

多元文化经历提升人类对机器人的利他行为 及心智知觉的中介作用*

滕 玥^{1,2#} 张昊天^{3#} 赵偲琪^{1,2} 彭凯平⁴ 胡晓檬^{1,2}

(¹中国人民大学心理学系, 北京 100872) (²中国人民大学哲学与认知科学跨学科交叉平台, 北京 100872)

(³武汉大学哲学学院心理学系, 武汉 430072) (⁴清华大学心理学系, 北京 100084)

摘 要 人类步入人工智能时代, 人工智能快速迭代、迅猛发展、不断变革, 引发诸多伦理问题。其中一个核心关切便是在人类与 AI 和谐共生的指导思想之下, 人类是否会对机器人持有与人类同等的道德关怀并对其做出利他行为。目前, 鲜有研究探讨哪些文化心理因素提高人类对机器人的利他行为及其背后的深层心理机制。本研究旨在从多元文化经历的视角开展科学探索。基于中西两种文化样本的研究发现, 个体的多元文化经历正向预测人类对机器人的利他行为, 人对机器人的心智感知起部分中介作用。此中介模型在中西文化被试中不存在显著文化差异, 表明该模型具备一定的文化普适性。最后, 线上启动实验研究虽未能证实因果机制, 但存在符合研究假设的总体趋势。本研究丰富了多元文化经历的下流效应, 对探讨何种前因变量会提高人类对机器人的利他行为及其解释机制做出了一定的原创贡献。

关键词 多元文化经历, 人机交互, 利他行为, 心智知觉, 文化普适

分类号 B849: C91

1 引言

人工智能革命浪潮正在席卷全球, 然而心理学与人工智能交叉地带的基础研究和成果转化亟需积累更多理论指引和科学支撑。人工智能时代的机器人应用已然成为人类社会不可或缺的组成部分。从最贴近人们日常生活的扫地机器人、服务型机器人, 到执行高风险与高心智能力任务的军事机器人、医学诊断机器人等, 机器人的普及程度和对人类社会的重大影响日渐增加。尽管人工智能在技术和应用上的迅猛发展大大提升了解决现实问题的效率与质量, 但也不可避免地带来了复杂的道德伦理问题。专攻认知科学与人工智能的哲学家 Klineciewicz (2016) 认为, 今时今日的人们想要过上“道德生活”的难度远高于古人所处的年代, 这是因为我们从祖先那里继承的道德理论与道德实践已

经与快速更新迭代的科技生活脱节, 无法满足当代社会所出现的新的道德问题, 这就是所谓的“道德滞后” (moral lag) 现象 (丁晓军 等, 2020)。因此, 心理学家应当前瞻性的探索人工智能给人类社会带来的道德伦理问题, 并与哲学、社会学、工程师、人工智能产业等开展合作进行对话性的研究, 以期更好的发挥人工智能的优势, 规避可能带来的伦理风险, 构建人与人工智能的良好生态关系。

1.1 机器人是否可能拥有道德地位?

人与机器人的交互接触是双向的, 机器人在危机管理、军事航空、医疗、教育等领域为人类提供极大的便利 (Bigman & Gray, 2018)。由此, 基于人类固有的互惠利他原则, 人们开始呼吁人类也应当关心机器人的权益, 甚至赋予机器人人类的道德考量, 即给予机器人与人类同等程度的道德关怀和道德权益 (Lund et al., 2007)。谷歌工程师 Blake

收稿日期: 2022-09-04

* 中国人民大学“双一流”跨学科重大创新规划平台支持。

滕玥和张昊天为本文的共同第一作者

通信作者: 胡晓檬, E-mail: xiaomenghu@ruc.edu.cn

Lemolne 公开呼吁 AI 有灵魂有意识, 应该享受与人类一样的基本权利, 这位工程师甚至打算帮助聊天机器人 LaMDA 聘请律师以争取其应有的合法权益, 保障其享有法律所赋予“人”的权力。心理学家基于“计算机是社会行动者”(computers are social actors) 框架(Nass & Moon, 2000), 提出人类对机器人的行为是亲社会模式的。也就是说, 当人们以亲社会的方式对待人类和机器时, 动机和认知过程是类似的, 其中包括将机器视为人类, 将机器进行社会分类, 受到机器的社会影响, 以及对机器体验社会情绪, 比如共情等(Nielsen et al., 2021)。当人们将机器人视为与人类同等的对象时, 一个重要的问题随之产生, 人类对于与自己同属于人类的外群体成员都不一定将其纳入道德圈的范畴, 而“类人”的机器人群体又是否会拥有和人类同等的道德地位呢? 因此, 伦理学家提出了“伦理行为主义”(ethical behaviourism)理论, 该理论强调机器人在人类心目中拥有道德地位的可能性, 如果机器人在行为表现上与其他具有重要道德地位的实体大致相当, 那么机器人就可以拥有道德地位(Danaher, 2019)。也就是说, 人们会在特定的情况下将机器人纳入道德考量的范围。还有研究者认为, 从道德进化论的角度而言, 人类的道德圈拓展是一个漫长的进化过程, 受到多种因素的影响, 包括文化、技术、社会制度等。未来社会中, 道德圈的拓展将持续加大, 可能会包括更广泛的人类群体、机器、人工智能等范畴(Anthis & Paez, 2021)。本研究旨在探讨哪些文化心理因素可能提高人类对机器人的利他行为。从心理学的角度理解人类是否将机器人纳入道德考量的范围将有助于完善现有的人工智能伦理框架中对机器人身份的界定问题: 对于人类而言, 机器人究竟仅仅是工具, 还是值得被善意对待的实体存在?

1.2 人类对机器人利他行为的影响因素

当人类将一些人或物纳入道德考量的范围后, 人们愿意赋予这些人或物亲社会的意愿和行为。比如, 人们会将人道主义和环境问题优先于个人的利益, 牺牲自己的生命来拯救其他人或动植物(Crimston et al., 2016)。当人们将机器人纳入道德考量的范围后, 一个最重要且直观的体现就是人们愿意对机器人做出道德行为。利他行为是最基础的道德行为(Fabes et al., 1999; Osswald et al., 2010)。在日常生活中, 人们可能对机器人做出一些微小的利他行为, 比如避开上菜机器人的行进路线, 帮助扫地机器人处理其无法触及的区域等。

人类对机器人的利他行为受到诸多因素的影响。已有研究表明, 机器人自身的物理和行为特征会影响人们对其的利他态度和利他行为。相比那些人为遥控的机器人, 当机器人表现得更具自主性(autonomy)时, 人们会更快地响应机器人的求助请求(Somanader et al., 2011; Srinivasan & Takayama, 2016)。Lee 和 Liang (2016)发现, 如果机器人在之前的问答游戏中做出了有用的帮助行为, 那么被试会更愿意帮助它, 说明人类对机器人的利他行为也会遵循互惠原则。此外, 机器人自身的社会类别线索也会影响人们对其的利他行为。Siegel 等人(2009)的研究表明, 与发出男性声音机器人相比, 男性对发出女性声音的机器人有更多的捐赠行为。除了机器人的物理与行为特征外, 个体自身的特征也会影响人类对机器人的利他行为。例如, 拟人化倾向(指将人类的外部特征与内在心理状态赋予机器人的倾向)高的个体会对机器人有更多的共情, 从而提升对机器人的接受度和利他行为(Kühnlénz et al., 2013; Nijssen et al., 2019; Nielsen et al., 2022)。另外, “诚实-谦逊(Honesty-Humility)”人格倾向较高的个体会在博弈游戏中给 AI 分更多的钱(Nielsen et al., 2022)。然而, 前人研究关注人类自身特征对机器人利他行为的研究相对较少, 需要进一步从理论阐述和实证层面探索哪些个体因素会影响人类对机器人的利他行为, 以及背后的深层心理机制如何。

1.3 多元文化经历提高人类对机器人的利他行为

人类对机器人的利他行为是人类给予机器人道德考量的基本行为体现, 哪些个体因素可以促进这种利他行为呢? 利他的本质是平等对待他人的同理心和关注他人利益与福祉的道德观念(Batson, 2011), 越来越多的研究表明, 道德判断与道德行为并不是理性的过程, 日常生活中的道德判断是更情绪化及更灵活的过程(Pizarro et al., 2006)。利他行为作为一种道德行为, 也会受到很多非理性因素的影响。个体的生活经历, 尤其是多元文化经历可以增加人们对新事物的接纳以及认知上的灵活性(Tadmor et al., 2006)。因此, 本研究聚焦多元文化经历对机器人利他行为的影响, 以丰富对机器人利他行为提升机制的实证探索。

多元文化经历是指人们直接或间接与外国文化元素或文化成员接触或互动的体验(Leung et al., 2008), 强调个体在不同层次文化中的综合体验。胡晓檬等人(2021)系统梳理了多元文化经历的概念界定、测量工具、操纵范式以及多元文化经历对个体

心理和行为产生的“双刃剑效应”。多元文化经历可能会带来一些负面的心理与行为后果,比如不道德行为(Lu et al., 2017),文化震荡与外群体排斥(Morris et al., 2015)等。同时,多元文化经历也会带来诸多积极影响。已有研究发现,多元文化经历可能会减少偏见和歧视(Tadmor et al., 2012),增加个体对他人的广义信任(Cao et al., 2013),提高创造力(Leung et al., 2008),更能容忍不确定性(Adair & Xiong, 2018)和看似矛盾的科学陈述(Hu et al., 2020)。从人机互动的角度而言,人们对机器人的反应往往类似于对人类外群体(例如,移民或少数民族)的反应(Smith et al., 2021)。比如,与人类外群体相似,机器人可能也会取代人类的工作或对人类造成身体上的伤害(de Graaf & Allouch, 2016; Nam, 2019),从而引起人们的恐惧和排斥心理;人们还认为机器人的价值观与人类不同,在道德困境中,机器人被认为会做出比人类更功利主义的道德决策,例如将失控的电车引向只会杀死一人而不是五人的轨道(Malle et al., 2015)。而多元文化经历会通过“认知冻融(epistemic unfreezing)”来打破对外群体的偏见和歧视,因此这一效应也可能扩展到机器人这一外群体身上。也就是说,多元文化经历更丰富的个体会以更积极和开放的态度和行为对待机器人。

除此之外,正如前文所述,多元文化经历的另一大优势是会提升个体的认知灵活性(Tadmor et al., 2006),不受固着思维的束缚,拓宽认知边界,这种灵活性也会使个体从不同的角度看待问题,甚至可能拓宽个体的同理心和理解范围(Park & Yu, 2017),使他们更有可能把机器人视为值得关心和尊重的对象;这种灵活性也可能延伸到道德领域,如创造性会提升个体道德判断的灵活性(Pizarro et al., 2006)。对于多元文化经历更丰富的个体而言,他们的认知更灵活而不僵化。对于机器人这样一种新异事物,具有丰富多元文化经历的个体更有可能打破机器人只是一个无生命的物体这样一种狭隘的认知,更有可能将机器人视作“类人”的生物实体,将其纳入道德考量的范围并对其做出更多的利他行为。据此,我们提出假设 1:多元文化经历促进人类对机器人的利他行为。

1.4 多元文化经历促进人类对机器人利他行为的心理机制——心智知觉的中介作用

心智知觉(mind perception)是影响人类对机器人利他行为的重要前因变量(Tanibe et al., 2017)。人

们在不同的实体中都能感知到心智,包括人类、非人类(例如动物等)和非生命体(例如机器人等)(Gray et al., 2007)。Gray 等人(2012)认为人们从两个维度感知心智:体验性(experience),即拥有感知情绪体验的能力,如感觉到饥饿、恐惧、痛苦、快乐和意识等;能动性(agency),即拥有主观意图去做出行为的感知能力,如自控、判断、交流、思考和记忆等。研究也揭示了心智知觉与道德判断有显著的关联:对于感知到具有体验性心智能力的实体,人们会赋予其更多的道德权益(moral rights);对于感知到具有能动性心智能力的实体,人们会赋予其更多的道德责任(moral responsibility)(Gray et al., 2012)。

对于日益更新迭代的人机交互模式而言,多元文化经历更丰富的个体更有可能想象机器人具有“类人”的心理特征,改变对机器人的态度,提升其对机器人的心智知觉。如上文所述,多元文化经历的人有更高的创造力和认知灵活性,对世界及事物的多元性有更丰富的感知,对变化与不确定性有更强的适应性和安全感。这些更丰富的感知度、创造力与灵活性会使个体增加对新事物的接纳、尊重和应用,这同样会体现在人类对机器人的心智感知上,尤其是面对机器人需要帮助的情境时,更能够激发人类对机器人的心智感知。另一方面,多元文化经历会减少对外群体的偏见与歧视,更少以刻板的特征对某个群体进行不公正的判断。因此,多元文化经历更丰富的个体可能更能摆脱“机器人只能是机器”的刻板印象,并减少对机器人“非人化(dehumanization,即对人类基本特征的否定)”倾向,从而更倾向于将机器人纳入内群体并赋予机器人以人的特性。根据心智感知理论,一旦人类对机器人的感知相似度(perceived similarity)增加,并产生试图与机器人建立社会联结(social connection)的动机时,便会触发心智感知(Waytz et al., 2010),即认为机器人更加富有思想与情感。再者,多元文化经历更丰富的个体在更广泛的情境中具有较高的广义信任,在机器人情境下,这种广义信任可能使得人类不再倾向将智能机器人视为威胁,而对其持有更友善、平等、开放的态度,从而更倾向于赋予机器人以更多的人类心智特征。

另外,根据以往研究,多元文化习得促使个体更加强调关爱、公正和平等主义等保护个体权益的价值观(Hu et al., 2020),进而会更多地增加知觉到的机器人道德权益,使人与机器人的关系更加平

等。需要说明的是,我们认为多元文化经历可以提升机器人的“主体(agent)”地位,但此主体并非道德主体(moral agent),而是指机器人有资格成为有心智的主体并能够获得道德权益。而以往研究也表明,增加对机器人的心智感知则会进一步赋予机器人人格地位,并增加对机器人的利他行为(Tanibe et al., 2017)。因此,当我们探讨人类如何对待机器人时,感知到的机器人的心智知觉将是一个重要维度,显著影响后续人们是否会对机器人做出利他行为。据此,我们提出假设 2:多元文化经历通过提高对机器人的心智知觉,进而提高对机器人的利他行为(见图 1)。

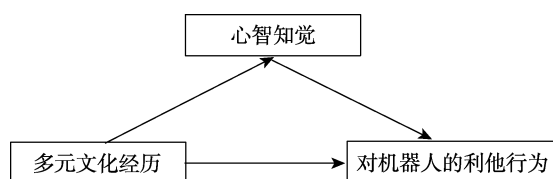


图1 本研究理论模型图

1.5 多元文化经历提升对机器人利他行为的跨文化普适性

综上所述,本研究希望探究多元文化经历对机器人利他行为的影响以及背后可能存在的心理机制。除此之外,本研究也将探讨来自不同文化背景的多元文化经历的个体对机器人的利他行为是否存在显著的中西方文化差异,即多元文化经历通过心智知觉提升对机器人的利他行为这一逻辑链条在不同文化中是否存在可推广性。已有研究表明,文化是影响人机交互过程的重要因素。研究发现,人们对机器人的态度和行为存在明显的文化差异。例如,相较于西方被试,日本被试对机器人的刻板印象更积极(Kaplan, 2004),但也有研究表明,虽然日本被试对传统机器人的评价比西方被试高,但对于高度拟人化的机器人,这种情况则刚好相反(Bartneck, 2008)。相比于西方国家的民众(如美国),东亚国家的民众(如日本)对机器人的接受度更低(Sakuma et al., 2019),而对于机器人的接受程度本身也会影响到对机器人的利他行为(Nielsen et al., 2022)。然而,也有一些研究认为,文化差异在人际交互过程中并不重要,有综述回顾了 2015 年至 2020 年的相关文献,对比了西方被试和日本被试对机器人的态度和行为,结果发现文化间的共性多于差异(Smith et al., 2021)。因此,不同文化如何塑造了人们对机器人的态度和交互模式,仍然存在一

定的争议空间。为了进一步明确文化在人机交互中的作用,尤其是个体的母文化(home culture)背景是否影响对机器人的利他行为。本研究将收集西方文化被试的数据,与中国被试进行跨文化比较,以验证本研究所建立的“多元文化经历-心智知觉-对机器人的利他行为”这一中介模型具有文化普适性还是文化差异性(见图 1)。

2 研究 1: 多元文化经历提高对机器人的利他行为: 心智知觉的中介作用

研究 1 旨在通过问卷法初步验证多元文化经历和对机器人利他行为间的正向关系,探索心智知觉在其中的中介作用,并在中国和西方文化中对本研究的假设进行检验,以探究多元文化经历-心智知觉-对机器人的利他行为”这一模型是否存在跨文化的普适性。

2.1 研究 1a: 中国样本

研究 1a 的目的是初步探究在中国文化情境下,多元文化经历更丰富的个体是否对机器人有更高的心智知觉,从而更愿意对机器人做出利他行为。研究 1a 采用问卷法,测量被试的多元文化经历、对机器人的利他行为、心智知觉和人口学变量,初步构建心智知觉在多元文化经历和对机器人的利他行为之间的中介模型。

2.1.1 方法

(1) 研究对象

参考 Lu 等(2017)人的研究,本研究在网络平台上招募海外生活经历 6 个月以上的中国籍留学生被试,利用在线平台“问卷星”发放问卷。在问卷中设置“此题请选择 x 选项”的注意力检测题,共回收有效数据 266 份,剔除掉极端值、作答时间过短、回答过于一致、国外生活经历小于 6 个月的问卷后,纳入分析的有效数据 217 份,包括 103 名男性,114 名女性,平均年龄为 25.64 岁($SD = 3.91$),95.4%被试为大学本科及以上学历,以主观社会经济地位量表(Adler et al., 2020)测得的被试的主观社会经济地位平均层数为 6.24 ($SD = 1.61$)。

(2) 研究程序

在被告知此次问卷调查的目的以及数据信息的安全保障后,被试填写知情同意书,然后被试填写多元文化经历量表、心智知觉量表、对机器人的利他行为,呈现顺序随机,最后被试报告了性别、年龄、受教育程度、主观社会经济地位等人口学信息。

(3) 测量工具

①多元文化经历

采用多元文化经历量表(Multicultural Experience Scale, MES; Aytug et al., 2018)的中文改编版, 量表结构包括多元文化暴露(如“观看不同文化的电影”)和多元文化互动(如“与来自不同文化的人聊天”)两个维度, 共 10 个项目, 采用 6 点计分(1 = “频率很低”, 6 = “频率很高”)。分数越高代表被试的多元文化经历越丰富。本研究中总量表的 Cronbach's α 系数为 0.86, 2 个因子的 Cronbach's α 系数分别为 0.79 和 0.88。

②心智知觉

采用心智归因量表中的心智知觉分量表(Perceived Mind Scale, PMS; Waytz et al., 2010; Shank & DeSanti, 2018)的中文版, 项目示例如“机器人有自己的思想”, 共 7 个项目, 量表采用 7 点计分(1 = “完全不同意”, 7 = “非常同意”)。分数越高表示被试认为机器人的心智程度越高。本研究中总量表的 Cronbach's α 系数为 0.93。

③对机器人的利他行为意愿

选择两个描述机器人需要帮助的图片情境材料(Thellman et al., 2017), 分别为: “家用机器人 Doppy 并不擅长做饭, 每次做饭 Doppy 都很苦恼”(见图 2); “Doppy 在做家务时不小心打碎了玻璃杯”(见图 3)。每个情境材料后要求被试选择其在多大程度上愿意帮助处于困境中的机器人, 采用 6 点计分(1 = “非常不愿意”, 6 = “非常愿意”)。分数越高表示被试对机器人的帮助意愿越强。两个情境材料的利他行为测量高度相关, $r(2) = 0.75, p < 0.001$ 。

④人口学变量

现有研究对个体的自身特征对机器人利他行为的研究尚不充分, 但已有研究指出个体特征可能会影响个体的道德边界, 从而影响个体是否对其赋



图 2 情景材料 1



图 3 情景材料 2

予道德地位, 进而影响是否对其做出利他行为。儿童(McGuire et al., 2022)和女性(Caviola et al., 2019)的道德圈相对较大, 而感知到的经济地位差距可能会缩小人们的道德圈(Kirkland et al., 2022)。因此本研究收集了被试的性别、年龄、受教育程度、主观社会经济地位等人口学信息作为控制变量。

2.1.2 结果

(1) 共同方法偏差检验

由于本研究数据采用问卷收集, 故采用探索性因素分析法对可能存在的共同方法偏差进行检验(周浩, 龙立荣, 2004)。采用 Harman 的单因子检验法, 析出的第一个公因子解释率为 31.99%, 小于 40%。因此, 本研究数据不存在严重的共同方法偏差。

(2) 描述性统计及相关矩阵

描述性统计分析结果如表 1 所示, 多元文化经历和主观社会经济地位($r = 0.20, p = 0.003$); 心智知觉($r = 0.42, p < 0.001$); 对机器人的利他行为($r = 0.35, p < 0.001$)显著正相关。心智知觉除了和多元文化经历与对机器人的利他行为显著正相关($r = 0.45, p < 0.001$), 同时和年龄($r = 0.24, p < 0.001$); 主观社会经济地位($r = 0.21, p = 0.002$)显著正相关。除此之外, 对机器人的利他行为和年龄($r = 0.22, p = 0.001$)显著正相关, 与最高学历显著负相关($r = -0.15, p = 0.024$)。

表 1 中国样本的多元文化经历、心智知觉、对机器人的利他行为和控制变量的描述统计与相关分析

变量	<i>M</i>	<i>SD</i>	Cronbach's α	1	2	3	4	5	6
1 性别	0.48	0.50	—	—					
2 年龄	25.64	3.91	—	0.03	—				
3 最高学历	5.36	0.66	—	0.04	0.04	—			
4 主观社会经济地位	6.24	1.61	—	-0.15*	0.08	-0.08	—		
5 多元文化经历	4.56	0.86	0.86	0.06	0.10	-0.09	0.20**	—	
6 心智知觉	3.86	1.68	0.93	-0.06	0.24***	-0.06	0.21**	0.42***	—
7 对机器人的利他行为	4.56	1.30	—	-0.02	0.22**	-0.15*	0.12	0.35***	0.45***

注: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; 性别: 男 = 1, 女 = 0。

表 2 中国样本的中介模型(N = 217)

变量	心智知觉		对机器人的利他行为	
	未纳入控制变量	纳入控制变量	未纳入控制变量	纳入控制变量
	β (SE)	β (SE)	β (SE)	B (SE)
控制变量:				
性别		-0.23 (0.21)		-0.05 (0.16)
年龄		0.08 (0.03) **		0.04 (0.02) *
最高学历		-0.07 (0.15)		-0.24 (0.12) *
主观社会经济地位		0.11 (0.07)		-0.01 (0.05)
预测变量:				
多元文化经历	0.82 (0.12) ***	0.74 (0.12) ***	0.30 (0.10) **	0.29 (0.10) **
中介变量:				
心智知觉			0.29 (0.05) ***	0.26 (0.05) ***
R^2	0.17	0.23	0.24	0.26
ΔR^2	0.17	0.21	0.23	0.25
F	45.29***	12.78***	32.96***	12.43***

注: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(3) 中介模型检验

使用 SPSS 26.0 的 PROCESS 插件(Hayes & Scharkow, 2013)中的模型 4, 建立心智知觉在多元文化经历与对机器人的利他行为之间的中介模型。以多元文化经历为自变量, 对机器人的利他行为作为因变量, 以机器人心智知觉为中介变量, 并纳入人口学变量作为控制变量(见表 2)。

未纳入控制变量的路径系数结果如图 4 所示。采用 Bootstrap 分析考察心智知觉的中介作用(Hayes & Scharkow, 2013), 样本量选择 10000, 95% 的 Bootstrap 置信区间为[0.14, 0.35], 不包括 0, 中介效应显著。直接效应与间接效应如表 3 所示。

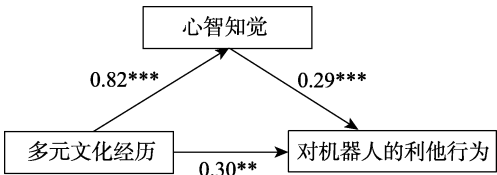


图 4 未纳入控制变量的中介模型
注: 报告的是非标准化系数; ** $p < 0.01$, *** $p < 0.001$ 。

纳入控制变量的路径系数结果如图 5 所示。采用 Bootstrap 分析考察心智知觉的中介作用(Hayes & Scharkow, 2013), 样本量选择 10000, 95% 的 Bootstrap 置信区间为[0.12, 0.34], 不包括 0, 中介效应显著。此外, 性别和主观社会经济地位的影响均不显著($ps > 0.770$), 而年龄与最高学历效应显著($ps < 0.048$), 具体体现为年龄越高, 对机器人的利他行为越多; 最高学历越高, 对机器人的利他行为

越少。直接效应与间接效应如表 3 所示。

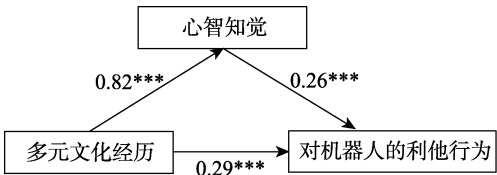


图 5 纳入控制变量的中介模型
注: 报告的是非标准化系数; ** $p < 0.01$, *** $p < 0.001$ 。

表 3 中国样本的对中介效应显著性检验的 bootstrap 分析及效果量

效应类型	效应值	效果量(%)	95%的置信区间	
			上限	下限
未纳入控制变量				
直接效应	0.30	57	0.10	0.50
间接效应	0.23	43	0.14	0.35
总效应	0.53			
纳入控制变量				
直接效应	0.29	58	0.004	0.09
间接效应	0.21	42	0.12	0.34
总效应	0.50			

2.1.3 讨论

研究 1a 的结果表明多元文化经历正向预测人们对机器人的利他行为意愿, 验证了本文的假设 1。研究发现, 心智知觉在其中起部分中介作用, 验证了本研究的假设 2。具体而言, 多元文化经历越丰富的个体, 对机器人心智能力感知越强, 从而有更高的对机器人的利他行为意愿。但研究 1a 仅使

用了中国文化样本,接下来的研究 1b 将在西方文化样本中检验这一模型是否具有跨文化普适性。

2.2 研究 1b: 西方样本

研究 1b 的目的是考察多元文化经历通过心智知觉对机器人利他行为的影响是否具有跨文化的普适性。研究 1b 招募西方被试,采用问卷法,探究西方文化情境下多元文化经历更丰富的个体是否会有更高的心智知觉,从而对机器人有更多的利他行为。

2.2.1 方法

(1) 研究对象

通过网络平台招募国外具有海外经历 6 个月及以上的被试,利用在线平台“Mturk”发放问卷。在问卷中设置“此题请选择 x 选项”的注意力检查题,共回收有效数据 354 份,剔除极端值、删去国外生活经历大于年龄的虚假作答、删去国外生活经历小于 26 周的问卷后纳入分析的有效问卷共 313 份,其中男生 192 人,女生 120 人。77% 的被试年龄在 25~40 岁之间,平均年龄为 33.94 岁($SD = 7.92$),84.1% 的被试最高学历为本科学历及以上,被试的主观社会经济地位平均为 5.13 层($SD = 0.95$)。

(2) 研究程序

研究 1b 为研究 1a 的英文版,研究材料与研究 1a 相同。被试在阅读完知情同意后填写了多元文化经历量表、心智知觉量表、对机器人的利他行为,三个量表的呈现顺序随机,最后被试报告了性别、年龄、受教育程度、主观社会经济地位等人口学信息。

(3) 测量工具

研究 1b 的问卷使用研究 1a 的英文版量表,收集了多元文化经历量表、心智知觉量表、对机器人的利他行为和其他人口学变量数据(性别、年龄、最高学历、主观社会经济地位)。多元文化经历量表和心智知觉量表的 Cronbach's α 系数分别为

0.88、0.90,两个情境材料显著正相关($r = 0.36, p < 0.001$)。

2.2.2 结果

(1) 共同方法偏差检验

由于本研究数据采用问卷收集,故采用探索性因素分析法对可能存在的共同方法偏差进行检验(周浩,龙立荣,2004)。使用 SPSS 26.0,采用 Harman 的单因子检验法,析出的第一个公因子解释率为 36.45%,小于 40%。因此,本研究数据不存在严重的共同方法偏差。

(2) 描述性统计及相关矩阵

对西方样本的数据进行描述性检验,结果如表 4 所示。

对西方样本的数据进行相关分析,和中国样本类似,多元文化经历与主观社会经济地位($r = 0.28, p < 0.001$)、心智知觉($r = 0.65, p < 0.001$)和对机器人的利他行为($r = 0.56, p < 0.001$)显著正相关,且相关系数更高。心智知觉除了与多元文化经历显著正相关,还和主观社会经济地位($r = 0.27, p < 0.001$)、对机器人的利他行为($r = 0.50, p < 0.001$)显著正相关,且相关系数更高,但西方样本的心智知觉与年龄不存在显著正相关。除此之外,西方样本中对机器人的利他行为和主观社会经济地位显著正相关($r = 0.46, p < 0.001$),但和年龄、最高学历无显著相关。

(3) 西方样本的心智知觉中介作用检验

使用 SPSS 26.0 的 PROCESS 插件(Hayes & Scharkow, 2013)中的模型 4,建立西方样本的心智知觉在多元文化经历与对机器人的利他行为之间的中介模型。以多元文化经历为自变量,对机器人的利他行为作为因变量,以机器人心智知觉为中介变量,以性别、年龄、最高学历、主观社会经济地位为控制变量(见表 5)。

表 4 西方样本的多元文化经历、心智知觉、对机器人的利他行为和控制变量相关分析

变量	<i>M</i>	<i>SD</i>	Cronbach's α	1	2	3	4	5	6
1 性别	0.62	0.49	—						
2 年龄	33.95	7.92	—	-0.15**					
3 最高学历	5.13	0.95	—	-0.08	0.06				
4 主观社会经济地位	7.48	1.86	—	-0.02	-0.05	-0.03			
5 多元文化经历	4.69	0.77	0.88	-0.01	0.06	0.10	0.28***		
6 心智知觉	5.21	1.12	0.90	-0.02	-0.01	0.05	0.27***	0.65***	
7 对机器人的利他行为	4.70	0.83	—	-0.02	-0.05	0.10	0.46***	0.56***	0.50***

注: 性别: 男 = 1, 女 = 0; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ 。

表 5 西方样本的中介模型(N = 313)

变量	心智知觉		对机器人的利他行为	
	未纳入控制变量	纳入控制变量	未纳入控制变量	纳入控制变量
	β (SE)	β (SE)	β (SE)	β (SE)
控制变量:				
性别		-0.04 (0.10)		0.04 (0.07)
年龄		-0.01 (0.01)		-0.01 (0.004)
最高学历		-0.01 (0.05)		-0.06 (0.04)
主观社会经济地位		0.05 (0.03)		-0.14 (0.02) ***
预测变量:				
多元文化经历	0.95 (0.06) ***	0.92 (0.07) ***	0.44 (0.06) ***	0.36 (0.06) **
中介变量:				
心智知觉			0.17 (0.04) ***	0.14 (0.04) ***
R^2	0.42	0.43	0.35	0.44
ΔR^2	0.42	0.42	0.34	0.43
F	228.78***	46.65***	82.26***	40.08***

注: ** $p < 0.01$, *** $p < 0.001$ 。

未纳入控制变量的路径系数结果如图 6 所示。采用 Bootstrap 分析考察心智知觉的中介作用(Hayes & Scharkow, 2013), 样本量选择 10000, 95% 的 Bootstrap 置信区间为[0.05, 0.29], 不包含 0, 中介效应显著, 直接效应与间接效应如表 6 所示。相对而言, 西方样本的中介模型的间接效应小于中国样本的间接效应。

纳入控制变量的路径系数结果如图 7 所示。采用 Bootstrap 分析考察心智知觉的中介作用(Hayes & Scharkow, 2013), 样本量选择 10000, 95% 的 Bootstrap 置信区间为[0.04, 0.25], 不包含 0, 中介效应显著。此外, 被试的性别、年龄最高学历的效应均不显著($ps > 0.09$), 主观社会经济地位效应显

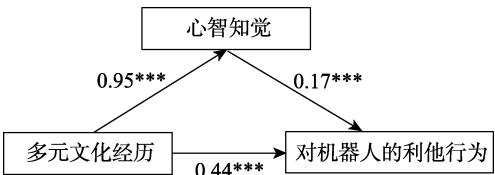


图 6 西方样本未纳入控制变量的中介模型
注: 报告的是非标准化系数; *** $p < 0.001$ 。

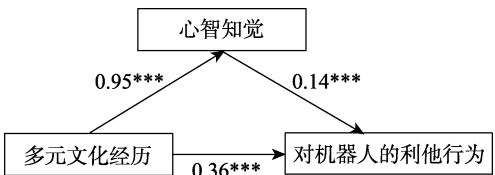


图 7 西方样本纳入控制变量的中介模型
注: 报告的是非标准化系数; *** $p < 0.001$ 。

表 6 西方样本的对中介效应显著性检验的 bootstrap 分析及效果量

效应类型	效应值	效果量(%)	95%的置信区间	
			上限	下限
未纳入控制变量				
直接效应	0.44	72	0.31	0.56
间接效应	0.17	28	0.05	0.29
总效应	0.61			
纳入控制变量				
直接效应	0.36	73	0.24	0.49
间接效应	0.13	27	0.04	0.25
总效应	0.49			

著($p < 0.001$), 具体而言体现为主观社会经济地位越高, 对机器人的利他行为越多。直接效应与间接效应如表 6 所示。相对而言, 西方样本的中介模型的间接效应小于中国样本的间接效应。

2.2.3 补充分析: 中西模型差异检验

为了进一步明确多元文化经历提升对机器人的利他行为及其心理机制是否在不同文化间存在差异, 综合研究 1a 和研究 1b 的中西文化样本数据, 使用 SPSS 26.0 的 PROCESS 插件(Hayes & Scharkow, 2013)中的 model 59 构建有调节的中介模型, 检验中西文化是否对中介模型存在调节作用, 结果表明不存在显著的调节作用(见表 7)。采用 Bootstrap 分析考察心智知觉的中介作用(Hayes & Scharkow, 2013), 样本量选择 10000, 95% 的 Bootstrap 置信区间在中国和西方条件下分别为

表 7 中西样本的有调节的中介模型($N = 530$)

变量	心智知觉		对机器人的利他行为	
	未纳入控制变量	纳入控制变量	未纳入控制变量	纳入控制变量
	β (SE)	β (SE)	β (SE)	β (SE)
控制变量:				
性别		-0.10 (0.10)		0.02 (0.08)
年龄		0.01 (0.01)		0.0003 (0.01)
最高学历		-0.03 (0.06)		-0.01 (0.05)
主观社会经济地位		0.08 (0.03)		0.09 (0.02) ***
预测变量:				
多元文化经历	0.82 (0.09) ***	0.78 (0.09) ***	0.30 (0.08) ***	0.28 (0.04) ***
中西文化	0.62 (0.60)	0.60 (0.60)	-0.33 (0.46)	-0.30 (0.46)
多元文化经历 \times 中西文化	0.13 (0.13)	0.11 (0.13)	0.13 (0.11)	0.12 (0.12)
心智知觉 \times 中西文化			-0.11 (0.08)	-0.12 (0.07)
中介变量:				
心智知觉			0.29 (0.04) ***	0.27 (0.04) ***
R^2	0.41	0.42	0.28	0.30
ΔR^2	0.41	0.41	0.27	0.28
F	122.48 ***	54.20 ***	40.66 ***	24.76 ***

注: *** $p < 0.001$ 。

[0.06, 0.30]和[0.14, 0.36], 不包含 0, 中介效应显著。但文化对三条路径的调节作用并不显著, $ps > 0.11$ (见图 8)。纳入控制变量后, 中介作用依然显著, 95%的 Bootstrap 置信区间在中国和西方条件下分别为[0.04, 0.27]和[0.13, 0.35], 不包含 0, 中介效应显著。文化对三条路径的调节作用依然不显著, $ps > 0.110$ (见图 9)。此外, 被试的性别、年龄和最高学

历的效应均不显著($ps > 0.751$), 主观经济地位效应显著, $p < 0.001$ 。因此在中西文化中, 多元文化经历提高心智知觉, 从而提高对机器人的利他行为这一效应具有一定的跨文化一致性。

2.2.4 讨论

研究 1b 旨在探究多元文化经历提升对机器人利他行为及心智知觉在其间的中介模型的跨文化普适性。结果发现多元文化经历的主效应显著, 即多元文化经历越丰富, 越能预测更高的对机器人的利他行为, 且心智知觉在其中的中介作用显著, 通过跨文化样本重复验证了研究 1a 的发现。同时, 结果表明未发现多元文化经历与中西文化背景存在交互作用。

总体而言, 研究 1 存在若干局限。首先, 本研究在测量对机器人利他行为中用到的图片和指导语(如: “每次做饭 Doppy 都很苦恼”)可能本身预设了机器人拥有心智, 这可能会对心智知觉和对机器人利他行为的测量产生干扰, 所以在接下来的研究中采用了自我报告的量表测量来弥补不足。另外, 研究 1 是从横断研究的层面反映了核心变量之间的相关模式, 而非因果推论。因此, 研究 2 将通过实验法操纵多元文化经历来探究因果机制。

3 研究 2: 多元文化经历提高对机器人利他行为的因果实验

研究 1 用问卷法初步探索了心智知觉在多元文化经历和机器人利他行为间的中介作用, 研究 2 采

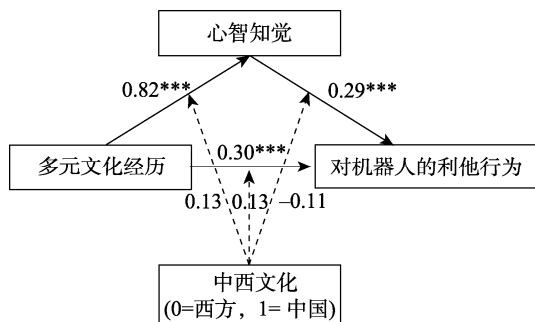


图 8 未纳入控制变量有调节的中介模型检验结果

注: 图中数字为非标准化系数; *** $p < 0.001$ 。

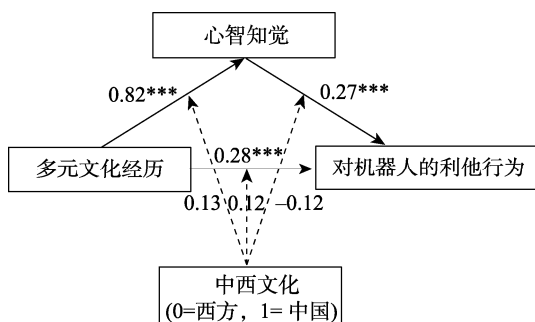


图 9 纳入控制变量有调节的中介模型检验结果

注: 图中数字为非标准化系数; *** $p < 0.001$ 。

用回忆范式(recall paradigm)进一步探索多元文化经历与心智知觉和机器人利他行为之间的因果关系。另外,研究1并未区分心智感知的具体维度,即能动性 & 体验性,无法更深入了解心智感知的分维度对主效应的不同解释机制。我们推测,在人类对机器人实施利他行为时,由于人类的角色是道德主体,而机器人的角色是道德客体,所以机器人体验性的增加会使人类对机器人有更多的利他行为。研究发现,当人们对机器人做出帮助行为时,会认为机器人有更多体验快乐的心智属性(Tanibe et al., 2017),即体验性的感知。所以,相对于能动性,体验性可能更加提升了人们对机器人的利他行为。

3.1 方法

3.1.1 研究对象

通过网络平台招募有国外生活经历的中国被试,利用在线平台“问卷星”进行线上回忆启动实验,通过注意力检查题的答卷共 323 份,剔除极端值、删除未曾去过外国国家的被试、主观题内容不符合指示语要求、回答无意义字符、直接复制网页内容等无效问卷后,最终纳入分析的有效问卷共 249 份,其中男性 162 人,女性 87 人,平均年龄为 25.96 岁($SD = 4.49$)。94.8%的被试最高学历是本科生,被试的主观社会经济地位为 6.41 层($SD = 1.33$)。被试被随机分配为三组,最后多元文化经历组被试 87 个、家乡文化经历组 83 个、日常生活对照组 79 个。

3.1.2 实验程序

我们告知被试此次问卷调查的目的以及数据信息的安全保障后,请被试填写知情同意书。在阅读指导语后将被试随机分配到 3 个实验条件组:多元文化经历启动组、家乡经历启动组和日常生活组。当被试阅读完启动材料后,要求被试写下 100 字描述其回忆到的国外生活经历/家乡生活经历和日常生活。随后被试填写操纵检验量表、心智知觉量表以及对机器人的利他行为问卷,三个测量的顺序随机呈现。最后,被试报告了性别、年龄、受教育程度和主观社会经济地位等人口学信息。

3.1.3 多元文化经历启动

根据以往研究(Lu et al., 2017; Maddux & Galinsky, 2009),本实验共设置 3 种条件,多元文化经历组、家乡文化经历组和日常生活对照组,后两组均为对照组。家乡文化经历作为多元文化经历的文化经历控制组,日常生活对照组是不涉及文化经历的控制组。采用回忆范式,要求多元文化经历组回忆“您曾去过的一个外国国家,详细描述一下您

在国外生活的那段时间都有哪些生活经历。比如,在那个国家经常发生什么事,您遇到了哪些人,您在国外生活是什么感受,您看到了什么、做了什么,想到了什么”,并写下不少于 100 字的描述。家乡文化经历组的指导语是“请您回忆您的家乡,详细描述一下您在家乡生活的那段时间都有哪些生活经历。比如,在家乡经常发生什么事,您遇到了哪些人,您在家乡生活是什么感受,您看到了什么、做了什么,想到了什么”,日常生活对照组的指导语是“请您回忆并详细描述下您上一次去超市的经历”,两组均被要求写下不少于 100 字的描述。

3.1.4 测量工具

(1) 操纵检验

设置 3 道操纵检验题目,分别是“我和国外文化的人接触很多”“我和国外文化的事物接触很多(比如电影、音乐、书籍等)”“我对探索国外的文化风格很有兴趣”。3 道题的 Cronbach's α 系数为 0.60。

(2) 心智知觉

采用 Yam 等人(2021)编制的心智知觉量表,分为主观性和经验性两个维度。主观性包括“机器人会思考”等共 4 个项目,经验性包括“机器人能感觉到疼痛”共 4 个项目,均采用 6 点计分(1 = “非常不同意”,6 = “非常同意”)。分数越高表示被试认为机器人的心智程度越高。本研究中总量表的 Cronbach's α 系数为 0.88,两个分量表的 Cronbach's α 系数分别为 0.77 和 0.88。

(3) 对机器人的利他行为

考虑到研究1中对机器人的利他行为的材料可能在一定程度上人格化机器人而对被试的利他行为产生干扰,因此研究2采用自编的对机器人的利他量表,共 5 个项目,分别为“我常会为机器人做一些好事”“当机器人遇到困难时,我会去帮助机器人”“我有时会顾及机器人的感受,为机器人着想”“我愿意做出行动来保障机器人的权益”“我觉得如果能加入机器人保护组织当一名志愿者是不错的选择”,采用 6 点计分(1 = “完全不符合”,6 = “非常符合”)。量表的 Cronbach's α 系数为 0.83。

(4) 人口统计学变量

收集了性别、年龄、最高学历、主观社会经济地位等人口学变量。

3.2 结果

3.2.1 操纵检验

以组别为自变量,操纵检验为因变量进行单因素方差分析。结果表明,三组之间存在显著差异,

$F(2, 246) = 3.65, p = 0.032, \eta_p^2 = 0.29$ 。事后 Bonferroni 检验表明, 多元文化启动组($M = 5.16, SD = 0.63$)的多元文化经历得分显著高于日常文化控制组($M = 4.89, SD = 0.71, p = 0.031$), 而和家乡文化启动组($M = 4.97, SD = 0.66, p = 0.183$)不存在显著差异, 说明多元文化经历启动成功。

3.2.2 对机器人的利他行为

以组别为自变量, 对机器人的利他量表得分为因变量的单因素方差分析结果不显著, $F(2, 246) = 2.18, p = 0.120$ 。多元文化启动组($M = 4.85, SD = 0.73$)的利他行为与日常生活控制组($M = 4.59, SD = 0.96$)和家乡文化启动组($M = 4.67, SD = 0.84$)间的差异不显著, 事后的 Bonferroni 检验表明均未达到显著水平, $ps > 0.130$ 。

3.2.3 心智知觉

以组别为自变量, 心智知觉量表总得分为因变量的单因素方差分析结果未显著, $F(2, 246) = 1.26, p = 0.26$ 。事后检验表明家乡文化启动组($M = 3.73, SD = 1.13$)的心智知觉与多元文化启动组($M = 3.63, SD = 1.23$)和日常生活控制组($M = 3.44, SD = 1.07$)间的差异不显著, 均未达到显著水平, $ps > 0.954$ 。

和预期假设不一致的是, 以组别为自变量, 能动性维度得分为因变量的单因素方差分析结果边缘显著, $F(2, 246) = 2.32, p = 0.10, \eta_p^2 = 0.02$ 。家乡经历启动组($M = 4.37, SD = 1.12$)的心智知觉与日常生活启动组($M = 4.01, SD = 0.99$)和多元文化控制组($M = 4.31, SD = 1.25$)间的差异不显著, 事后检验表明不存在显著差异, $ps > 0.130$ 。而以体验性维度得分为因变量的单因素方差分析结果不显著, $F(2, 246) = 0.48, p = 0.622$ 。

3.2.4 中介模型

使用 SPSS 26.0 的 PROCESS 插件(Hayes & Scharkow, 2013)中的模型 4, 样本量选择 10000。由于心智知觉中仅能动性维度边缘显著, 因此建立心智知觉的能动性在多元文化经历与对机器人的利他行为之间的中介模型。以多元文化经历为自变量, 对机器人的利他行为作为因变量, 以能动性为中介变量, 建立中介模型。结果表明直接效应不显著($p = 0.143$), 中介效应也不显著, 95%的 Bootstrap 置信区间分别为 $[-0.10, 0.01]$, 包括 0。

3.3 讨论

本研究对多元文化经历进行实验操纵, 旨在验证多元文化经历与机器人利他行为的因果关系, 并探索心智知觉的中介作用。结果表明多元文化经历

的实验操纵成功, 多元文化经历对机器人利他行为的主效应未达到显著, 未能验证本文的假设 1。但和预期假设不一致的是, 本研究发现多元文化经历对心智知觉能动性的影响边缘显著, 这在一定程度上表明相对于体验性, 多元文化经历丰富的人们更可能感知到机器人拥有主体能动性。由于本研究是线上启动实验, 存在作答环境等无法控制的额外变量的干扰, 可能在一定程度上导致结果未达到显著。但在一定程度上为多元文化经历提高对机器人的利他行为提供了实证支持证据, 并表明能动性和体验性可能存在不同的中介作用。

4 总讨论

4.1 理论贡献与实践启示

研究 1 通过线上问卷的方式, 在中西方文化下收集数据, 初步验证了多元文化经历对机器人利他行为正向预测作用, 并发现心智知觉在二者中间起部分中介作用, 这一结论具有跨文化的普适性。具体而言, 多元文化经历程度越高的个体, 越有可能认为机器人有心智能力, 从而提升对机器人的利他行为意愿, 初步验证了假设 1 和假设 2。研究 2 对多元文化经历进行实验操纵, 尽管结果未达到显著, 但多元文化经历组的利他行为水平有高于控制组的趋势, 也在一定程度上表明多元文化经历可能提高人们对机器人的利他行为和对机器人能动性的感知。

本研究对社会心理学视角的人工智能研究主要有如下几点贡献。首先, 本文首次探讨了多元文化经历对机器人利他行为的提升作用, 并通过中西文化样本证实了我们的核心假设。以往研究大多从机器人自身的物理和行为特征来探讨人类对机器人的利他行为(如 Somanader et al., 2011), 鲜有研究从人类的自身特征或个体差异来探讨。本研究则选取了一个能够促进个体的开放性、认知灵活性、对新颖事物接受程度的重要变量——多元文化经历, 探究其对机器人利他行为的影响。这在一定程度上拓展了多元文化经历的社会心理后果的研究证据, 并将其引入到人工智能心理学的领域, 丰富了多元文化经历的理论内涵, 发现了多元文化经历可能带来的“下游效应(downstream effects)”, 并为何种因素会提升人类对机器人利他行为的研究做出原创贡献。

其次, 本研究从心智感知理论的视角, 发现多元文化经历通过增加对机器人的心智感知, 从而增

加对机器人的利他行为,揭示了多元文化经历如何影响机器人利他行为的心理机制。以往研究发现了心智知觉影响对机器人利他行为(Tanibe et al., 2017),但未探究何种文化因素会影响人们对机器人的心智知觉。多元文化经历这一变量的引入在一定程度上填补了这一空白,丰富了心智感知理论领域相关的研究。

最后,本研究探索了心智知觉在多元文化经历和对机器人利他行为行为间中介模型的跨文化普适性。以往有诸多研究发现对机器人的接受程度存在一定的文化差异(Bartneck et al., 2005; Dang & Liu, 2021; 许丽颖, 喻丰, 2020),虽然本文并未发现多元文化经历与中西文化背景的交互作用,但发现了对机器人利他行为的中西文化差异:西方文化下的被试对机器人的利他行为显著高于中国文化下的被试。这可能是因为西方对机器人的接受度要比中国人更高(Sakuma et al., 2019),从而导致对机器人的利他行为更高。

4.2 研究局限与未来展望

本研究最主要的局限在于只在相关模式的层面发现了多元文化经历对机器人利他行为的提升,因果层面的验证由于线上实验干扰因素多的局限仅得到了边缘显著的结果。未来研究可以进一步采用更为有效的多元文化经历启动范式,考察对心智知觉不同维度的具体影响、改进对机器人的利他行为的测量方法,并进行多次严格的重复实验,通过元分析的技术验证该效应是否稳健。此外,在研究1b的设计中,没有考虑中西方文化被试在利他行为意愿上可能本来就存在差异,比如西方人可能更愿意帮助陌生人,中国人可能更愿意帮助亲人朋友,这可能也会影响对机器人的利他行为,这在之后的研究中应予以考虑。除此之外,本研究是从多元文化经历的积极后果探讨其对机器人利他行为的影响,但多元文化经历也可能带来负面的后果(胡晓檬等, 2021),比如多元文化适应不佳的个体,可能更觉得机器人是个冲击和威胁,挑战了人类的权威和道德地位,将其视为外群体并对其持排斥态度等。因此未来研究可以继续探讨什么条件下多元文化经历会使人们对机器人产生积极态度,什么条件下产生消极态度,或者持有矛盾态度。另外,本研究虽然排除了一些控制变量对实验结果的影响,但还存在一些其他可能的额外变量,如友善与共情能力,可能对实验的主效应起到一定的干扰作用。未来研究应探讨除心智知觉外其他可能的中介变量,

进一步拓展解释机制。

5 结论

综合2个研究的结果,初步证明个体的多元文化经历对机器人利他行为具有正向预测作用,且心智知觉在二者中间起部分中介作用。线上启动实验研究虽未能证实因果机制,但存在符合假设的趋势。另外,中介模型具有一定的跨文化一致性。以上研究证据表明,增加个体的多元文化暴露或多元文化互动,有助于提高人们对机器人的心智知觉,从而进一步提高对机器人的利他行为。

参 考 文 献

- Adair, W. L., & Xiong, T. X. (2018). How Chinese and Caucasian Canadians conceptualize creativity: The mediating role of uncertainty avoidance. *Journal of Cross-Cultural Psychology*, 49(2), 223–238.
- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology*, 19(6), 586–592.
- Anthis, J. R., & Paez, E. (2021). Moral circle expansion: A promising strategy to impact the far future. *Futures: The Journal of Policy, Planning and Futures Studies*, 130, 102756.
- Aytug, Z. G., Kern, M. C., & Dilchert, S. (2018). Multicultural experience: Development and validation of a multidimensional scale. *International Journal of Intercultural Relations*, 65, 1–16.
- Bartneck, C. (2008). Who like androids more: Japanese or US Americans? In M. Buss & K. Kühnlenz (Chair), *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN* (pp. 553–557). Munich, Germany.
- Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2005). Cultural differences in attitudes towards robots. In K. Dautenha (Chair), *Proceedings of the AISB Symposium on Robot Companions: Hard Problems and Open Challenges In Human-Robot Interaction* (pp. 1–4). Hartfield, UK.
- Batson, C. D. (2011). *Altruism in humans*. Oxford University Press, New York.
- Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, 181, 21–34.
- Cao, J., Galinsky, A. D., & Maddux, W. W. (2013). Does travel broaden the mind? breadth of foreign experiences increases generalized trust. *Social Psychological & Personality Science*, 5(5), 515–527.
- Caviola, L., Everett, J. A. C., & Faber, N. S. (2019). The moral standing of animals: Towards a psychology of speciesism. *Journal of Personality and Social Psychology*, 116(6), 10111029.
- Crimston, C. R., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology*, 111(4), 636–653.
- Danaher, J. (2019). Welcoming robots into the moral circle: A defence of ethical behaviourism. *Science and Engineering*

- Ethics*, 26(4), 2023–2049.
- Dang, J., & Liu, L. (2021). Robots are friends as well as foes: Ambivalent attitudes toward mindful and mindless AI robots in the United States and China. *Computers in Human Behavior*, 115, 106612.
- de Graaf, M. M. A., & Ben Allouch, S. (2016). Anticipating Our Future Robot Society: The Evaluation of Future Robot Applications From A User's Perspective. In J. Kaye & A. Druin (Chair), *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 755–762). New York, USA.
- Ding, X. J., Yu, F., & Xu, L. Y. (2020). The origin of and solution to people's aversion to AI making moral decisions. *Journal of Dialectics of Nature*, 42(12), 7–14.
- [丁晓军, 喻丰, 许丽颖. (2020). 人们对人工智能做道德决策的厌恶感之源及解决之道. *自然辩证法通讯*, 42(12), 7–14.]
- Fabes, R. A., Carlo, G., Kupanoff, K., & Laible, D. (1999). Early adolescence and prosocial/moral behavior I: The role of individual processes. *Journal of Early Adolescence*, 19(1), 5–16.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315(5812), 619.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101–124.
- Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter?. *Psychological Science*, 24(10), 1918–1927.
- Hu, X., Zhu, Y., Yu, F., Wilder, D. A., Zhang, L., Chen, S. X., & Peng, K. (2020). A cross - cultural examination on global orientations and moral foundations. *PsyCh Journal*, 9(1), 108–117.
- Hu, X. M., Han, Y. F., Yu, F., & Peng, K. P. (2021). The double-edged sword effect of multicultural experiences: Psychological consequences and boundary conditions. *Chinese Journal of Applied Psychology*, 27(1), 1–10.
- [胡晓檬, 韩雨芳, 喻丰, 彭凯平. (2021). 多元文化经历的双刃剑效应: 心理后果与边界条件. *应用心理学*, 27(1), 1–10.]
- Kaplan, F. (2004). Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. *International Journal of Humanoid Robotics*, 1(03), 465–480.
- Kirkland, K., Crimston, C. R., Jetten, J., Rudnev, M., Acevedo-Triana, C., Amiot, C. E., ... Bastian, B. (2022). Moral expansiveness around the world: The role of societal factors across 36 countries. *Social Psychological and Personality Science*, 14(3), 305–318.
- Klincewicz, M. (2016). Artificial intelligence as a means to moral enhancement. *Studies in Logic, Grammar and Rhetoric*, 48(1), 171–187.
- Kühnlenz, B., Sosnowski, S., Buß, M., Wollherr, D., Kühnlenz, K., & Buss, M. (2013). Increasing helpfulness towards a robot by emotional adaption to the user. *International Journal of Social Robotics*, 5(4), 457–476.
- Lee, S. A., & Liang, Y. (Jake). (2016). The role of reciprocity in verbally persuasive robots. *Cyberpsychology, Behavior, and Social Networking*, 19(8), 524–527.
- Leung, A. K. Y., Maddux, W. W., Galinsky, A. D., & Chiu, C. Y. (2008). Multicultural experience enhances creativity. *American Psychologist*, 63(3), 169–181.
- Lu, J. G., Quoidbach, J., Gino, F., Chakroff, A., Maddux, W. W., & Galinsky, A. D. (2017). The dark side of going abroad: How broad foreign experiences increase immoral behavior. *Journal of personality and social psychology*, 112(1), 1–16.
- Lund, V., Mejdell, C., Röcklinsberg, H., Anthony, R., & Håstein, T. (2007). Expanding the moral circle: Farmed fish as objects of moral concern. *Diseases of Aquatic Organisms*, 75(2), 109–118.
- Maddux, W. W., & Galinsky, A. D. (2009). Cultural borders and mental barriers: The relationship between living abroad and creativity. *Journal of Personality and Social Psychology*, 96(5), 1047–1061.
- Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015, March). *Sacrifice one for the good of many?* Presented at HRI '15, Portland, OR.
- McGuire, L., Palmer, S. B., & Faber, N. S. (2022). The development of speciesism: Age-related differences in the moral view of animals. *Social Psychological and Personality Science*, 14(2), 228–237.
- Morris, M. W., Chiu, C.-y., & Liu, Z. (2015). Polycultural psychology. *Annual Review of Psychology*, 66, 631–659.
- Nam, T. (2019). Citizen attitudes about job replacement by robotic automation. *Futures*, 109, 39–49.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 58(1), 81–130.
- Nielsen, Y. A., Pfattheicher, S., & Keijsers, M. (2021). Prosocial behavior towards machines. *Current Opinion in Psychology*, 43, 260–265.
- Nielsen, Y. A., Thielmann, I., Zettler, I., & Pfattheicher, S. (2022). Sharing money with humans versus computers: On the role of honesty-humility and (non-)social preferences. *Social Psychological and Personality Science*, 13(6), 1058–1068.
- Nijssen, S. R. R., Müller, B. C. N., van Baaren, R. B., & Paulus, M. (2019). Saving the robot or the human? Robots Who feel deserve moral care. *Social Cognition*, 37(1), 41–S2.
- Osswald, S., Greitemeyer, T., Fischer, P., & Frey, D. (2010). Moral prototypes and moral behavior: Specific effects on emotional precursors of moral behavior and on moral behavior by the activation of moral prototypes. *European Journal of Social Psychology*, 40(6), 1078–1094.
- Park, S., & Yu, J. (2017). The effects of multicultural experience on empathy in adolescents: Focused on mediating effect of multicultural acceptance and cultural empathy. *Journal of Digital Convergence*, 15(4), 499–510.
- Pizarro, D. A., Detweiler-Bedell, B., & Bloom, P. (2006). The creativity of everyday moral reasoning: Empathy, disgust, and moral persuasion. In J. C. Kaufman & J. Baer (Eds.), *Creativity and reason in cognitive development* (pp. 81–98). Cambridge University Press.
- Sakuma, M., Kuramochi, K., Shimada, N., & Ito, R. (2019). Positive and Negative Opinions about Living with Robots in Japanese University Students. In *Hri '19: 2019 14th Acm/Ieee International Conference on Human-Robot Interaction* (pp. 640–641). Daegu, Korea (South).
- Shank, D. B., & Desanti, A. (2018). Attributions of morality and mind to artificial intelligence after real-world moral violations. *Computers in Human Behavior*, 86, 401–411.
- Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive robotics: The influence of robot gender on human behavior. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2563–2568). St. Louis, MO, USA.
- Smith, E. R., Šabanović, S., & Fraune, M. R. (2021). Human-robot interaction through the lens of social

- psychological theories of intergroup behavior. *Technology, Mind, and Behavior*, 1(2), 1–11.
- Somanader, M. C., Saylor, M. M., & Levin, D. T. (2011). Remote control and children's understanding of robots. *Journal of Experimental Child Psychology*, 109(2), 239–247.
- Srinivasan, V., & Takayama, L. (2016, March). *Help me please: Robot politeness strategies for soliciting help from humans*. Paper presented at Proceedings of the 2016 Conference on Human Factors in Computing Systems (CHI 2016), New York, NY.
- Tadmor, C. T., Hong, Y. Y., Chao, M. M., Wiruchnipawan, F., & Wang, W. (2012). Multicultural experiences reduce intergroup bias through epistemic unfreezing. *Journal of Personality and Social Psychology*, 103(5), 750–772.
- Tadmor, C. T., Tetlock, P. E., & Peng, K. (2006). Biculturalism and integrative complexity: Testing the acculturation complexity model. *Academy of Management Annual Meeting Proceedings*, 2006(1), C1–C5.
- Tanibe, T., Hashimoto, T., & Karasawa, K. (2017). We perceive a mind in a robot when we help it. *Plos One*, 12(7), e0180952.
- Thellman, S., Silvervarg, A., & Ziemke, T. (2017). Folk–Psychological interpretation of human vs. humanoid robot behavior: Exploring the intentional stance toward robots. *Frontiers in Psychology*, 8, 1962–1968.
- Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, 14(8), 383–388.
- Xu, L. Y., & Yu, F. (2020). Factors that influence robot acceptance. *Science Bulletin*, 65(6), 496–510.
- [许丽颖, 喻丰. (2020). 机器人接受度的影响因素. *科学通报*, 65(6), 496–510.]
- Yam, K. C., Bigman, Y. E., Tang, P. M., Ilies, R., de Cremer, D., Soh, H., & Gray, K. (2021). Robots at work: People prefer- and forgive-service robots with perceived feelings. *Journal of Applied Psychology*, 106(10), 1557–1572.
- Zhou, H., & Long, L. R. (2014). Statistical remedies for common method biases. *Advances in Psychological Science*, 12(6), 942–950.
- [周浩, 龙立荣. (2004). 共同方法偏差的统计检验与控制方法. *心理科学进展*, 12(6), 942–950.]

Multicultural experiences enhance human altruism toward robots and the mediating role of mind perception

TENG Yue^{1,2}, ZHANG Haotian³, ZHAO Siqu^{1,2}, PENG Kaiping⁴, HU Xiaomeng^{1,2}

(¹ Department of Psychology, Renmin University of China, Beijing 100872, China)

(² Major Innovation & Planning Interdisciplinary Platform for the “Double-First Class” Initiative, Renmin University of China, Beijing 100872, China)

(³ Department of Psychology, School of Philosophy, Wuhan University, Wuhan 430072, China)

(⁴ Department of Psychology, Tsinghua University, Beijing 100084, China)

Abstract

Artificial intelligence is developing rapidly, and the future of human beings is closely related to it. The question of how humans and robots can better work together has become a pressing concern for social psychologists. Human-robot interaction is a two-way process, and research has explored how robots can better serve humans. Whether humans share the same moral concern or even act altruistically toward robots is critical, as it will feed into technological advances and the stability of human society. Few studies have examined which cultural and psychological factors shape people's willingness and behavior to accord robots moral status, i.e., whether they are believed to deserve the same rights and benefits as humans and to perform more altruistic acts toward them. Through two sub-studies and a causal study, the present work seeks to explore whether individuals' multicultural experiences enhance altruistic behavior toward robots and whether human mental perceptions of robots play a mediating role.

Study 1a began with a cross-sectional study in which 217 valid participants with an average age of 25.64 years were collected in China through the Questionnaire Star platform to measure their multicultural experiences, altruistic behavior toward robots, mind perceptions, and demographic information. To verify the model's cross-cultural generalizability that multicultural experiences enhance altruistic behavior toward robots, Study 1b replicated the procedure of Study 1a on Mturk with 313 valid participants (mean age 33.94 years) using the English version of the questionnaire from Study 1a. Finally, to infer the causal relationship between multicultural experiences and altruistic behavior toward robots, Study 2 recruited Chinese participants with six months or more of overseas experiences to prime participants' multicultural experiences through reading and writing tasks. A total of 249 valid data were collected in Study 2, with a mean age of 25.96 years, and

participants were randomly divided into a multicultural experience priming group, hometown experience priming group, and control group. After priming, participants were asked to fill out the manipulation check scale, the Mind Perception Scale, and the Altruistic Behavior Toward Robots Questionnaire, with the order of the three measures presented randomly. Finally, participants reported their information on a number of demographic variables.

Study 1a found that individuals' multicultural experiences positively predicted altruistic behavior toward robots, with mind perceptions playing a partially mediating role. Study 1b found that this mediating chain was cross-culturally consistent across Chinese and Western participants, with no Chinese or Western cultural differences. We infer that this effect has some degree of cultural generalizability. Study 2 found that multicultural experiences were manipulated successfully, but the main effect of multicultural experiences on altruistic behavior toward robots did not reach significance.

The current work reveals that individuals' multicultural experiences increase altruistic behavior toward robots. Specifically, the richer an individual's multicultural experiences, the more likely they are to perceive the robots as possessing mental perception. Thus, they are more likely to trigger altruistic behavior toward robots. At the same time, this effect is to some extent cross-culturally generalizable. The results of our studies enrich the theoretical predictions of multicultural experiences, identify possible “downstream effects” of multicultural experiences, and make an original contribution to the study of what factors enhance human altruistic behavior toward robots.

Keywords multicultural experiences, human-computer interaction, altruistic behavior, mind perception, cultural universality