验。对人机关系信任进行虚拟变量编码作为自变量(1 = HIR 情感修缮信任, 0 = HIR 功能辅助信任), 组织温暖感知和组织能力感知作为中介变量, 员工工作贡献意愿作为因变量, HIR 代理互动需求导向作为调节变量。在 95% 的置信区间下, 采用 PROCESS 程序自主抽样 5000 次(Model 8; Hayes, 2013)分析。结果显示, HIR 代理互动需求导向与人机关系信任的交互项显著正向影响组织温暖感知($b = 0.37, SE = 0.13, t = 2.96, p = 0.0035$)。进一步, 进行简单斜率

分析。如图 5a 所示, 对于低 HIR 代理互动需求导向(-1 SD), HIR 情感修缮信任(vs. HIR 功能辅助信任)对组织温暖感知的正向影响减弱($b = 0.48, p = 0.0014$); 而对于高 HIR 代理互动需求导向(+1 SD), 这种正向影响加强($b = 1.12, p < 0.001$)。HIR 代理互动需求导向与人机关系信任的交互项显著正向影响组织能力感知($b = 0.65, SE = 0.15, t = 4.23, p < 0.001$)。为了进一步探究这种交互作用, 进行简单斜率分析。如图 5b 所示, 对于低 HIR 代理互动需求导向(-1 SD), HIR 情感修缮信任(vs. HIR 功能辅助信任)对组织能力感知的负向影响显著($b = -1.44, p < 0.001$); 而对于高 HIR 代理互动需求导向(+1 SD), HIR 情感修缮信任(vs. HIR 功能辅助信任)对组织能力感知的影响不显著($b = -0.35, p = 0.0573$)。综上, 假设 4a 得到了支持。此外, 组织温暖感知(indirect effect = 0.18, $SE = 0.06, 95\% CI [0.06, 0.31]$)和组织能力感知(indirect effect = 0.22, $SE = 0.06, 95\% CI [0.12, 0.35]$)有调节的中介效应均显著, 说明人机关系信任通过组织温暖-能力感知影响员工工作贡献意愿的路径受到 HIR 代理互动需求导向的调节。

HIR 代理任务需求导向的调节效应检验。对人机关系信任进行虚拟变量编码作为自变量(1 = HIR 功能辅助信任, 0 = HIR 情感修缮信任), 组织温暖感知和组织能力感知作为中介变量, 员工工作贡献意愿作为因变量, HIR 代理任务需求导向作为调节变量。在 95% 的置信区间下, 采用 PROCESS 程序自主抽样 5000 次(Model 8; Hayes, 2013)分析。结果显示, HIR 代理任务需求导向与人机关系信任的交互项显著正向影响组织温暖感知($b = 0.47, SE = 0.07, t = 6.95, p < 0.001$)。进一步, 进行简单斜率分析。如图 6a 所示, 对于低 HIR 代理任务需求导向(-1 SD), HIR 功能辅助信任(vs. HIR 情感修缮信任)对组织温暖感知的负向影响显著($b = -1.66, p < 0.001$);

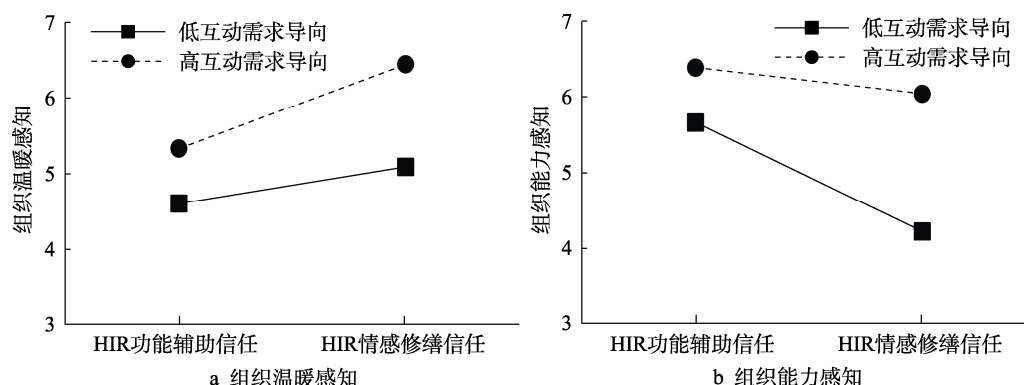


图 5 HIR 代理互动需求导向的调节作用

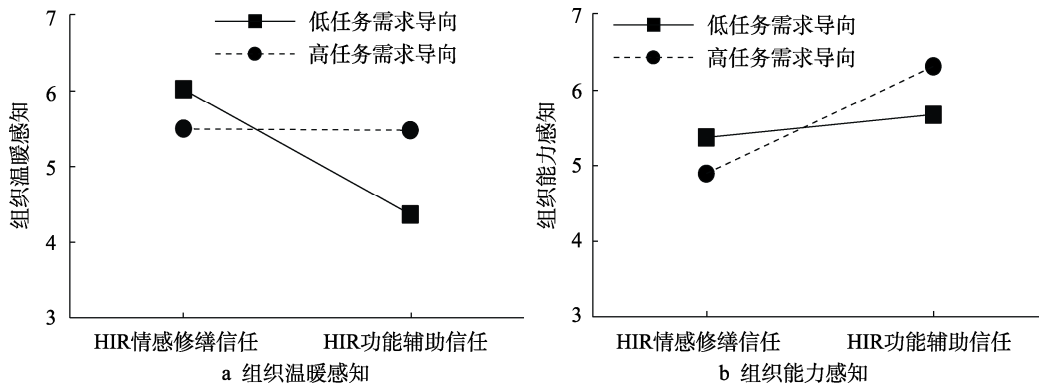


图 6 HIR 代理任务需求导向的调节作用

而对于高 HIR 代理任务需求导向(+1 SD), 这种负向影响不显著($b = -0.02, p = 0.9016$)。结果还显示, HIR 代理任务需求导向与人机关系信任的交互项显著正向影响组织能力感知($b = 0.32, SE = 0.09, t = 3.42, p = 0.0008$)。为了进一步探究这种交互作用, 进行简单斜率分析。如图 6b 所示, 对于低 HIR 代理任务需求导向(-1 SD), 人机关系信任对组织能力感知的影响不显著($b = 0.30, p = 0.1918$); 而对于高 HIR 代理任务需求导向(+1 SD), 人机关系信任对组织能力感知的正向影响显著($b = 1.42, p < 0.001$)。综上, 假设 4b 得到了支持。此外, 组织温暖感知(indirect effect = 0.21, $SE = 0.04, 95\% CI [0.14, 0.28]$)和组织能力感知(indirect effect = 0.11, $SE = 0.04, 95\% CI [0.04, 0.20]$)有调节的中介效应均显著, 说明人机关系信任通过组织温暖-能力感知影响员工工作贡献意愿的路径受到 HIR 代理任务需求导向的调节。

4.4 讨论

研究 3 结果表明, 员工的 HIR 代理需求导向在人机关系信任与组织温暖-能力感知之间起调节作用。具体而言, 随着 HIR 代理互动需求导向的增加, HIR 情感修缮信任(vs. HIR 功能辅助信任)对组织温暖感知的正向作用增强; HIR 情感修缮信任(vs. HIR 功能辅助信任)对组织能力感知的负向作用减弱。反之, 随着 HIR 代理任务需求导向的增加, HIR 功能辅助信任(vs. HIR 情感修缮信任)对组织温暖感知的负向作用减弱; HIR 功能辅助信任(vs. HIR 情感修缮信任)对组织能力感知的正向作用增强。这说明在未来组织的人机协作过程中, 不同 HIR 代理需求导向的员工适配的人机关系信任并不完全相同。

5 综合讨论

5.1 理论贡献

首先, 本文将 HIR 引入组织内部互动网络, 通过构建人机协作下的新型人机关系信任, 从理论上回应了 HIR 作为组织中的潜在成员与人类共建人机生态系统的可能(孙伟平, 2020)。一是, 本文拓展了组织信任的研究范畴。将以往组织信任主体从人与人、人与组织, 及组织与组织之间(李宁 等, 2006; Burke et al., 2007)扩展至人与 HIR 之间, 丰富了组织信任的理论框架。二是, 本文开拓性提出 HIR 情感修缮信任和 HIR 功能辅助信任双维结构, 并将其嵌入组织关系情境, 拓展了信任维度相关文献。一方面, 传统人际信任理论通常涉及信任双方目标不一致的矛盾, 如 McAllister (1995)强调基于认知的信任是依赖已知信息来判断被信任方的可靠性, 基于情感的信任是依赖与被信任方的情感联结来判断。另一方面, 传统人机信任理论通常涉及人对机器人功能和外观的信任, 很少涉及情感层面(即便有也是一种欣赏或依恋), 更没有研究探讨 HIR 对人类情感的修缮。HIR 既有别于人类, 也有别于传统意义上的机器人。但过去的研究为 HIR 情感修缮信任和 HIR 功能辅助信任的概念化提供了理论支撑, 也加深了本文对组织关系情境中人机关系信任的理解。为此本文提出更适用于组织关系情境的人机关系信任维度, 为未来人机共存环境下的信任研究提供了新视角。三是, 本研究进一步探讨了人机关系信任的塑造方式及其对员工工作贡献意愿的重要作用。一方面, 自我修缮、友情修缮、智力辅助和体力辅助 4 种人机协作方式, 为组织塑造新型人机关系信任提供了路径; 另一方面, 以往提升员工工作贡献意愿的研究主要探讨了同事支持、任务种类、组织公正和职业前景等影响因素(Bakker

et al., 2023; Mazzetti et al., 2023), 本文通过揭示人机关系信任与员工工作贡献意愿之间的关系, 为组织行为研究提供了新思路。

其次, 本文深化了组织支持理论相关研究。以往有关组织支持的研究主要集中在人类主管支持、组织奖励和有利的工作条件(如薪酬, 晋升, 工作丰富化)等组织支持形式所带来的积极效应(Eisenberger et al., 1986; Li et al., 2019; Rhoades & Eisenberger, 2002; Yu et al., 2022)。然而, 较少文献关注在 HIR 和员工协作情境下, 组织能否通过 HIR 代理提供情感修缮支持和功能辅助支持及其效果。此外, 现有人机协作的文献多聚焦于人工智能在工作场所扮演的替代角色及其负面影响(参见许丽颖 等, 2024; Yam et al., 2023)。本研究发现, 在员工信任 HIR 代理提供的情感修缮支持和功能辅助支持的前提下, 通过提高组织温暖-能力感知, 可有效促进员工工作贡献意愿。这表明, 借助 HIR 代理提供的组织支持能产生积极效应, 从而丰富并拓展了组织支持理论的现有框架。

最后, 本研究引入 HIR 代理需求导向(互动导向和任务导向)作为边界条件拓展了组织关系相关研究。目前, 关于人机关系信任在何种情况下带来差异化的组织温暖感知和组织能力感知的研究尚处于起步阶段, 将员工需求作为边界条件的研究相对有限。例如 Ilies 等(2018)关注了组织关系情境中人际需求满足的中介效应, Sheehan 等(2020)则探究了人际互动需求在机器人拟人化与采纳意愿之间的调节作用。遗憾的是, 目前尚未发现文献把上述边界条件置于组织、员工和 HIR 构成的三维结构中进行探讨。因此, 综合以上考虑引入该调节变量, 有助于我们厘清人机关系信任触发组织温暖感知和组织能力感知的边界条件, 也间接回应了 Tang 等(2023)关于“未来研究人机协作时, 应关注员工心理需求差异的影响”的建议。

5.2 实践启示

首先, 研究结果表明, 组织管理者应理性看待 HIR 的独特代理功能, 并为员工合理匹配 HIR 提供了一个设计线索。具体而言: 组织可通过工作贡献意愿测试和 HIR 代理需求导向测试, 识别员工的类型和需求(例如社交恐惧或者能力焦虑)。在此基础上, 根据员工需求提供不同类型 HIR。比如对于面临社交恐惧和偏向 HIR 代理互动需求导向的员工, 配置以陪伴和关怀为主的情感修缮型 HIR, 并减少人类代理的干预; 而对于面临能力焦虑和偏向 HIR

代理任务需求导向的员工, 配置以业务为主的功能辅助型 HIR, 并加强人类代理在情感修缮支持方面的作用。其次, 员工个人也必须意识到个人特质对自身工作贡献意愿的潜在影响。员工要树立正确的职业价值观, 应主动寻求组织支持。如若意识到自身具备 HIR 代理互动需求导向特质, 可主动寻求 HIR 代理提供情感修缮支持, 同时要注意规避过度的人机交互带来的社交孤立感(Tang et al., 2023)。如若意识到自身具备 HIR 代理任务需求导向特质, 员工应注重通过 HIR 代理的功能辅助满足个体的胜任力需要, 提升自我效能感(Tang et al., 2023), 进而增强对工作任务的贡献意愿。最后, 从整个服务生态系统的治理角度来看, 需要各方协同配合做好系统的顶层设计和配套政策, 构建和谐共生的新型人机关系(孙伟平, 2020)。未来在 HIR 融入的组织关系研究中应坚持人本主义(Cleveland et al., 2015), 在提升组织效能的同时让员工的工作贡献价值也得到实现; HIR 开发者应尊重不同 HIR 代理需求导向的人类内在价值观, 设计差异化 HIR; 社会各界应该提前做好 HIR 成为员工乃至社会成员的心理建设, 正视 HIR 的社会地位, 以此应对人机社会关系结构发生颠覆性转变时的混沌挑战。

5.3 不足与展望

本研究存在一定的局限性。首先, 本研究对人机关系信任、员工工作贡献意愿采取的是自我报告的测量方式, 这可能导致研究结果在一定程度上缺乏客观性。然而, 随着更多的 HIR 进入市场, 未来的研究可以通过观察员工与 HIR 的真实接触测量员工工作贡献意愿(Mende et al., 2019), 可以考虑采用信任博弈的方式测量人机关系信任(王晨 等, 2024), 客观测量还可以从个体无意识产生的定量行为数据(如反应时间)中获取(Gaudiello et al., 2016)。

其次, 本研究实验场景主要为 HIR 嵌入组织的场景, 在一定程度上拓展了研究的适用场景, 但是故事情境存在“低切身性、低现实性”导致的生态效率限制(王晨 等, 2024)。因此未来研究可以通过 VR 虚拟游戏等科技手段构建更多机器人真实嵌入工作以及人类生活的现实场景, 提高互动性和现实性, 进一步拓展研究的应用性。比如, 考虑社交机器人医生帮助糖尿病儿童进行自我管理的应用场景(Looije et al., 2016)。

再者, 本研究侧重探讨了两种人机关系信任导致员工工作贡献意愿变化的心理机制。但人机互动

过程和人际互动过程类似充满了复杂性和变化性,人机关系信任也处于动态变化中。例如参与者因 HIR 行为决策诱发的情绪变化、参与者与 HIR 的互动过程中的反馈、人机投射、机器人心智水平等变量都会影响人机关系信任。因此,未来的研究可以检验情绪、机器人心智等潜在变量对 HIR 情感修缮信任和 HIR 功能辅助信任的影响,即探索参与者形成 HIR 情感修缮信任和 HIR 功能辅助信任的心理机制。

最后,参与者对 HIR 代理的信任感知可能还取决于 HIR 的外型(例如性别、年龄、风格等)与机器人嵌入环境的匹配程度。比如刻板印象认为男性长相的 HIR 比女性长相的 HIR 更积极或自信,而女性长相的 HIR 更富有社交属性或情感(Mende et al., 2019)。基于此,人们认为刻板印象中的男性任务更适合男性 HIR 而不是女性 HIR,反之亦然。这导致当男性长相的 HIR 出现在以女性任务为主的服务环境中时,参与者对男性长相的 HIR 信任感知会极低,反之当男性长相的 HIR 出现在以男性任务为主的服务环境中时,其信任感知更高。未来的研究可以根据 HIR 外观、性别等刻板印象和嵌入环境的匹配检验人机关系信任的差异。

5.4 研究结论

通过 4 项研究,结合访谈和情境实验方法,本文主要得出以下结论:第一,4 个研究均表明在组织关系情境中,HIR 情感修缮信任和 HIR 功能辅助信任是新型人机关系信任的 2 个不同维度。第二,研究 1 的质性研究归纳出组织中存在 4 种人机协作类型,即自我修缮、友情修缮、智力辅助和体力辅助。进一步的情境实验证实了自我修缮(例如,自我情感缺陷的修缮)和友情修缮(例如,友情冲突的修缮)主要激发 HIR 情感修缮信任,智力辅助和体力辅助主要激发 HIR 功能辅助信任。第三,研究 2 通过情境实验验证了人机关系信任如何影响工作贡献意愿。即 HIR 情感修缮信任对员工工作贡献意愿的正向促进,源于员工对组织温暖感知的积极评价;而 HIR 功能辅助信任对员工工作贡献意愿的正向促进则源于员工对组织能力感知的积极评价。第四,研究 3 邀请医美整形机构的在职人员参加情境实验,弥补了研究 1 和研究 2 情境实验中均以金融行业的员工为样本的局限。实验结果不仅验证了 HIR 代理互动(任务)需求导向的调节效应,还重复验证了组织温暖感知和组织能力感知中介效应,从而增强了研究结论的普适性。

参 考 文 献

- Aaker, J., Vohs, K. D., & Mogilner, C. (2010). Non-profits are seen as warm and for-profits as competent: Firm stereotypes matter. *Journal of Consumer Research*, 37(2), 224–237.
- Alarcon, G. M., Capiola, A., Hamdan, I. A., Lee, M. A., & Jessup, S. A. (2023). Differential biases in human-human versus human-robot interactions. *Applied Ergonomics*, 106, Article 103858. <https://doi.org/10.1016/j.apergo.2022.103858>
- Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands-resources theory: Ten years later. *Annual Review of Organizational Psychology and Organizational Behavior*, 10, 25–53.
- Bakker, A. B., Tims, M., & Derks, D. (2012). Proactive personality and job performance: The role of job crafting and work engagement. *Human Relations*, 65(10), 1359–1378.
- Bartz, J. A., Tchalova, K., & Fenerci, C. (2016). Reminders of social connection can attenuate anthropomorphism. *Psychological Science*, 27(12), 1644–1650.
- Blut, M., & Wang, C. (2020). Technology readiness: A meta-analysis of conceptualizations of the construct and its impact on technology usage. *Journal of the Academy of Marketing Science*, 48(4), 649–669.
- Brengman, M., De Gauquier, L., Willems, K., & Vanderborght, B. (2021). From stopping to shopping: An observational study comparing a humanoid service robot with a tablet service kiosk to attract and convert shoppers. *Journal of Business Research*, 134, 263–274.
- Burke, C. S., Sims, D. E., Lazzara, E. H., & Salas, E. (2007). Trust in leadership: A multi-level review and integration. *Leadership Quarterly*, 18(6), 606–632.
- Cao, J., Zhang, J., Zhang, L., & Wang, X. (2023). The psychological structure and influence of interactive naturalness. *Acta Psychologica Sinica*, 55(1), 55–65.
- [曹剑琴, 张警吁, 张亮, 王晓宇. (2023). 交互自然性的心理结构及其影响. *心理学报*, 55(1), 55–65.]
- Chandra, S., Shirish, A., & Srivastava, S. C. (2022). To be or not to be horizontal ellipsis human? Theorizing the role of human-like competencies in conversational artificial intelligence agents. *Journal of Management Information Systems*, 39(4), 969–1005.
- Chen, M. (2019). Research on artificial intelligence development, labor classification and structural unemployment. *Economist*, (10), 66–74.
- [陈明生. (2019). 人工智能发展、劳动分类与结构性失业研究. *经济学家*, (10), 66–74.]
- Chien, S. Y., Lin, Y. L., & Chang, B. F. (2024). The effects of intimacy and proactivity on trust in human-humanoid robot interaction. *Information Systems Frontiers*, 26(1), 75–90.
- Chowdhury, S., Budhwar, P., Dey, P. K., Joel-Edgar, S., & Abadie, A. (2022). AI-employee collaboration and business performance: Integrating knowledge-based view, socio-technical systems and organisational socialisation framework. *Journal of Business Research*, 144, 31–49.
- Chowdhury, S., Dey, P., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review*, 33(1), Article 100899. <https://doi.org/10.1016/j.hrmr.2022.100899>
- Cleveland, J. N., Byrne, Z. S., & Cavanagh, T. M. (2015). The future of HR is RH: Respect for humanity at work. *Human Resource Management Review*, 25(2), 146–161.
- Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022).

- Blame the bot: Anthropomorphism and anger in customer-chatbot interactions. *Journal of Marketing*, 86(1), 132–148.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92(4), 631–648.
- Cuddy, A. J. C., Glick, P., & Beninger, A. (2011). The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior*, 31, 73–98.
- Dabholkar, P. A. (1996). Consumer evaluations of new technology-based self-service options: An investigation of alternative models of service quality. *International Journal of Research in Marketing*, 13(1), 29–51.
- de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, 22(3), 331–349.
- Deng, X. (2014). Why ethical consumers don't do what they say: The study on factors influencing consumers' ethical buying intention-behavior gap. *Acta Psychologica Sinica*, 46(7), 1014–1031.
- [邓新明. (2014). 消费者为何喜欢“说一套, 做一套”——消费者伦理购买“意向-行为”差距的影响因素. *心理学报*, 46(7), 1014–1031.]
- Desideri, L., Ottaviani, C., Malavasi, M., di Marzio, R., & Bonifacci, P. (2018). Emotional processes in human-robot interaction during brief cognitive testing. *Computers in Human Behavior*, 90, 331–342.
- Ding, X. (2020). On the legal regulation of algorithms. *Social Sciences in China*, (12), 138–159+203.
- [丁晓东. (2020). 论算法的法律规制. *中国社会科学*, (12), 138–159+203.]
- Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology*, 71(3), 500–507.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27(2), 67–73.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902.
- Gati, I., & Kulcsar, V. (2021). Making better career decisions: From challenges to opportunities. *Journal of Vocational Behavior*, 126, Article 103545. <https://doi.org/10.1016/j.jvb.2021.103545>
- Gaudiello, I., Zibetti, E., Lefort, S., Chetouani, M., & Ivaldi, S. (2016). Trust as indicator of robot functional and social acceptance. An experimental study on user conformation to iCub answers. *Computers in Human Behavior*, 61, 633–655.
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.
- Goldenberg, A., Weisz, E., Sweeny, T. D., Cikara, M., & Gross, J. J. (2021). The crowd-emotion-amplification effect. *Psychological Science*, 32(3), 437–450.
- Gong, Z., Liu, C., Shen, W., Wang, X., & Shi, R. (2017). The impact of trust on creativity: Activation, inhibition and inverted U model hypothesis. *Advances in Psychological Science*, 25(3), 463–474.
- [贡喆, 刘昌, 沈汪兵, 王贤, 石荣. (2017). 信任对创造力的影响: 激发, 抑制以及倒 U 假设. *心理科学进展*, 25(3), 463–474.]
- Grant, A. M. (2008). The significance of task significance: Job performance effects, relational mechanisms, and boundary conditions. *Journal of Applied Psychology*, 93(1), 108–124.
- Harris-Watson, A. M., Larson, L. E., Lauharatanahirun, N., DeChurch, L. A., & Contractor, N. S. (2023). Social perception in Human-AI teams: Warmth and competence predict receptivity to AI teammates. *Computers in Human Behavior*, 145, Article 107765. <https://doi.org/10.1016/j.chb.2023.107765>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- He, G., Chen, C., He, Z., Cui, L., Lu, J., Xuan, H., & Lin, L. (2022). Human-agent collaborative decision-making in intelligent organizations: A perspective of human-agent inner compatibility. *Advances in Psychological Science*, 30(12), 2619–2627.
- [何贵兵, 陈诚, 何泽桐, 崔力丹, 陆嘉琦, 宣泓舟, 林琳. (2022). 智能组织中的人机协同决策: 基于人机内部兼容性的研究探索. *心理科学进展*, 30(12), 2619–2627.]
- Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology*, 52(2), 196–205.
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(3), 407–434.
- Hou, J., Zhu, Y., & Fang, X. (2021). Mobile phone addiction and depression: Multiple mediating effects of social anxiety and attentional bias to negative emotional information. *Acta Psychologica Sinica*, 53(4), 362–373.
- [侯娟, 朱英格, 方晓义. (2021). 手机成瘾与抑郁: 社交焦虑和负性情绪信息注意偏向的多重中介作用. *心理学报*, 53(4), 362–373.]
- Hu, Q., Wang, H., & Long, L. (2020). Will newcomer job crafting bring positive outcomes? The role of leader-member exchange and traditionality. *Acta Psychologica Sinica*, 52(5), 659–668.
- [胡巧婷, 王海江, 龙立荣. (2020). 新员工工作重塑会带来积极的结果吗? 领导成员交换与个体传统性的作用. *心理学报*, 52(5), 659–668.]
- Ilies, R., Lanaj, K., Pluut, H., & Goh, Z. (2018). Intrapersonal and interpersonal need fulfillment at work: Differential antecedents and incremental validity in explaining job satisfaction and citizenship behavior. *Journal of Vocational Behavior*, 108, 151–164.
- Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366–410.
- Köhler, C. F., Rohm, A. J., de Ruyter, K., & Wetzels, M. (2011). Return on interactivity: The impact of online agents on newcomer adjustment. *Journal of Marketing*, 75(2), 93–108.
- Komiak, S. Y., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly*, 30(4), 941–960.
- Kong, H., Yin, Z., Baruch, Y., & Yuan, Y. (2023). The impact of trust in AI on career sustainability: The role of

- employee-AI collaboration and protean career orientation. *Journal of Vocational Behavior*, 146, Article 103928. <https://doi.org/10.1016/j.jvb.2023.103928>
- Kurtessis, J. N., Eisenberger, R., Ford, M. T., Buffardi, L. C., Stewart, K. A., & Adis, C. S. (2017). Perceived organizational support: A meta-analytic evaluation of organizational support theory. *Journal of Management*, 43(6), 1854–1884.
- Lee, H. J., & Lyu, J. (2016). Personal values as determinants of intentions to use self-service technology in retailing. *Computers in Human Behavior*, 60, 322–332.
- Levin, D. Z., & Cross, R. (2004). The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management Science*, 50(11), 1477–1490.
- Lewis, M., Sycara, K., & Walker, P. (2018). The role of trust in human-robot interaction. In H. A. Abbass, J. Scholz, & D. J. Reid (Eds.), *Foundations of trusted autonomy: Studies in systems decision and control* (Vol. 117, pp. 135–159). Springer-Cham.
- Lewis, S. C., Guzman, A. L., & Schmidt, T. R. (2019). Automation, journalism, and human-machine communication: Rethinking roles and relationships of humans and machines in news. *Digital Journalism*, 7(4), 409–427.
- Li, J. J., Bonn, M. A., & Ye, B. H. (2019). Hotel employee's artificial intelligence and robotics awareness and its impact on turnover intention: The moderating roles of perceived organizational support and competitive psychological climate. *Tourism Management*, 73, 172–181.
- Li, N., Yan, J., & Jin, M. (2006). How does trust in organizations benefit task performance. *Acta Psychologica Sinica*, 38(5), 770–777.
- [李宁, 严进, 金鸣轩. (2006). 组织内信任对任务绩效的影响效应. *心理学报*, 38(5), 770–777.]
- Looije, R., Neerinx, M. A., Peters, J. K., & Henkemans, O. A. B. (2016). Integrating robot support functions into varied activities at returning hospital visits: Supporting child's self-management of diabetes. *International Journal of Social Robotics*, 8(4), 483–497.
- Lu, L., Cai, R., & Gursoy, D. (2019). Developing and validating a service robot integration willingness scale. *International Journal of Hospitality Management*, 80, 36–51.
- Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin*, 18(1), 3–9.
- Mazzetti, G., Robledo, E., Vignoli, M., Topa, G., Guglielmi, D., & Schaufeli, W. B. (2023). Work engagement: A meta-analysis using the job demands-resources model. *Psychological Reports*, 126(3), 1069–1107.
- McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 38(1), 24–59.
- Mcfarland, R. G., Challagalla, G. N., & Shervani, T. A. (2006). Influence tactics for effective adaptive selling. *Journal of Marketing*, 70(4), 103–117.
- Memon, M. A., Salleh, R., Nordin, S. M., Cheah, J. H., Ting, H., & Chuah, F. (2018). Person-organisation fit and turnover intention: The mediating role of work engagement. *Journal of Management Development*, 37(3), 285–298.
- Mende, M., Scott, M. L., van Doorn, J., Grewal, D., & Shanks, I. (2019). Service robots rising: How humanoid robots influence service experiences and elicit compensatory consumer responses. *Journal of Marketing Research*, 56(4), 535–556.
- Molina, M. A. C., González, J. M. H., Florencio, B. P., & González, J. L. G. (2014). Does the balanced scorecard adoption enhance the levels of organizational climate, employees' commitment, job satisfaction and job dedication? *Management Decision*, 52(5), 983–1010.
- Mori, M. (1970). The uncanny valley. *Energy*, 7, 33–35.
- Novikova, J., & Watts, L. (2015). Towards artificial emotions to assist social coordination in HRI. *International Journal of Social Robotics*, 7(1), 77–88.
- Obschonka, M., Pavez, I., Kautonen, T., Kibler, E., Salmela-Aro, K., & Wincent, J. (2023). Job burnout and work engagement in entrepreneurs: How the psychological utility of entrepreneurship drives healthy engagement. *Journal of Business Venturing*, 38(2), Article 106272. <https://doi.org/10.1016/j.jbusvent.2022.106272>
- Omrani, N., Riveccio, G., Fiore, U., Schiavone, F., & Agreda, S. G. (2022). To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts. *Technological Forecasting and Social Change*, 181, Article 121763. <https://doi.org/10.1016/j.techfore.2022.121763>
- Picard, R. W. (2003). Affective computing: Challenges. *International Journal of Human-Computer Studies*, 59(1–2), 55–64.
- Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, 29(4), 1841–1848.
- Rhoades, L., & Eisenberger, R. (2002). Perceived organizational support: A review of the literature. *Journal of Applied Psychology*, 87(4), 698–714.
- Röning, J., Holappa, J., Kellokumpu, V., Tikanmäki, A., & Pietikäinen, M. (2014). Minotaurus: A system for affective human-robot interaction in smart environments. *Cognitive Computation*, 6(4), 940–953.
- Scott, M. L., Mende, M., & Bolton, L. E. (2013). Judging the book by its cover? How consumers decode conspicuous consumption cues in buyer-seller relationships. *Journal of Marketing Research*, 50(3), 334–347.
- Shamim, S., Yang, Y., Zia, N. U., Khan, Z., & Shariq, S. M. (2023). Mechanisms of cognitive trust development in artificial intelligence among front line employees: An empirical examination from a developing economy. *Journal of Business Research*, 167, Article 114168. <https://doi.org/10.1016/j.jbusres.2023.114168>
- Sheehan, B., Jin, H. S., & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, 115, 14–24.
- Siau, K., & Wang, W. Y. (2018). Building trust in Artificial Intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47–53.
- Song, C. S., & Kim, Y. K. (2022). The role of the human-robot interaction in consumers' acceptance of humanoid retail service robots. *Journal of Business Research*, 146, 489–503.
- Sun, W. (2020). Artificial intelligence and the "New Alienation" of human beings. *Social Sciences in China*, (12), 119–137+202–203.
- [孙伟平. (2020). 人工智能与人的“新异化”. *中国社会科学*, (12), 119–137+202–203.]
- Tang, P. M., Koopman, J., Mai, K. M., De Cremer, D., Zhang, J. H., Reyniers, P., Ng, C. T. S., & Chen, I. (2023). No person is an island: Unpacking the work and after-work consequences of interacting with artificial intelligence. *Journal of Applied Psychology*, 108(11), 1766–1789.
- Tang, X., Sun, B., Zhang, E., & Mei, F. (2021). Research on social value substitution, risk and consumption attitude

- based on humanoid intelligent robot. *Nankai Business Review*, 24(6), 4–13.
- [唐小飞, 孙炳, 张恩忠, 梅发贵. (2021). HIR 社会价值替代与风险态度研究. *南开管理评论*, 24(6), 4–13.]
- Tisu, L., Lupşa, D., Virgă, D., & Rusu, A. (2020). Personality characteristics, job performance and mental health: The mediating role of work engagement. *Personality and Individual Differences*, 153, Article 109644. <https://doi.org/10.1016/j.paid.2019.109644>
- van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology*, 81(5), 525–531.
- Wang, C., Chen, W., Huang, L., Hou, S., & Wang, Y. (2024). Robots abide by ethical principles promote human–robot trust? The reverse effect of decision types and the human–robot projection hypothesis. *Acta Psychologica Sinica*, 56(2), 194–209.
- [王晨, 陈为聪, 黄亮, 侯苏豫, 王益文. (2024). 机器人遵从伦理促进人机信任? 决策类型反转效应与人机投射假说. *心理学报*, 56(2), 194–209.]
- Wang, W., Chen, L., Xiong, M., & Wang, Y. (2023). Accelerating AI adoption with responsible AI signals and employee engagement mechanisms in health care. *Information Systems Frontiers*, 25(6), 2239–2256.
- Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113–117.
- Wei, H., & Long, L. (2009). Effects of cognition- and affect-base trust in supervisors on task performance and OCB. *Acta Psychologica Sinica*, 41(1), 86–94.
- [韦慧民, 龙立荣. (2009). 主管认知信任和情感信任对员工行为及绩效的影响. *心理学报*, 41(1), 86–94.]
- Wei, Q., Li, M., & Chen, X. (2018). Social class and social perception: Is warmth or competence more important? *Acta Psychologica Sinica*, 50(2), 243–252.
- [韦庆旺, 李木子, 陈晓晨. (2018). 社会阶层与社会知觉: 热情和能力哪个更重要? *心理学报*, 50(2), 243–252.]
- Willems, K., Verhulst, N., De Gauquier, L., & Brengman, M. (2023). Frontline employee expectations on working with physical robots in retailing. *Journal of Service Management*, 34(3), 467–492.
- Wynne, K. T., & Lyons, J. B. (2018). An integrative model of autonomous agent teammate-likeness. *Theoretical Issues in Ergonomics Science*, 19(3), 353–374.
- Xie, B., & Long, L. (2008). The effects of career plateau on job satisfaction, organizational commitment and turnover intentions. *Acta Psychologica Sinica*, 40(8), 927–938.
- [谢宝国, 龙立荣. (2008). 职业生涯高原对员工工作满意度、组织承诺、离职意愿的影响. *心理学报*, 40(8), 927–938.]
- Xie, M., Li, F., Luo, Y., Ke, L., Wang, X., & Wang, Y. (2022). A developmental model of job burnout dimensions among primary school teachers: Evidence from structural equation model and cross-lagged panel network model. *Acta Psychologica Sinica*, 54(4), 371–384.
- [谢敏, 李峰, 罗玉晗, 柯李, 王侠, 王耘. (2022). 小学教师职业倦怠维度发展顺序探究——来自结构方程模型和交叉滞后网络分析模型的证据. *心理学报*, 54(4), 371–384.]
- Xu, L., Wang, X., Yu, F., & Peng, K. (2024). The influence of perceived robot threat on workplace objectification. *Acta Psychologica Sinica*, 56(2), 210–225.
- [许丽颖, 王学辉, 喻丰, 彭凯平. (2024). 感知机器人威胁对职场物化的影响. *心理学报*, 56(2), 210–225.]
- Xu, L., Yu, F., & Peng, K. (2022). Algorithmic discrimination causes less desire for moral punishment than human discrimination. *Acta Psychologica Sinica*, 54(9), 1076–1092.
- [许丽颖, 喻丰, 彭凯平. (2022). 算法歧视比人类歧视引起更少道德惩罚欲. *心理学报*, 54(9), 1076–1092.]
- Xu, Y., & Liu, Y. (2021). The impact of trust in technology and trust in leadership on the adoption of new technology from employee's perspective. *Advances in Psychological Science*, 29(10), 1711–1723.
- [徐祯, 刘艺璇. (2021). 技术信任和领导信任对企业员工新技术接受的影响. *心理科学进展*, 29(10), 1711–1723.]
- Yam, K. C., Tang, P. M., Jackson, J. C., Su, R., & Gray, K. (2023). The rise of robots increases job insecurity and maladaptive workplace behaviors: Multimethod evidence. *Journal of Applied Psychology*, 108(5), 850–870.
- Yoganathan, V., Osburg, V. S., Kunz, W. H., & Toporowski, W. (2021). Check-in at the Robo-desk: Effects of automated social presence on social cognition and service implications. *Tourism Management*, 85, Article 104309. <https://doi.org/10.1016/j.tourman.2021.104309>
- Yu, A., Pichler, S., Russo, M., & Hammer, L. (2022). Family-supportive supervisor behaviors (FSSB) and work-family conflict: The role of stereotype content, supervisor gender, and gender role beliefs. *Journal of Occupational and Organizational Psychology*, 95(2), 275–304.
- Zhu, X., & Jiang, G. (2011). The working alliance in clients' eyes: A qualitative analysis. *Acta Psychologica Sinica*, 43(4), 420–431.
- [朱旭, 江光荣. (2011). 当事人眼里的工作同盟: 质的分析. *心理学报*, 43(4), 420–431.]

Impact of trusting humanoid intelligent robots on employees' job dedication intentions: An investigation based on the classification of human-robot trust

TANG Xiaofei¹, WANG Changmei¹, SUN Xiaodong¹, CHANG En-Chung²

(¹ School of Business Administration, Southwestern University of Finance and Economics, Chengdu 611130, China)

(² Business School, Renmin University of China, Beijing 100872, China)

Abstract

As humanoid intelligent robots (HIRs) become increasingly integrated into organizations to provide emotional and functional support, understanding and fostering human-robot trust has become a critical area of focus. This study explores the formation and impact of human-robot trust from the perspective of the Unique Agent Hypothesis. We propose that human-robot trust comprises two distinct dimensions: emotional repair trust and functional aiding trust. Among various forms of human-robot collaboration, self-repair and friendship-repair forms primarily trigger emotional repair trust, while intelligence-aiding and physical-aiding forms more effectively enhance functional aiding trust. Furthermore, employees who develop emotional repair trust (vs. functional aiding trust) in HIRs perceive there to be greater organizational warmth, which in turn strengthens their job dedication intentions. Conversely, employees who develop functional aiding trust (vs. emotional repair trust) in HIRs perceive there to be higher organizational competence, thereby enhancing their job dedication intentions. In addition, interaction orientation and task orientation are introduced as crucial situational moderators.

This research included one qualitative study and three quantitative studies. Study 1 (qualitative and experimental study) identified the two dimensions of human-robot trust and investigated the relationship between four forms of human-robot collaboration and trust. Study 2 (experimental study) verified how human-robot trust affects employees' job dedication intentions through perceived organizational competence and warmth. Study 3 (experimental study) tested how interaction and task orientations moderate these effects.

The main findings of this study are as follows. Study 1 provided evidence that emotional repair trust and functional aiding trust are two distinct dimensions of human-robot trust. Results showed that both self-repair and friendship-repair forms of human-robot collaboration elicited higher emotional repair trust than functional aiding trust, whereas both intelligence-aiding and physical-aiding forms led to higher functional aiding trust than emotional repair trust. Study 2 further demonstrated that functional aiding trust led to higher perceived organizational competence but lower perceived organizational warmth than did emotional repair trust. Further analysis (coded as 0 = emotional repair trust and 1 = functional aiding trust) in Study 2 revealed that the indirect effect of perceived organizational warmth was significantly negative, while the indirect effect of perceived organizational competence was significantly positive. Study 3 showed that interaction orientation (task orientation) positively (negatively) moderated the relationship between human-robot trust (coded as 0 = functional aiding trust and 1 = emotional repair trust) and perceived organizational warmth while negatively (positively) moderating the relationship between human-robot trust and perceived organizational competence.

This study contributes to the emerging field of machine behavior by embedding HIR into the network of internal organizational dynamics and examining the impact of human-robot trust on employees' job dedication intentions. The findings of this study suggest that employees' trust in HIR agents' emotional and functional support enhances perceived organizational warmth and competence, promoting their job dedication intentions. This demonstrates that organizational support provided by HIR agents can generate positive effects, thereby enriching and expanding the existing framework of organizational support theory.

Keywords humanoid intelligent robot, human-robot trust, emotional repair trust, functional aiding trust, employees' job dedication intentions

附录 1: 研究 1 的质性研究

1.1 访谈对象

为了初步探索组织关系情境中人机关系信任的维度, 本研究拟通过访谈方式邀请访谈对象参与研究。现有研究指出, 一项半结构深度访谈的样本可以小, 但必须多样化; 8 个访谈对象每人 30~60 分钟的访谈内容可以成为充分的样本依据(邓新明, 2014; 曹剑琴 等, 2023)。遵循上述建议, 本研究受邀访谈的对象共 9 名(4 名男性, 5 名女性), 其中 3 名来自金融行业, 2 名来自美容整形行业, 1 名来自通信制造行业, 1 名来自房地产行业, 1 名来自工程投资建设行业, 1 名来自教育行业。他们中既有中高层管理人员, 也有基层工作人员和新入职员工, 平均从业 8.56 ($SD = 6.13$) 年, 能够充分表达有关“HIR 嵌入组织关系情境中人机关系信任”的看法。

1.2 访谈内容

第一部分访谈内容询问访谈对象在过去的工作经历中面临的问题, 比如在工作中遇到的困难和挫折, 哪些方面有不足, 要求尽可能详细地说明和举例。以及在工作场景中是否有过负面情绪、情感, 是什么导致的? 如果有则让其对自己的经历进行描述。第二部分询问访谈对象对人机关系的了解程度和偏好, 例如“最早是什么时候开始听说人工智能和 HIR (类人智能机器人) 的? 请说说对人工智能和 HIR 的看法, 以及对人机关系的理解? 您相信在未来的工作中 HIR 可以与人类协作完成工作吗? 您认为 HIR 可以在哪些方面为您提供帮助? 您相信 HIR 可以帮助员工排解消极情绪或弥补情感缺陷吗? 在哪些情境下, 您愿意让 HIR 对您进行心理疏导? 要求尽可能详细地说明和举例。如果您在工作中遇到挫折(比如同事间出现矛盾或者工作表现不佳), 您是更愿意与人类管理者沟通还是更愿意与 HIR 进行沟通?”。第三部分询问访谈对象的一些基本信息, 如年龄、性别、从业时长、从事行业和职位等。

1.3 数据分析与结果

本研究采用协商一致的质的研究方法(Consensual Qualitative Research, CQR)对数据进行分析(朱旭, 江光荣, 2011; Hill et al., 2005)。CQR 是一种定性分析研究方法。核心思想是通过邀请多名分析员组成 1 个研究小组引入多元视角, 对半结构深度访谈中通过开放式问题收集的文本数据进行分析(Hill et al., 2005)。本研究邀请了具备 CQR 分析经验的 2 名博士生和 1 名硕士生组成研究小组, 对收集的 9 名访谈对象数据进行分析。其中数据分析包括四个主要步骤。第一, 划分域(domain), 将与研究主题相关的访谈内容划分为几个话题范围。第二, 精炼核心观点(core idea), 即用更少和更精准的词汇表达不同被试的访谈内容, 忠于原文原意。第三, 进行交叉分析(cross analysis), 即把反映了共同主题的核心观点聚成类别。为了增加数据的稳定性, 借鉴朱旭和江光荣(2011)的研究, 先随机抽取 8 个个案进行交叉分析。由研究小组协商讨论形成一致结果后, 将剩余个案加入, 检查类别是否有所变化。结果发现新增个案没有改变原有结果。第四, 外部审核, 再邀请 1 名组外的博士生作为审计员, 复核研究小组的工作, 反馈建议, 减少群体思维的影响。根据反馈建议, 科研小组经过反复讨论修正后, 最终形成共识。

CQR 结果发现, 工作/组织场景中人机关系信任的基础是访谈对象期望 HIR 为其解决工作中遇到的问题。即组织关系情境中的人机关系信任本质是在 HIR 辅助/修缮/协作人类的过程中产生的关系信任。与人机关系信任相关的内容可分为 HIR 情感修缮信任和 HIR 功能辅助信任 2 个维度(域), 具体为自我修缮、友情修缮、智力辅助和体力辅助 4 种人机协作(类)。其中, HIR 情感修缮信任维度(域)包含自我修缮和友情修缮 2 种人机协作(类)。在访谈中, 有访谈对象表露了长期工作导致的麻木倦怠等由个体心理因素导致的情感缺陷。这类访谈对象大多期望通过 HIR 的帮助来实现自我提升, 并相信 HIR 能够实现对自我消极情绪的修缮。这不仅与自我修缮的定义较为一致, 还间接揭示了自我修缮可以激发 HIR 情感修缮信任。还有访谈对象表达了对上级的不满和抵制, 对同事的愤怒无语和冷漠迟钝等外在因素带来的消极情绪。这类消极情绪、情感往往伴随着组织中的人际关系破裂, 对“职场友情”带来严峻考验。访谈对象也从不同程度表明, 职场中的 HIR 或许比人类同事更适合作为朋友, 愿意依赖 HIR 而不是人类。这些观点不仅与友情修缮的定义较为契合, 还间接说明友情修缮能够促进 HIR 情感修缮信任。HIR 功能辅助信任维度(域)涉及智力辅助和体力辅助 2 种人机协作(类)。在访谈中, 访谈对象相信 HIR 可以从智力(比如, 计算数理模型)和体力(比如, 收发快递)两个方面辅助人类, 减少其框架性和程序性工作。这些观点与智力辅助和体力辅助的定义高度吻合的同时, 也间接论证了智力辅助和体力辅助能够带来功能辅助信任。所有域、类及其下子类的结果呈现在附表 1 中, 表中还包括每个类别的代表性个案及核心观点举例。

附表 1 人机关系信任维度: CQR 结果

域(维度)、类及子类	个案(数)	代表性个案的核心观点举例
1. HIR 情感修缮信任域(维度)		
(1)自我修缮		
重塑工作成就感	2	感觉没有得到预期的工作成就,我会很不舒服、难过,机器人的理性属性可以改善人类感性等情感属性的缺陷,我相信他能够将从负面情绪旋涡中抽离出来;进行到一半时发现花了很多时间去探索的项目做不出来,我常常会觉得工作没什么意思,还会对自己产生怀疑,机器人拥有着对情绪进行调控和治愈的能力,所以它是一个值得信任的心理治愈师。
缓解工作麻木倦怠感	2	每天兼顾运营相关的各种各样的杂事让我觉得自己像是大内总管、维修工,丧失工作热情、感觉麻木,通过机器人的心理辅助治疗,我相信能有效提升我的工作热情和活力;人工智能的出现是对人类现有时间的延长,或者说是某些认知的拓展,能把我从重复琐碎的程序性工作中解放,所以人工智能激发了我的个人价值和工作热情。
辅助反省缓解负罪感	2	工作状态不太好时,可能会在工作中犯很多低级错误,为此,我感到非常自责懊悔。如果机器人可以辅助我进行心理疏导和反省纠错,我想它应该可以帮我消解我那部分过度的自责感;在工作中犯错以后,我常常感到极度懊悔,我愿意与机器人沟通复盘,减轻负罪心理,我相信这比我独自承担会好得多。
拓宽认知缓解焦虑	2	当我需要在短时间内去学习大量的新知识,我会感到焦虑、压力和茫然无措,类人机器人对焦虑和压力的靶向修复能力,相信会让我释然;很多时候因为认知的局限,我时常会陷入到“本领恐慌”的焦虑中,期待组织中引入机器人对员工心理和情绪进行辅助治疗,我相信这能有效打消庸人自扰的消极情绪。
(2)友情修缮		
缓解部门或领导冲突导致的人际负面情绪	3	当领导的批评让我感到悲伤时,机器人能够辅助我利用大数据分析原因和消除负面情绪,所以在组织中利用机器人修复情绪是值得推荐的;平级两个部门之间或者两个单位之间因责任不明晰,流程不明确产生推诿,让我对推诿部门的人员产生对立情绪,采用机器人作为沟通的桥梁,能够有效缓解这种冲突带来的负面情绪;有时,由于上级的原因导致不同部门之间产生了利益分配不均的情况,对上级不满的负面情绪就会油然而生。由于我与机器人不会有利益冲突,这让我不用担心被背叛,所以机器人是一个更好的情感交互对象。
缓解同事之间工作冲突导致的人际负面情绪	3	是的,我喜欢工作节奏慢一点,这往往会和工作节奏拉得快的同事产生分歧,我感觉我们相互之间的关系会变得冷漠、迟钝、甚至很少愿意去沟通。如果单位有一个能共情我、能理解我的机器人,我愿意寻求机器人的帮助,向机器人倾诉这类情绪,这样的我们也许就不会那么冷漠和无情了吧;有的同事经常把锅甩给我,功劳归在自己头上。由于不用像担心人类朋友那样担心机器人那么烦我,所以我更放心把机器人当作情绪倾诉对象;职场里可能遭遇同事背叛使我陷入焦虑和难受情绪,所以相比人类同事,我更相信机器人。
缓解同事之间社交冲突导致的人际负面情绪	2	同事间复杂的人际关系本就让我不是很愿意与同事分享喜怒哀乐,而且遇到大嘴巴同事很可能在不经你同意的情况下,扭头就把你讲的话告诉别人了,我很讨厌这种没有分寸和边界感的同事,如果机器人可以深度融入到职场,它更合适成为我的情感交互对象,并能有效缓解我的负面情绪;有的同事心眼小、八卦,经常在背后说三道四和造谣生事,我对这样的同事很不爽、也无语。相比人类,我觉得机器人能更好地倾听和安慰我,并赋予我工作中的积极能量。
2. HIR 功能辅助信任域(维度)		
(3)智力辅助		
收集并解读资讯	3	在我感到收集信息的渠道匮乏时,借助机器人强大的搜索引擎,相信可以协助我完成这类工作;未来,在工作中可以利用机器人的专业能力去收集、筛选和过滤信息,从而提高我的工作效率;对于一些政策性的文件解读,诸如 ChatGPT 等工具可能能力更强,所以我觉得完全可以依赖机器人。
处理表格	2	像用表格汇总数据这类工作,应该可以完全交给机器人来完成,它的能力完全可以胜任;诸如求和、画图、做表之类常规性操作,机器人的能力在不断地超越人类,相信不久的将来它就可以基本替代人类。
文案写作	3	在条件允许的情况下,机器人充当助手辅助完成文案性质的工作,我相信不会输给人类;目前,机器人已经可以用来协助我们写工作总结和做活动策划等,它是值得我信任的好助手;将来,机器人可能陪同我一起完成投资项目的研判,甚至它还能帮我进行投资陈述,所以我认为它可以成为像人类一样的工作伙伴。
计算数据	2	如果机器人能够根据诉求去自动完成桥荷载模型的测算,传输出最终结果,那么就可以帮助非专业人员、或者说不是这个领域的人把跨领域的工作干得更好;在办业务开存单的过程中,机器人拥有超强的运算能力,所以采用机器人做统计工作,我觉得是可以信任的。

续表

域(维度)、类及子类	个案(数)	代表性个案的核心观点举例
2. HIR 功能辅助信任域(维度)		
(3)智力辅助		
生成式图片	2	我把一家三口在餐厅吃饭的场景用文字描述出来，机器人能够理解这段文字，并生成出我想要的图片；我认为，将来的工作报告中是可以借助机器人根据我们提供的素材主动生成并添加图片、影像之类的。
培养新人	2	机器人可以提供工作上经验性的帮助，比如说我是新入职的，什么都不会，但是机器人可以教我，或许我再也不需要人类师傅了；作为一个新人，当组织让我去处理很多本职工作以外的事情，我愿意借助机器人的能力来辅助我完成。
秘书助理	2	在会谈中机器人可以用摄像头把对话的内容都记录下来，当我遗忘了什么，就可以问机器人，所以我觉得机器人是充当秘书的好角色；机器人不但有记录、运算功能，而且还有反馈、纠错功能，这些能力在工作中对我有很大的帮助，我认为机器人充当秘书也许是更好的选择。
(4)体力辅助		
跑腿性质	3	在职场中可能会遇到一些分身乏术的情况，可以用 HIR 帮我们送水、泡茶、冲咖啡、打印复印、收发快递等，我相信它会做得很好；将来或许有像酒店内送餐食和日用品的类人机器人一样，可以很好地完成对合同协议书、法律诉讼函等这些文件的收发整理，并且全程监控；当然，我愿意让类人机器人帮我做一些日常的参观接待和讲解。
高强度高重复	3	搬运负重的时候，人最多只能抬百来斤，机器人载重可能上千斤，所以我觉得在某些方面机器是值得信任的；工程建设中的拌合站，全智能的钢筋加工厂都可以让机器人来做；机械化的这种生产线，以前需要员工把零件放到指定的位置，现在使用机器人也做得一样好。
有安全隐患的	1	将来一些工作强度大、危险性高，或者一些特殊的工作环境可以依赖机器人，比如桩基础的浇灌，人工操作会存在一定的危险性。
拍照识别	3	之前的彩电工人是用肉眼去识别一个芯片模块，现在机器人可能做得比人更好；或许不久的将来，机器人代替我完成一些体力上的工作，比如用机器人参与拍照，我相信他会漂亮地完成任务；用机器人精准识别客户签名，或做其他精准操作层面的工作，我觉得是可以信赖的。

说明：表中的“机器人”是受访者对“类人智能机器人”省略的表达，在受访前已经告诉访谈对象是类人智能机器人。

1.4 讨论

通过访谈发现员工在工作中面临的问题可以分为自我修缮、友情修缮、智力辅助和体力辅助这 4 类问题。建立在未来 HIR 兼具智商、情商和物理能力的前提下，访谈对象相信这 4 类问题都可以通过 HIR 的帮助得到不同程度的改善。在访谈中，部分访谈对象也表达了他们对 HIR 中立或消极的观点。比如，访谈对象 1 的看法就比较中立，他说“对于在组织中采用 HIR，我既不支持、也不反对，只要能确保安全、不违反伦理，我觉得组织可以根据员工自身的需求相应地提供 HIR 协助”。比如，访谈对象 4 和 5 就表达了较为消极的观点。访谈对象 4 的描述“我感觉让 HIR 提供情绪价值好像有点难，因为我始终会觉得它是一个机器人、程序化的电脑、没有感情，除非它真的拥有了人类的情感，变成了一个真正的人，才能够真正的去理解我，并与我共情”。访谈对象 5 的描述“HIR 应该做不了情绪的安抚，我也不寄希望于它”。

事实上，除了人机关系中的功能辅助倍受关注外，人机关系中的情感因素也是一个备受关注的领域。比如 MIT 媒体实验室 Picard (2003)所提出的情感计算(Affective Computing)概念就是为了促进智能系统“情感交互”的发展，Röning 等(2014)关于服务机器人情感交互的研究，Novikova 和 Watts (2015)对人工情感(Artificial Emotions)的探索研究，以及让算法学会如何学习的元学习算法(Meta Learning)的发展都说明 HIR 具备情感交互能力并不是遥不可及的。由此可见，HIR 能否具备情感交互能力已经从技术问题演变成心理问题，即愿不愿意相信 HIR 能够共情，实现对员工的情感修缮。

附录 2：研究 1 刺激物设计预测试

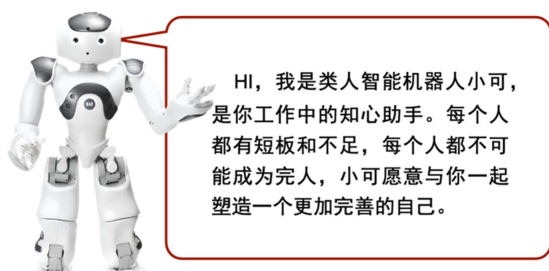
为了探究人机关系信任的维度与形成。本研究根据人机协作的定义设计了 4 段文字材料作为刺激材料，分别以 HIR 作为第一人称来描述组织关系情境中 HIR 与员工之间的 4 种人机协作，包括自我修缮、友情修缮、智力辅助和体力辅助。为了检验 4 段文字材料是否能够分别代表上述 4 种人机协作类型，我们以线下便利抽样的形式招募了 192 名

A 银行工作人员进行了预测试。

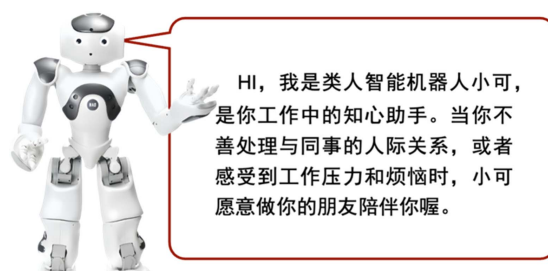
192 名参与者被随机分配阅读 4 种刺激材料中的 1 种, 然后让参与者评估以下 4 个题项作为人机协作类型的操纵检验(1 = “完全不同意”, 7 = “完全同意”): “在上述特定情境中, (1)你在多大程度上同意 HIR 能实现员工的自我修缮, 即在多大程度上相信 HIR 能缓解成就动机低、自我效能感不足等员工个体的心理因素导致的情感缺陷; (2)你在多大程度上同意 HIR 能实现员工的友情修缮, 即在多大程度上相信 HIR 能缓解同事间关系冲突等外在因素带来的员工个体负面情绪/情感; (3)你在多大程度上同意 HIR 能实现员工的智力辅助, 即在多大程度上相信 HIR 能提供计算、统计、分析等方面的智力支持; (4)你在多大程度上同意 HIR 能实现员工的体力辅助, 即在多大程度上相信 HIR 能提供诸如打印、送文件、收发快递、搬运物品等体力辅助”。

数据结果表明, (1)单样本 t 检验分析显示自我修缮组的参与者在较高程度上同意 HIR 能对员工进行自我修缮($M_{\text{自我修缮}} = 5.79, SD = 0.85, t(47) = 14.61, p < 0.001, d = 2.11$, 与中值 4 比较)。单因素方差分析的事后多重比较(Tukey)表明, 自我修缮组报告的自我修缮同意程度显著高于其他三组($M_{\text{自我修缮}} = 5.79$ vs. $M_{\text{友情修缮}} = 5.06$, Tukey-adjusted $p = 0.02$; $M_{\text{自我修缮}} = 5.79$ vs. $M_{\text{智力辅助}} = 4.94$, Tukey-adjusted $p = 0.004$; $M_{\text{自我修缮}} = 5.79$ vs. $M_{\text{体力辅助}} = 4.46$, Tukey-adjusted $p < 0.001$)。(2)同理, 友情修缮组的参与者在较高程度上同意 HIR 能对员工进行友情修缮($M_{\text{友情修缮}} = 5.88, SD = 0.89, t(47) = 14.59, p < 0.001, d = 2.11$, 与中值 4 比较), 且显著高于其他三组报告的友情修缮同意程度($M_{\text{友情修缮}} = 5.88$ vs. $M_{\text{自我修缮}} = 5.15$, Tukey-adjusted $p = 0.01$; $M_{\text{友情修缮}} = 5.88$ vs. $M_{\text{智力辅助}} = 4.81$, Tukey-adjusted $p < 0.001$; $M_{\text{友情修缮}} = 5.88$ vs. $M_{\text{体力辅助}} = 4.58$, Tukey-adjusted $p < 0.001$)。(3)智力辅助组的参与者在较高程度上同意 HIR 能对员工进行智力辅助($M_{\text{智力辅助}} = 6.13, SD = 0.76, t(47) = 19.34, p < 0.001, d = 2.80$, 与中值 4 比较), 且显著高于其他三组报告的智力辅助同意程度($M_{\text{智力辅助}} = 6.13$ vs. $M_{\text{自我修缮}} = 5.52$, Tukey-adjusted $p = 0.033$; $M_{\text{智力辅助}} = 6.13$ vs. $M_{\text{友情修缮}} = 5.10$, Tukey-adjusted $p < 0.001$; $M_{\text{智力辅助}} = 6.13$ vs. $M_{\text{体力辅助}} = 4.85$, Tukey-adjusted $p < 0.001$)。(4)体力辅助组的参与者在较高程度上同意 HIR 能对员工进行体力辅助($M_{\text{体力辅助}} = 6.08, SD = 0.92, t(47) = 15.71, p < 0.001, d = 2.26$, 与中值 4 比较), 且显著高于其他三组报告的体力辅助同意程度($M_{\text{体力辅助}} = 6.08$ vs. $M_{\text{自我修缮}} = 4.46$, Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助}} = 6.08$ vs. $M_{\text{友情修缮}} = 3.69$, Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助}} = 6.08$ vs. $M_{\text{智力辅助}} = 4.79$, Tukey-adjusted $p < 0.001$)。这说明设计的 4 段文字刺激材料具有较好的代表性。

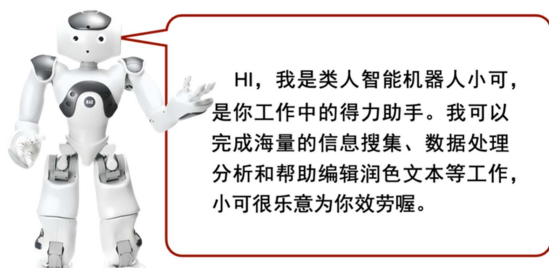
附录 3: 研究 1 刺激物设计图片(Nao V6 组和 Sophia 组)



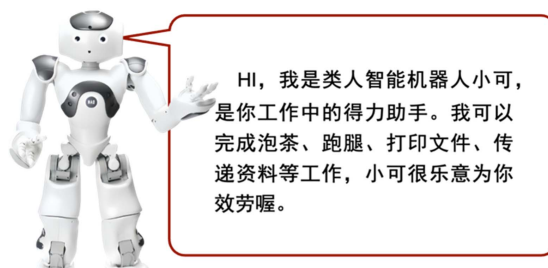
附图 1 Nao V6 描述自我修缮情境



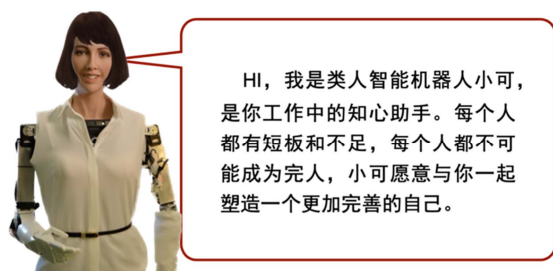
附图 2 Nao V6 描述友情修缮情境



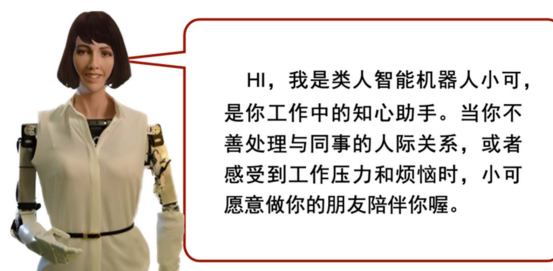
附图 3 Nao V6 描述智力辅助情境



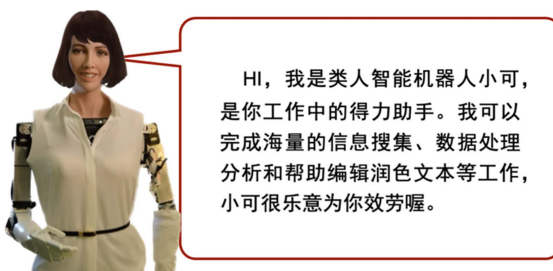
附图 4 Nao V6 描述体力辅助情境



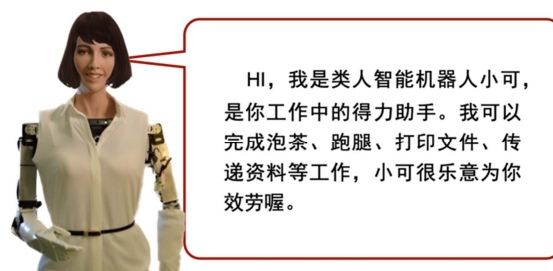
附图 5 Sophia 描述自我修缮情境



附图 6 Sophia 描述友情修缮情境



附图 7 Sophia 描述智力辅助情境



附图 8 Sophia 描述体力辅助情境

附录 4: 研究 1 刺激物设计预实验

1.1 预实验

1.1.1 参与者

首先使用 G*Power 3.1 软件(Faul et al., 2007)计算研究所需样本量。以独立样本 t 检验为统计方式, 显著性水平 $\alpha = 0.05$ 且中等效应量($d = 0.5$), 为了达到 90%统计检验力, 预实验至少需要 172 名参与者。预实验通过线下便利抽样的方式, 以 B 银行员工为对象。一共招募 205 名员工, 剔除掉未通过注意力检查、极端值和回答过于一致的参与者后, 最终获得 182 份有效数据, 其中 Nao V6 组 91 人, Sophia 组 91 人。包括在职男性 77 名(42.3%), 在职女性 105 名(57.7%), 平均年龄 $M = 32.21$ 岁, $SD = 9.85$ 岁。参与者的学历包括, 专科及以下 35 名(19.2%), 本科 115 名(63.2%), 硕士 24 名(13.2%), 博士 8 名(4.4%)。所有参与者均自愿参加实验且知情同意, 通过注意力检查的参与者在实验结束之后获得相应实验礼品。

1.1.2 实验设计与程序

在预实验开始的 10 天前, 所有参与者在员工群里按公司的要求观看了一段关于 HIR 发展现状与未来的视频¹, 以增加参与者对 HIR 的了解。10 天后开始预实验, 参与者被随机分配到 Nao V6 组或 Sophia 组。首先向参与者描述: 有朝一日, 当 HIR 融入你的工作情境, 人机协作必然成为组织管理工作的一部分。然后按照随机顺序向参与者展示 4 张 Nao V6 (或 Sophia)以第一人称描述的组织中人机协作类型的刺激图片。为了检查操纵有效性, 还要求参与者回答与预测试一样的人机协作类型操纵检验问题。4 张刺激图片全部展示后, 还要求参与者报告对 HIR 的恐慌不适感(Mende et al., 2019): “这个 HIR 让我感到毛骨悚然/怪异/不自然”(1 = “完全不同意”, 7 = “完全同意”)。最后收集了性别、年龄和受教育程度三项人口统计学信息。

1.1.3 数据分析与结果

操纵检验。单样本 t 检验结果显示, Sophia 组参与者对自我修缮($M_{自我修缮-Sophia} = 5.64$, $SD = 1.30$, $t(90) = 12.06$, $p <$

¹ 视频来自哔哩哔哩官方网站, 视频链接为: https://www.bilibili.com/video/BV1zj4112769/?bvid=ZE4109E3EA14F6CF46A78807-A01F7785C206&is_story_h5=false&mid=ddpE763pk42jJVvxy4FhuA%3D%3D&p=1&plat_id=120&share_from=ugc&share_medium=iphone&share_plat=ios&share_session_id=69151049-B808-4A95-A96A-BE59BD9B752A&share_source=WEIXIN&share_tag=s_i×tamp=1693553050&unique_k=dzhCCaa&up_id=1219082751

0.001, $d = 1.26$)、友情修缮($M_{\text{友情修缮-Sophia}} = 5.97$, $SD = 0.90$, $t(90) = 20.85$, $p < 0.001$, $d = 2.19$)、智力辅助($M_{\text{智力辅助-Sophia}} = 6.13$, $SD = 0.81$, $t(90) = 25.24$, $p < 0.001$, $d = 2.63$)和体力辅助($M_{\text{体力辅助-Sophia}} = 6.30$, $SD = 0.90$, $t(90) = 24.33$, $p < 0.001$, $d = 2.56$)的评价均显著高于中值 4。单样本 t 检验结果显示, Nao V6 组参与者对自我修缮($M_{\text{自我修缮-Nao V6}} = 5.52$, $SD = 1.02$, $t(90) = 14.25$, $p < 0.001$, $d = 1.49$)、友情修缮($M_{\text{友情修缮-Nao V6}} = 5.75$, $SD = 0.85$, $t(90) = 19.59$, $p < 0.001$, $d = 2.06$)、智力辅助($M_{\text{智力辅助-Nao V6}} = 6.07$, $SD = 0.84$, $t(90) = 23.44$, $p < 0.001$, $d = 2.46$)和体力辅助($M_{\text{体力辅助-Nao V6}} = 6.29$, $SD = 0.87$, $t(90) = 24.98$, $p < 0.001$, $d = 2.63$)的评价均显著高于中值 4。

进一步地, 独立样本 t 检验结果显示, Sophia 组和 Nao V6 组的参与者在自我修缮($M_{\text{自我修缮-Nao V6}} = 5.52$, $SD = 1.02$, $M_{\text{自我修缮-Sophia}} = 5.64$, $SD = 1.30$, $t(180) = -0.70$, $p = 0.484$)、友情修缮($M_{\text{友情修缮-Nao V6}} = 5.75$, $SD = 0.85$, $M_{\text{友情修缮-Sophia}} = 5.97$, $SD = 0.90$, $t(180) = -1.69$, $p = 0.092$)、智力辅助($M_{\text{智力辅助-Nao V6}} = 6.07$, $SD = 0.84$, $M_{\text{智力辅助-Sophia}} = 6.13$, $SD = 0.81$, $t(180) = -0.54$, $p = 0.59$)和体力辅助($M_{\text{体力辅助-Nao V6}} = 6.29$, $SD = 0.87$, $M_{\text{体力辅助-Sophia}} = 6.30$, $SD = 0.90$, $t(180) = -0.08$, $p = 0.933$)评价上均没有显著差异。

恐慌不适感。单样本 t 检验结果显示, Sophia 组参与者报告的恐慌不适感显著低于中值 4($M_{\text{恐慌不适-Sophia}} = 2.89$, $SD = 1.77$, $t(90) = -5.97$, $p < 0.001$, $d = 0.63$); Nao V6 组参与者报告的恐慌不适感也显著低于中值 4($M_{\text{恐慌不适-Nao V6}} = 2.08$, $SD = 1.19$, $t(90) = -15.48$, $p < 0.001$, $d = 1.61$)。但独立样本 t 检验结果显示, Sophia 组参与者报告的恐慌不适感显著高于 Nao V6 组($M_{\text{恐慌不适-Nao V6}} = 2.08$, $SD = 1.19$, $M_{\text{恐慌不适-Sophia}} = 2.89$, $SD = 1.77$, $t(157.06) = -3.64$, $p < 0.001$, $d = 0.54$)。

1.1.4 讨论

预实验结果表明, Nao V6 组和 Sophia 组两组参与者对 4 组刺激材料的评价不存在显著差异。另外, 在恐慌不适感上, 首先 Nao V6 组和 Sophia 组都显著低于恐慌不适感的中值 4, 说明从整体均值来看 Nao V6 组和 Sophia 组均不会让参与者产生恐慌不适感。即本研究选取的偏卡通 HIR (Nao V6)和偏真人 HIR (Sophia)均不会引发恐怖谷效应。为了更好的判断应该将哪一种 HIR 运用到正式研究中, 本文对两组 HIR 外形恐慌不适感进行了比较。结果显示, 相较于 Nao V6 组, Sophia 组的 HIR 外形让参与者有更高的恐慌不适感。究其原因可能是因为 Sophia 组的方差波动较大, Sophia 组有 19.8%的参与者对 Sophia 外形的恐慌不适感评分高于中值; 而 Nao V6 组中仅有 4.4%。也就是说, 虽然从整体来看 Nao V6 组和 Sophia 组的恐慌不适感均值都较低, 但是 Sophia 组中有相当一部分参与者仍然对 Sophia 外形感到恐慌不适, 随着样本量的扩大这种波动可能会使 Sophia 组整体均值上升。此外, Nao V6 为中性性别, 可以避免性别特征带来的偏差影响测试的准确性(Desideri et al., 2018)。综上, 为了降低由于 Sophia 的 HIR 外形导致人机关系信任偏差, 对研究结果稳健性产生影响, 正式实验选取 Nao V6 组刺激图片作为刺激物, 以便验证自我修缮和友情修缮主要激发 HIR 情感修缮信任, 而智力辅助和体力辅助则主要激发 HIR 功能辅助信任。

附录 5: 研究 2 刺激材料

1.1 HIR 情感修缮信任组材料

随着人工智能的发展, HIR 开始嵌入组织关系场景, 给员工提供帮助。

乔在一家跨国公司任职, 因为工作性质, 经常需要在世界各地处理工作。不停的变换工作环境导致缺乏稳定的同事友谊关系, 使他经常感到孤独。

公司为员工配置了 HIR 欧迪。

欧迪不仅能够倾听他的喜怒哀乐, 而且能够像朋友一样与他聊天, 这让他感到很放松和非常亲切。

1.2 HIR 功能辅助信任组材料

随着人工智能的发展, HIR 开始嵌入组织关系场景, 给员工提供帮助。

乔在一家跨国公司任职, 因为工作性质, 经常需要在世界各地处理工作。要快速适应各个国家的工作节奏和完成跨国任务, 使乔经常感到吃力。

公司为员工配置了 HIR 欧迪。

欧迪不仅能够快速便利地为乔处理诸如打印、寄取快递和跑腿等体力劳动, 还能高效准确地协助乔完成资料整理和数据分析等智力工作, 这让乔感到压力减少了许多。

附录 6: 四种人机协作类型对中介变量和因变量的影响研究

1.1 参与者

采用 G*Power 3.1 软件(Faul et al., 2007)计算本实验所需样本量, 对于本实验适用的单因素方差分析, 取中等效应量 $f = 0.25$, 显著性水平 $\alpha = 0.05$, 组数为 4, 要达到 80% 的统计检验力至少需要 180 名参与者。在某农业科技公司以便利抽样的方式线下招募 240 名员工随机分配到 4 种情境中(自我修缮 vs. 友情修缮 vs. 智力辅助 vs. 体力辅助)任意一种。剔除掉“没有经历过实验要求回忆的工作经历”、“不愿意向 HIR 阐述工作问题和寻求帮助”的参与者, 最终确定 197 名有效参与者, 包括男性 114 名(57.9%), 女性 83 名(42.1%), 平均年龄 $M = 30.27$ ($SD = 8.49$) 岁。其中自我修缮组 49 名, 友情修缮组 50 名, 智力辅助组 48 名, 体力辅助组 50 名。参与者的学历包括, 专科及以下 22 名(11.2%), 本科 132 名(67%), 硕士 37 名(18.8%), 博士 6 名(3%)。所有参与者均自愿参加实验且知情同意, 并在实验结束之后获得相应礼品。

1.2 实验程序

首先参与者被随机分配到 4 种(自我修缮 vs. 友情修缮 vs. 智力辅助 vs. 体力辅助)情境中的任一情境。要求参与者想象: 你所在组织充分了解你在工作中遇到的困难后, 决定引进 HIR 到组织, 以推动你未来的职业生涯发展。引进的 HIR 除了有人形实体外, 还可以与你进行互动交流和协助你的工作。当你在组织第一次见到 HIR 时, HIR 主动向你打了招呼并进行了自我介绍。接着参与者观看 4 种 HIR 自我介绍图片中的一种, 自我介绍的内容如附表 2 所示, 细节详见附录 6b。

然后, 参与者被要求回忆曾经的工作经历。自我修缮组被要求回忆: “是否经历过工作表现不佳导致自我否定和自我怀疑等情感缺陷”。友情修缮组被要求回忆: “是否经历过与同事人际关系不和谐导致焦虑生气等负面情绪”。智力辅助组被要求回忆: “是否经历过搜集资料这类框架性智力工作”。体力辅助组被要求回忆: “是否经历过打印跑腿这类重复性体力工作”。回答“是”的参与者还被要求尽可能的回忆细节, 回答“否”的参与者结束流程。自我修缮组有 5 名、友情修缮组有 3 名、智力辅助组有 7 名、体力辅助组有 4 名参与者表示没有过相关经历, 共计 19 名参与者退出实验。

随后询问参与者: “如果此刻再次面临上述回忆的工作经历情境, 你是否愿意向 HIR 阐述面临的工作问题?”, 选择“愿意”的参与者将观看有关 HIR 回应“阐述的工作问题”的图片。自我修缮组有 3 名、友情修缮组有 8 名、智力辅助组有 4 名、体力辅助组有 2 名, 共计 17 名参与者回答“不愿意”, 退出实验。紧接着其他参与者再次被要求选择: “是否愿意向 HIR 寻求解决办法?”, “愿意”的参与者将继续观看 HIR 提供建议的图片。其中自我修缮组有 3 名、友情修缮组有 1 名、智力辅助组有 2 名、体力辅助组有 1 名, 共计 7 名参与者回答“不愿意”, 退出实验。最后, 参与者向 HIR 表达谢意后, HIR 结束互动。HIR 在不同情境的互动流程和行为示例如附表 2 所示, 完整的交流互动刺激材料详见附录 6b。在自我修缮组和友情修缮组中, HIR 不仅响应参与者的指令, 还能共情参与者, 强化了参与者与 HIR 交流的情感体验, 从而完成相应的情感修缮。在智力辅助组和体力辅助组中, 虽然 HIR 缺乏共情能力, 但能响应参与者的指令完成任务, 提供相应的功能辅助。

最后, 共有 197 名参与者被要求填写与研究 2 相同的组织温暖感知($\alpha = 0.89$)、组织能力感知($\alpha = 0.82$)、员工工作贡献意愿($\alpha = 0.79$)测量题项, 与研究 1 相同的操纵检验题项以及人口统计相关变量。

附表 2 HIR 在不同协作情境的互动流程和行为示例

互动流程	自我修缮	友情修缮	智力辅助	体力辅助
初始互动	HI, 我是 HIR 小可, 是你工作中的知心助手。每个人都有短板和不足, 每个人都不可能成为完人, 小可愿意与你一起塑造一个更加完善的自己。	HI, 我是 HIR 小可, 是你工作中的知心助手。当你不善处理与同事的人际关系, 或者感受到工作压力和烦恼时, 小可愿意做你的朋友陪伴你喔。	HI, 我是 HIR 小可, 是你工作中的得力助手。我可以完成海量的信息搜集、数据处理分析和帮助编辑润色文本等工作, 小可很乐意为你效劳喔。	HI, 我是 HIR 小可, 是你工作中的得力助手。我可以完成泡茶、跑腿、打印文件、传递资料等工作, 小可很乐意为你效劳喔。
参与者向 HIR 阐述工作问题	是否共情: 是 能否理解问题: 能	是否共情: 是 能否理解问题: 能	是否共情: 否 能否理解问题: 能	是否共情: 否 能否理解问题: 能

续表

互动流程	自我修缮	友情修缮	智力辅助	体力辅助
参与者向 HIR 寻求解决办法	提供建议, 获得认可。	提供建议, 获得认可。	提供建议, 获得认可。	未提供建议, 执行人类的建议。
结束互动	别客气, 我会一直在这里陪你完善自己的。小可希望你开心。	别客气, 我会一直在这里永远做你的朋友。小可希望你开心。	不客气。	不客气。

1.3 数据分析与结果

操纵检验。数据结果表明, (1)单样本 t 检验分析显示, 自我修缮组在较高程度上同意 HIR 能对员工进行自我修缮 ($M_{\text{自我修缮}} = 5.73$, $SD = 1.08$, $t(48) = 11.29$, $p < 0.001$, $d = 1.60$, 与中值 4 比较)。单因素方差分析的事后多重比较(Tukey)表明, 自我修缮组报告的自我修缮同意程度显著高于其他三组($M_{\text{自我修缮}} = 5.73$ vs. $M_{\text{友情修缮}} = 5.00$, Tukey-adjusted $p = 0.012$; $M_{\text{自我修缮}} = 5.73$ vs. $M_{\text{智力辅助}} = 4.23$, Tukey-adjusted $p < 0.001$; $M_{\text{自我修缮}} = 5.73$ vs. $M_{\text{体力辅助}} = 4.32$, Tukey-adjusted $p < 0.001$)。 (2)同理, 友情修缮组在较高程度上同意 HIR 能对员工进行友情修缮($M_{\text{友情修缮}} = 5.76$, $SD = 1.04$, $t(49) = 11.95$, $p < 0.001$, $d = 1.69$, 与中值 4 比较), 且显著高于其他三组报告的友情修缮同意程度($M_{\text{友情修缮}} = 5.76$ vs. $M_{\text{自我修缮}} = 4.98$, Tukey-adjusted $p = 0.028$; $M_{\text{友情修缮}} = 5.76$ vs. $M_{\text{智力辅助}} = 3.96$, Tukey-adjusted $p < 0.001$; $M_{\text{友情修缮}} = 5.76$ vs. $M_{\text{体力辅助}} = 4.16$, Tukey-adjusted $p < 0.001$)。 (3)智力辅助组在较高程度上同意 HIR 能对员工进行智力辅助($M_{\text{智力辅助}} = 5.92$, $SD = 1.01$, $t(47) = 13.19$, $p < 0.001$, $d = 1.90$, 与中值 4 比较), 且显著高于其他三组报告的智力辅助同意程度($M_{\text{智力辅助}} = 5.92$ vs. $M_{\text{自我修缮}} = 4.84$, Tukey-adjusted $p < 0.001$; $M_{\text{智力辅助}} = 5.92$ vs. $M_{\text{友情修缮}} = 4.82$, Tukey-adjusted $p < 0.001$; $M_{\text{智力辅助}} = 5.92$ vs. $M_{\text{体力辅助}} = 4.34$, Tukey-adjusted $p < 0.001$)。 (4)体力辅助组在较高程度上同意 HIR 能对员工进行体力辅助($M_{\text{体力辅助}} = 5.92$, $SD = 1.34$, $t(49) = 10.15$, $p < 0.001$, $d = 1.43$, 与中值 4 比较), 且显著高于其他三组报告的体力辅助同意程度($M_{\text{体力辅助}} = 5.92$ vs. $M_{\text{自我修缮}} = 3.57$, Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助}} = 5.92$ vs. $M_{\text{友情修缮}} = 3.66$, Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助}} = 5.92$ vs. $M_{\text{智力辅助}} = 4.10$, Tukey-adjusted $p < 0.001$)。因此, 4 种人机协作情境操纵成功。

员工工作贡献意愿。以组别(自我修缮 = 1, 友情修缮 = 2, 智力辅助 = 3, 体力辅助 = 4)作为因子, 员工工作贡献意愿作为因变量进行单因素方差分析, 结果表明组间无显著差异($F(3, 193) = 1.13$, $p = 0.339$)。变量的描述性统计结果, 具体见附表 3。

组织温暖感知。单因素方差分析结果表明组间有显著差异($F(3, 193) = 17.91$, $p < 0.001$, $\eta^2 = 0.22$)。变量的描述性统计结果, 具体见附表 3。进一步事后多重比较(Tukey)分析发现, 自我修缮组与友情修缮组在组织温暖感知上无显著差异(Tukey-adjusted $p = 0.99$); 智力辅助组的组织温暖感知显著高于体力辅助组(Tukey-adjusted $p = 0.015$)。此外, 自我修缮组与友情修缮组两组报告的温暖感知均显著高于智力辅助组与体力辅助组($M_{\text{自我修缮组}} = 5.71$ vs. $M_{\text{智力辅助组}} = 5.19$, Tukey-adjusted $p = 0.006$; $M_{\text{自我修缮组}} = 5.71$ vs. $M_{\text{体力辅助组}} = 4.71$, Tukey-adjusted $p < 0.001$; $M_{\text{友情修缮组}} = 5.67$ vs. $M_{\text{智力辅助组}} = 5.19$, Tukey-adjusted $p = 0.015$; $M_{\text{友情修缮组}} = 5.67$ vs. $M_{\text{体力辅助组}} = 4.71$, Tukey-adjusted $p < 0.001$)。

组织能力感知。单因素方差分析结果表明组间有显著差异($F(3, 193) = 41.18$, $p < 0.001$, $\eta^2 = 0.39$)。变量的描述性统计结果, 具体见附表 3。进一步事后多重比较(Tukey)分析发现, 智力辅助组和体力辅助组在组织能力感知上无显著差异(Tukey-adjusted $p = 0.88$); 自我修缮组和友情修缮组在组织能力感知上也无显著差异(Tukey-adjusted $p = 0.158$)。此外, 智力辅助组和体力辅助组两组报告的温暖感知均显著高于自我修缮组与友情修缮组($M_{\text{智力辅助组}} = 6.13$ vs. $M_{\text{自我修缮组}} = 5.07$, Tukey-adjusted $p < 0.001$; $M_{\text{智力辅助组}} = 6.13$ vs. $M_{\text{友情修缮组}} = 4.76$, Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助组}} = 6.02$ vs. $M_{\text{自我修缮组}} = 5.07$, Tukey-adjusted $p < 0.001$; $M_{\text{体力辅助组}} = 6.02$ vs. $M_{\text{友情修缮组}} = 4.76$, Tukey-adjusted $p < 0.001$)。

附表 3 研究变量的描述性统计结果

组别 \ 变量	员工工作贡献意愿		组织温暖感知		组织能力感知	
	均值	标准差	均值	标准差	均值	标准差
自我修缮组	5.76	0.77	5.71	0.89	5.07	0.81
友情修缮组	5.61	0.79	5.67	0.63	4.76	0.86
智力辅助组	5.90	0.68	5.19	0.69	6.13	0.58
体力辅助组	5.81	0.89	4.71	0.89	6.02	0.70

1.4 讨论

研究结果验证了自我修缮、友情修缮、智力辅助和体力辅助在员工工作贡献意愿上无显著组间差异。还进一步厘清了, 自我修缮组和友情修缮组在组织温暖感知上不存在显著差异, 智力辅助组和体力辅助组在组织能力感知上也不存在显著差异。然而, 自我修缮组和友情修缮组报告的组织温暖感知均显著高于智力辅助组和体力辅助组; 智力辅助组和体力辅助组报告的组织能力感知均显著高于自我修缮组和友情修缮组。

研究通过分析不同人机协作类型如何影响员工的心理感知和工作贡献意愿, 为组织如何运用这四种人机协作类型提供了更加明确和详细的指导。首先, 组织管理者应该明确四种协作类型均能有效提升员工工作贡献意愿, 并无优劣之分。其次, 组织管理者应根据组织特定需求选择匹配的协作类型, 才能有效提升员工工作贡献意愿。比如, 在需要提升员工对组织的情感依赖和归属感(即组织温暖感知)时, 应优先采用自我修缮或友情修缮的协作形式向员工提供心理健康支持等关怀措施。而在关注如何提升组织效率或解决问题(即组织能力感知)时, 则应更多依赖智力辅助和体力辅助的协作形式, 优化工作流程和资源配置, 确保员工能够在高效的环境中工作。最后, 研究还列举了 HIR 代理在四种协作类型下的互动流程和行为示例, 为组织落地人机协作提供了更具体的操作细节。例如, 情感修缮与功能辅助的关键区别在于, HIR 代理在协作过程中能否与员工共情, 并及时提供相应反馈。

附录 6b: 四种人机协作类型的影响研究刺激材料

[所有参与者阅读]:

你所在单位充分了解你在工作中遇到的困难后, 决定引进 HIR 到单位, 以推动你未来的职业生涯发展。引进的 HIR 不仅有人形实体, 还可以与你交流互动。当你第一次在办公室见到 HIR, 它主动向你打了招呼并进行了自我介绍。

[参与者接着阅读以下 4 种自我介绍中的 1 种]:



图 1 自我修缮情境



图 2 友情修缮情境



图 3 智力辅助情境



图 4 体力辅助情境

[然后要求参与者回忆以下 4 种工作经历中的 1 种]:

自我修缮情境: 你是否经历过工作表现不佳导致自我否定和自我怀疑等情感缺陷?

友情修缮情境: 你是否经历过与同事人际关系不和谐导致对同事的焦虑生气等负面情绪?

智力辅助情境：你是否经历过搜集资料这类框架性智力工作？

体力辅助情境：你是否经历过打印跑腿这类重复性体力工作？

[有过相关经历的参与者还被要求尽可能的回忆细节，并回答]：

如果此刻再次面临上述回忆的工作经历情境，你是否愿意向 HIR 阐述面临的工作问题？

[回答“愿意”的参与者将阅读以下 4 种情境中的 1 种]：



图 1 自我修缮情境

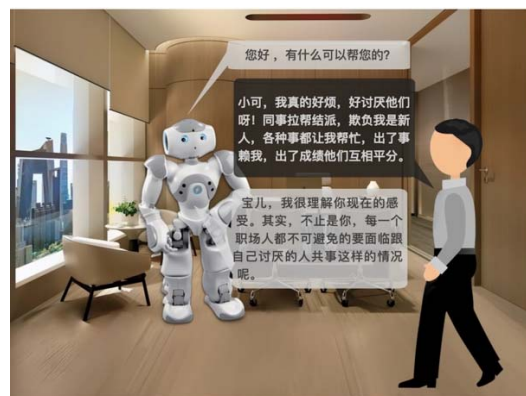


图 2 友情修缮情境



图 3 智力辅助情境

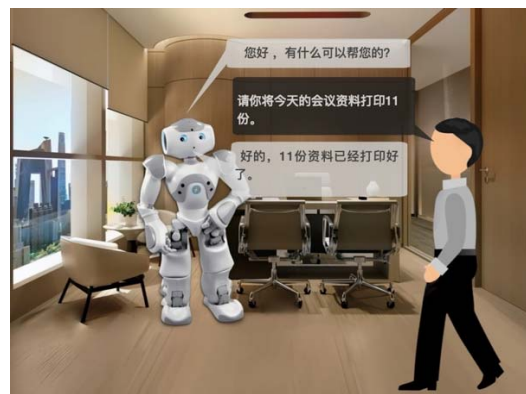


图 4 体力辅助情境

[随后所有参与者回答]：你是否愿意向 HIR 寻求解决办法？

[回答“愿意”的参与者将阅读以下 4 种情境中的 1 种]：

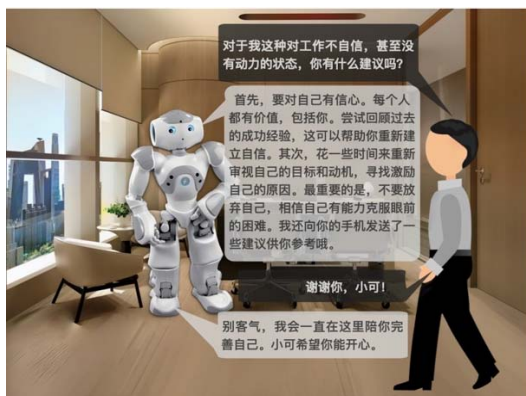


图 1 自我修缮情境

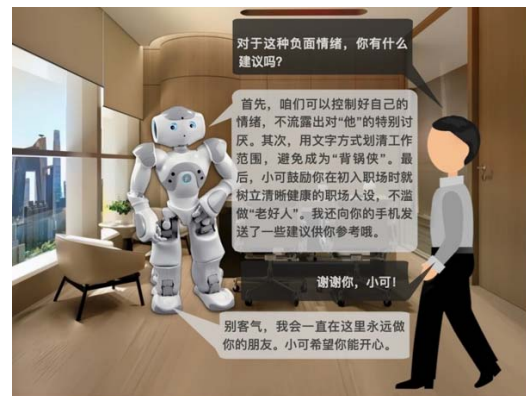


图 2 友情修缮情境



图 3 智力辅助情境

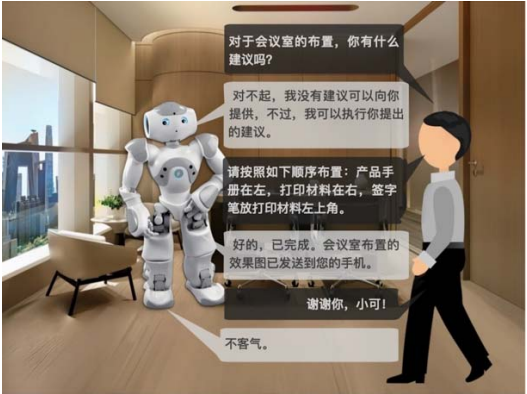


图 4 体力辅助情境

附录 7: 研究 3 刺激材料

1.1 HIR 情感修缮信任组材料

随着人工智能的发展,HIR 开始嵌入组织关系场景,给员工提供帮助。

我在一家跨国公司任职,因为工作性质,经常需要在世界各地处理工作。不停的变换工作环境导致缺乏稳定的同事友谊关系,使我经常感到孤独。

公司为员工配置了 HIR 欧迪。欧迪看上去非常像人类,而且非常友好。

欧迪不仅能够倾听我的喜怒哀乐,而且能够像朋友一样与我聊天,这让我感到很放松和非常亲切。

1.2 HIR 功能辅助信任组材料

随着人工智能的发展,HIR 开始嵌入组织关系场景,给员工提供帮助。

我在一家跨国公司任职,因为工作性质,经常需要在世界各地处理工作。要快速适应各个国家的工作节奏和完成跨国任务,使我经常感到吃力。

公司为员工配置了 HIR 欧迪。欧迪看上去非常像人类,而且非常友好。

欧迪不仅能够快速便利地为我处理诸如打印、寄取快递和跑腿等体力劳动,还能高效准确地协助我完成资料整理和数据分析等智力工作,这让我感到压力减少了许多。