

• 人工智能心理与治理专刊 •

智能时代的人类优势与心理变革 (代序)*

吴胜涛¹ 彭凯平²

(¹ 吉林大学哲学社会学院, 长春 130012) (² 清华大学心理与认知科学系, 北京 100084)

摘要 随着智能时代的来临, 人、机边界变得模糊, 如何重新认识和发展人类的独特优势变得更加凸显和紧迫; 同时, 随着技术进步以及心理学理论、范式的不断转型, 一个涵盖人、动物和机器心智与行为的广义心理学正在形成。研究者从人工智能影响、新型人机关系、人工智能方法和学科交叉赋能等角度, 开展了一系列人工智能心理与治理研究。未来的心理研究者要立足人类社会和未来发展, 反思人工智能冲击下的人类地位与人性尊严, 特别是人类长期进化而来的审美感、创造力、同理心等独特优势以及不断延伸的人类本质和身份边界; 真正掌握和利用人工智能技术赋能心理学发展, 让人类意识黑箱和复杂社会行为的研究更加精确和高效, 并推动人工智能在心智计算、决策与干预等领域的应用; 尤其要思考心理学如何赋能人工智能, 发挥前者在研究人性、价值观、社会关系等领域的优势, 研究智能认知及其与人和动物的异同, 促进人工智能的合理使用与人机共生社会的良善治理。

关键词 人工智能心理, 人类优势, 智能认知, 个性化, 科技向善

分类号 B842

随着信息技术的飞速发展, 人工智能(Artificial Intelligence, AI)正以前所未有的速度学习语言、模拟决策、生成文本和图像, 被赋予情感、创造和自我进化的潜力。尤其随着大语言模型技术的不断突破, 人与机器的边界正在变得模糊, 这促使我们重新思考 AI 的本质和人类自身的主体地位(孙伟平, 2020)。一方面, 越来越多的人开始担忧, AI 可能将很快成为具备自我意识的主体, 并威胁到人类的独特性和高贵性(Kieslich, Lünich, & Marcinkowski, 2021), 甚至碳基生命将最终被硅基生命所取代, 即 AI 从人类的“代理者”转变为“替代者”(李河, 2020)。另一方面, 一些研究者开始反思, 过去习惯性地 AI 描述为“思维机器”或“类人智能”, 可能遮蔽了人们对 AI 本质和局限的认知(Mitchell, 2024)。Farrell 等人(2025)进一步指出, 当前的 AI 模型本质上是文化与社会技术, 它们所表现出的

“智能”并非自发生成, 而是受其训练语料中人类价值观、语言结构和社会偏见的影响。

同时, 心理学作为一门探索人类心智与行为的科学, 自 19 世纪末创立以来经历了多次重大理论和范式变革, 从精神分析、行为主义到人文主义, 到认知革命、情感革命及神经科学的兴起, 再到后来的计算心理学和人工智能心理学。这一历史进程伴随着人类对自身认识的逐渐深化, 也标志着心理学研究对象的不断扩展、乃至重心转移, 即从关注生理、行为表现转向探讨其心理、认知过程, 以及通过跨学科方法来研究人类心智与行为的神经、计算和遗传进化机制, 反映了心理学研究从内省报告、行为操纵到内隐联想、生态系统分析等的发展趋势(Hothersall & Lovett, 2022; Ke et al., 2025)。当前, 正值生成式 AI 等重大技术变革和第四次工业革命的关口, 人类社会和心理学发展均面临诸多机

收稿日期: 2025-09-15

* 清华大学全球产业研究院项目(2024-06-18-LXHT001)和启明星辰心理与认知发展基金(100020002)的支持。

通信作者: 彭凯平, E-mail: pengkp@mail.tsinghua.edu.cn

遇和挑战,因此心理研究者将不得不重新思考人类的优势、心理学的理论和范式变革等问题。

首先,智能时代如何重新定位人类自身的独特优势以及心理学的研究对象?心理学的研究对象是否仍然只是局限于人和动物的行为?自轴心时代以来,思想家们就开始将人类自身设定为“万物的尺度”,这一思想在启蒙运动时期被发展成“以人为本”或“人类中心主义”,人类独特的优势(如劳动、理性、情感、道德等)不断被发现和证明——这些优势以一个反思性的自我为基础,是人类区别于自然、人造物等客观物体的重要标志(Lindgren, 2024; Wang, 2023)。所以,心理学自诞生之日起就以研究人的心智为己任。然而,以内省和精神分析为主要方法的心理学在科学届备受质疑,为了研究对象可以被客观地观察、测量和改变,人和动物行为被同时纳入到心理学的研究范围。接着,随着大数据和社交媒体的普及,人类行为的数字印记也逐渐成为心理学不可或缺的研究对象(Kosinski et al., 2015)。近年来,随着生成式 AI 的问世,其远超人类的记忆容量、推理能力、反应速度以及人格、决策、担责等类人心智和行为不断被发现和验证,越来越多的心理学家主张,应该将 AI 生成的内容及其智能体的互动行为作为研究对象,其行为反应、价值偏差、集群特点及其与人类行为的异同、风险值得深入研究和严肃对待(Aher et al., 2023; Dillion et al., 2023; Lin, 2025; Lu, Song, & Zhang, 2025; Shanahan et al., 2023; Wang et al., 2025)。

其次,在智能时代,人与 AI 如何和谐共生、共同发展,以促进人类命运共同体的建设?AI 与传统技术存在显著差异,尤其随着机器人的智能化水平不断提升,人与机器人的关系从人与机器的工具关系逐渐转向人与智能的社会关系(许为 等, 2024)。与此同时,人机交互的复杂性和拟人性越来越引起研究者的注意。例如,伴随技术的进步,人们对于机器的信任逐渐接近人际信任(Lewis & Marsh, 2022)。人与智能系统关系联结和团队协作,智能体自身如何在智能社区或人-机混合社区里互动与演化,以及人类如何利用 AI 和应对 AI 数据隐私、科学创新、产业升级、网络安全、行为成瘾、智能鸿沟、可持续发展等全球性治理问题,也引发了一系列的社会心理学、工程心理学及相关哲学社会科学研究(Farrel, 2025; Li & Qi, 2025; Mei et al., 2024; Veale, Matus, & Gorwa, 2023; Zhang et al., 2023)。

最后,心理学与 AI 科学如何交叉赋能、共同发展?这不仅是心理学如何利用 AI 技术促进心理学研究方法、测量和干预技术的进步的问题(Ke et al., 2025; Xu et al., 2024),也是心理学如何启发 AI 技术研发和应用、并实现两个学科彼此融合和协同发展的问题(Gigerenzer, 2024; Gray et al., 2025)。一方面,心理学本身就是科技进步的产物,只有逻辑推理足够系统、实验方法足够精细和精巧,人类的心智和行为才能够被客观地观察、测量和操纵。毋庸置疑,包括 AI 在内的每一次科技进步(如电生理技术、脑成像、基因检测、计算机、互联网),都推动了心理学观测技术、数据采集和分析方法的重大突破,这也使得以往不可能的数据样本、分析方法、实验场景以及科学发现成为可能(Meng, 2024)。另一方面,AI 的发展也与心理学存在深厚的历史渊源,例如心理学的神经网络、语言加工和认知建模研究对 AI 发展、特别是大语言模型的诞生有重要启发(黄林洁琼 等, 2025);尤为重要的是,AI 不同于以往的机器或工具,它是具有海量数据、超级算力和类人属性的特殊主体,因此,有必要从心理学的角度去分析这一主体的心智特征(Gray et al., 2025),以及借鉴比较心理学的范式、基于新型人机关系的发现来指导我们对 AI 本质的理解以及 AI 产品的设计和应用(Dang & Liu, 2024; Voudouris, Cheke, & Schulz, 2025)。

为了全面而深入地探讨上述主题,并推动 AI 领域的心理学研究和跨学科融合,《心理学报》“人工智能心理与治理”专刊汇聚了 11 篇论文,涵盖人工智能影响、新型人机关系、人工智能方法和学科交叉赋能等多个领域,涉及问卷调查、行为实验、计算建模、智能模拟等多种方法,形成了一个丰富而多元的人工智能心理与治理研究领域。

首先,在人工智能影响领域,研究者从社会创造策略、主体责任感、消费者响应模式角度探讨了生成式 AI 对人类社会的重大影响,并反思了人类的独特优势。在本专刊中,耿晓伟等(2025)通过 4 项实验探讨了 AI 对人类冒险行为的助长效应,发现尤其在人机合作的成功反馈条件下,个体知觉到的主体责任更大,从而在风险决策中更冒险。周详等(2025)基于社会认同理论系统探讨了生成式 AI 对个体社会创造策略的促进作用,发现个体会采用社会创造策略、而非社会竞争策略来应对生成式 AI 的凸显,且个体的人类中心主义显著增强生成式 AI 凸显情境下社会创造策略的使用程度。李斌

等(2025)基于刻板印象内容模型,探讨了AI设计与消费者响应模式的产品匹配效应,发现消费者更偏爱AI设计的创新产品、而厌恶AI设计的怀旧产品。这三项研究共同表明,AI助长了人类的主体责任,并激发或凸显了人类个体在社会创造、情感记忆上的独特优势。

其次,在人机关系领域,研究者从信任、接受意愿角度探讨了智能时代的新型主客关系及其影响因素。在本专刊中,唐小飞等(2025)区分了人机关系中的情感修缮信任和功能辅助信任,发现前者通过提升组织温暖感知促进员工的工作贡献意愿,而功能辅助信任后者通过提升组织能力感知促进员工的工作贡献意愿。由姗姗等(2025)聚焦安全风险这一人机信任的新挑战,发现人们对矮小、摄像头不明显的静态机器人安全信任更高,对运动较慢、摄像头关闭的动态机器人安全信任更高。魏心妮等(2025)考察了人们对AI环保系统的使用意愿和态度,发现低可持续性降低了人们对AI的接受意愿,且道德而非能动是影响可持续性和接受意愿的中介机制。这三项研究从不同角度揭示了AI态度或新型人机关系背后的情感、功能、安全及环境可持续性问题,为智能产品的设计和推广提供了重要的科学依据。

进而,研究者采用多种技术路线探讨了大语言模型在情感能力测评及心理咨询中的应用。在本专刊中,周子森等(2025)基于人物对话的图像、文本识别与推理,发现多模态大语言模型(MLLMs)在双人互动中已经初步展现出和人类被试类似的情绪识别与情绪推理能力;同时,人物对话视觉特征一定程度上制约MLLMs基本情绪识别的表现,但能够有效促进复合情绪识别。邓丽芳等(2025)结合扎根理论研究和机器学习方法,构建了工作紧张人群阈下抑郁影响因素的机器学习模型,发现了工作紧张人群阈下抑郁表征的5个主范畴(如意志减弱)和影响因素的8个主范畴(如工作因素、评价适应和自主选择)。黄峰等(2025)基于零样本学习和思维链提示策略构建了一个自助式AI心理咨询机器人系统,发现使用该系统可以有效改善用户的抑郁、焦虑和孤独情绪,尤其是拟人化的AI咨询师在缓解孤独感方面表现出显著优势,而非拟人化设计在减轻压力方面效果更为明显。这些研究推动了AI在情感计算和心理咨询领域的应用,对于AI与心理学的交叉具有重要意义。

最后,研究者还探讨了心理学对AI研究的启

发,尤其把大语言模型或智能体作为研究对象来分析其心理特征,这也是人工智能心理学的核心使命,是智能时代心理学发挥作用的重要突破点。武月婷等(2025)基于社会认知的刻板印象内容模型,考察了人类对大模型的感知模式,发现大模型感知与人类感知均通过热情和能力两个维度,且人类对人类的感知具有热情优先效应,而人类对大模型的感知却存在能力优先效应。许丽颖等(2025)探讨了AI作为主管的道德问题,发现人们对AI主管(vs. 人类主管)提出道德行为建议的遵从程度更低,但个体拟人化倾向或AI主管的拟人化程度可以提高人们对AI主管提出的道德行为建议的遵从意愿。这两项研究的突出贡献在于,研究者聚焦AI自身的心理特征,发现其能力维度和拟人化特征是影响AI态度的重要因素。

总之,随着智能时代的来临,人机边界变得模糊,如何重新认识和发展人类的独特优势变得更加凸显和紧迫;同时,随着技术进步以及心理学理论、范式的不断转型,一个涵盖人、动物和机器心智与行为的广义心理学正在形成。伴随这一学科变革而产生的人工智能心理学不仅要研究智能时代的人类和人机关系,还要真正掌握和利用AI技术,改进传统的心理学研究方法;更重要的是,立足心理学来研究AI,推动科技向善,实现人机和谐共生以及幸福、美好的人类命运共同体。当然,目前关于AI的心理学研究还处在初级阶段,真正采用AI技术、以AI为研究对象、尤其是以心理学赋能AI的研究相对较少,对人与智能本质及二者共性、个性的理论思考还相对匮乏,甚至存在诸多误解(Lin, 2025; Song & Lin, 2025)。

值得注意的是,尽管智能时代的心理学面临重大机遇,但也存在可能会被计算机科学、尤其是人工智能学科所替代或至少沦为工具配角的风险。因此,未来的心理研究者可以从三个方面推动AI与心理学的学科融合与交叉赋能,促进人类自身发展和人-机-环境三者协同进化,并以此维护和增强心理学在智能时代的话语权。具体如下:

首先,立足人类社会和未来发展,反思AI冲击下的人类地位与人性尊严(Kissinger, Schmidt, & Mundie, 2024),研究智能时代人类面临的挑战和机遇,尤其是人类长期进化而来的审美感、创造力、同理心等独特优势(彭凯平、闫伟, 2021; Pink, 2006),以及随技术进步而不断延伸的人类本质和身份边界(Song & Lin, 2025)。

同时,真正掌握和利用 AI 技术赋能心理学发展,让人类意识黑箱和复杂社会行为的研究更加精确和高效,并推动 AI 在跨时间、空间尺度和个性化的心智计算、决策与干预,及其在心理咨询、教育变革、社会治理等领域的广泛应用(童松 等, 2025; Meng, 2024)。

更重要的是,要思考心理学如何赋能 AI 发展,发挥前者在研究人性、价值观、社会关系等领域的优势,深入探讨机器的认知、情绪、行为、人格、道德、文化特点及其与人和动物的异同,尤其是 AI 心智模拟、AI 交互系统和 AI 变态行为,并促进 AI 技术的合理使用与人机共生社会的良善治理(Batool et al., 2025; Gray et al., 2025; 晋少雄, 刘超, 2025; Song & Lin, 2025; Voudouris, Cheke, & Schulz, 2025)。

致谢: 感谢傅小兰、朱廷劭研究员及秦绍正、胡晓晴、喻丰教授对本专刊组织和编审工作的付出,以及林志成教授、朱廷劭研究员对本文部分核心观点的宝贵意见。

参 考 文 献

- Aher, G., Arriaga, R. I., & Kalai, A. T. (2023). Using large language models to simulate multiple humans and replicate human subject studies. *Proceedings of the 40th International Conference on Machine Learning*, 1–35.
- Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. *AI and Ethics*, 5, 3265–3279.
- Dang, J., & Liu, L. (2024). Extended artificial intelligence aversion: People deny humanness to artificial intelligence users. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspi0000480>
- Deng, L. F., Pei, B., & Gao, T. A. (2025). The factors affecting subthreshold depression for people with occupational stress in the era of digital intelligence: Machine learning-based evidence. *Acta Psychologica Sinica*, 57(11), 2001–2021.
- [邓丽芳, 裴蓓, 高天艾. (2025). 数智时代工作紧张人群阈下抑郁的影响因素: 基于机器学习的证据. *心理学报*, 57(11), 2001–2021.]
- Dillion, D., Tandon, N., Gu, Y., & Gray, K. (2023). Can AI language models replace human participants? *Trends in Cognitive Sciences*, 27(7), 597–600.
- Farrell, H. (2025). AI as Governance. *Annual Review of Political Science*, 28(1), 375–392.
- Farrell, H., Gopnik, A., Shalizi, C., & Evans, J. (2025). Large AI models are cultural and social technologies. *Science*, 387(6739), 1153–1156.
- Geng, X. W., Liu, C., Su, L., Han, B. X., Zhang, Q. M., & Wu, M. Z. (2025). Human-AI cooperation makes individuals more risk seeking: The mediating role of perceived agentic responsibility. *Acta Psychologica Sinica*, 57(11), 1885–1900.
- [耿晓伟, 刘超, 苏黎, 韩冰雪, 张巧明, 吴明证. (2025). 人机合作使人更冒险: 主体责任感的中介作用. *心理学报*, 57(11), 1885–1900.]
- Gigerenzer, G. (2024). Psychological AI: Designing algorithms informed by human psychology. *Perspectives on Psychological Science*, 19(5), 839–848.
- Gray, K., Yam, K. C., Zhen'An, A. E., Dillion, D., & Waytz, A. (2025). The psychology of robots and artificial intelligence. In D. T. Gilbert, S. T. Fiske, E. J. Finkel, & W. B. Mendes (Eds.), *The handbook of social psychology* (6th ed.). Situational Press, Cambridge, MA.
- Hothersall, D., & Lovett, B. J. (2022). *History of psychology*. Cambridge University Press.
- Huang, F., Ding, H. M., Li, S. J., Han, N., Di, Y. Z., Liu, X. Q., Zhao, N., Li, L. Y., & Zhu, T. S. (2025). Self-help AI psychological counseling system based on large language models and its effectiveness evaluation. *Acta Psychologica Sinica*, 57(11), 2022–2042.
- [黄峰, 丁慧敏, 李思嘉, 韩诺, 狄雅政, 刘晓倩, 赵楠, 李林妍, 朱廷劭. (2025). 基于大语言模型的自助式 AI 心理咨询系统及其效果评估. *心理学报*, 57(11), 2022–2042.]
- Huang, L., Zhang, W., Chen, Z., Li, C., & Li, X. (2025). The historical origins of large language models and psychology. *Journal of Psychological Science*, 48(4), 773–781.
- [黄林洁琼, 张雯, 陈珍, 李晨曦, 李兴珊. (2025). 大语言模型与心理学的历史渊源. *心理科学*, 48(4), 773–781.]
- Jin, S., & Liu, C. (2025). From human mind to artificial intelligence: Advancing AI value alignment through psychological theories. *Journal of Psychological Science*, 48(4), 782–791.
- [晋少雄, 刘超. (2025). 由“心”及“智”: 心理学研究促进 AI 价值观对齐的路径探讨. *心理科学*, 48(4), 782–791.]
- Kieslich, K., Lünich, M., & Marcinkowski, F. (2021). The threats of artificial intelligence scale (TAI) development, measurement and test over three application domains. *International Journal of Social Robotics*, 13, 1563–1577.
- Kissinger, H. A., Schmidt, E., & Mundie, C. (2024). *Genesis: Artificial Intelligence, hope, and the human spirit*. Little, Brown.
- Ke, L., Tong, S., Cheng, P., & Peng, K. (2025). Exploring the frontiers of LLMs in psychological applications: A comprehensive review. *Artificial Intelligence Review*, 58, 305–340. <https://doi.org/10.1007/s10462-025-11297-5>
- Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American psychologist*, 70(6), 543–556.
- Lewis, P. R., & Marsh, S. (2022). What is it like to trust a rock? A functionalist perspective on trust and trustworthiness in artificial intelligence. *Cognitive Systems Research*, 72, 33–49.
- Li, B., Rui, J. X., Yu, W. N., Li, A. M., & Ye, M. L. (2025). When design meets AI: The impact of AI design products on consumers' response patterns. *Acta Psychologica Sinica*, 57(11), 1914–1932.
- [李斌, 芮建禧, 俞炜楠, 李爱梅, 叶茂林. (2025). 当设计遇见 AI: 人工智能设计产品对消费者响应模式的影响. *心理学报*, 57(11), 1914–1932.]
- Li, C., & Qi, Y. (2025). Toward accurate psychological simulations: Investigating LLMs' responses to personality and cultural variables. *Computers in Human Behavior*, 170, 108687.
- Li, H. (2020). From technology as agent to technology as substitute: Human obsolescence? *Social Sciences in China*, (10), 116–140+207.

- [李河. (2020). 从“代理”到“替代”的技术与正在“过时”的人类? *中国社会科学*, (10), 116–140+207.]
- Lin, Z. (2025). Six fallacies in substituting large language models for human participants. *Advances in Methods and Practices in Psychological Science*, 8(3), 1–19.
- Lindgren, S. (2024). *Critical theory of AI*. Cambridge: Polity Press.
- Lu, J. G., Song, L. L., & Zhang, L. D. (2025). Cultural tendencies in generative AI. *Nature Human Behaviour*, <https://doi.org/10.1038/s41562-025-02242-1>
- Mei, Q., Xie, Y., Yuan, W., & Jackson, M. O. (2024). A Turing test of whether AI chatbots are behaviorally similar to humans. *Proceedings of the National Academy of Sciences*, 121(9), e2313925121.
- Meng, J. (2024). AI emerges as the frontier in behavioral science. *Proceedings of the National Academy of Sciences*, 121(10), e2401336121.
- Mitchell, M. (2024). The metaphors of artificial intelligence. *Science*, 386(6723), eadt6140.
- Peng, K., & Yan, W. (2021). *Children's character: Positive psychology for parents*. Beijing: Citic Press.
- [彭凯平, 闫伟. (2021). *孩子的品格: 写给父母的积极心理学*. 北京: 中信出版社.]
- Pink, D. H. (2006). *A whole new mind: Why right-brainers will rule the future*. Penguin.
- Shanahan, M., McDonnell, K., & Reynolds, L. (2023). Role play with large language models. *Nature*, 623(7987), 493–498.
- Song, X., & Lin, Z. (2025). *Beyond the existence–utility binary: How AI reveals our hybrid self*. AI & Society.
- Sun, W. (2020). Artificial intelligence and the “new alienation” of human beings. *Social Sciences in China*, 12, 119–137.
- [孙伟平. (2020). 人工智能与人的“新异化”. *中国社会科学*, 12, 119–137.]
- Tang, X. F., Wang, C. M., Sun, X. D., & Zhang, E. Z. (2025). Impact of trusting humanoid intelligent robots on employees' job dedication intentions: An investigation based on the classification of human–robot trust. *Acta Psychologica Sinica*, 57(11), 1933–1950.
- [唐小飞, 王昌梅, 孙晓东, 张恩忠. (2025). 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. *心理学报*, 57(11), 1933–1950.]
- Tong, S., Chen, H., Ke, L., Ye, J., & Peng K. (2025). Psychoinformatics: Advances and perspectives in the computational cognition era. *Journal of Psychological Science*, 48(4), 792–803.
- [童松, 陈浩, 柯罗马, 叶俊楷, 彭凯平. (2025). 心理信息学: 计算认知时代的研究进展. *心理科学*, 48(4), 792–803.]
- Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, 19(1), 255–275.
- Voudouris, K., Cheke, L., & Schulz, E. (2025). Bringing comparative cognition approaches to AI systems. *Nature Reviews Psychology*, 4, 363–364.
- Wang, J. (2023). Self-awareness, a singularity of AI. *Philosophy Study*, 13(2), 68–77.
- Wang, Y., Zhao, J., Ones, D. S., He, L., & Xu, X. (2025). Evaluating the ability of large language models to emulate personality. *Scientific Reports*, 15(1), 519.
- Wei, X. N., Yu, F., & Peng, K. P. (2025). Unsustainability decreases acceptance of environmental artificial intelligence. *Acta Psychologica Sinica*, 57(11), 1973–1987.
- [魏心妮, 喻丰, 彭凯平. (2025). 低可持续性降低人工智能的接受意愿. *心理学报*, 57(11), 1973–1987.]
- Wu, Y. T., Wang, B., Bao, H. W. S., Li, R. N., Wu, Y., Wang, J. Q., Cheng, C., & Yang, L. (2025). Humans perceive warmth and competence in large language models. *Acta Psychologica Sinica*, 57(11), 2043–2059.
- [武月婷, 王博, 包寒吴霜, 李若男, 吴怡, 王嘉琪, 程诚, 杨丽. (2025). 人类对大语言模型的热情和能力感知. *心理学报*, 57(11), 2043–2059.]
- Xu, L. Y., Zhao, Y. J., & Yu, F. (2025). Employees adhere less to advice on moral behavior from artificial intelligence supervisors than human. *Acta Psychologica Sinica*, 57(11), 2060–2082.
- [许丽颖, 赵一骏, 喻丰. (2025). 人工智能主管提出的道德行为建议更少被遵从. *心理学报*, 57(11), 2060–2082.]
- Xu, R., Sun, Y., Ren, M., Guo, S., Pan, R., Lin, H., Sun, L., & Han, X. (2024). AI for social science and social science of AI: A survey. *Information Processing and Management*, 61(3), 103665.
- Xu, W., Gao, Z. F., & Ge, L. Z. (2024). New research paradigms and agenda of human factors science in the intelligence era. *Acta Psychologica Sinica*, 56(3), 363–382.
- [许为, 高在峰, 葛列众. (2024). 智能时代人因科学研究的新范式取向及重点. *心理学报*, 56(3), 363–382.]
- You, S. S., Qi, Y., Chen, J. T., Luo, L., & Zhang, K. (2025). Safety trust in intelligent domestic robots: Human and AI perspectives on trust and relevant influencing factors. *Acta Psychologica Sinica*, 57(11), 1951–1972.
- [由姗姗, 齐玥, 陈俊廷, 骆磊, 张侃. (2025). 人与 AI 对智能家居机器人的安全信任及其影响因素. *心理学报*, 57(11), 1951–1972.]
- Zhang, G., Chong, L., Kotovsky, K., & Cagan, J. (2023). Trust in an AI versus a human teammate: The effects of teammate identity and performance on Human-AI cooperation. *Computers in Human Behavior*, 139, 107536.
- Zhou, X., Bai, B. R., Zhang, J. J., & Liu, S. R. (2025). Unity without uniformity: Humans' social creativity strategy under generative artificial intelligence salience. *Acta Psychologica Sinica*, 57(11), 1901–1913.
- [周详, 白博仁, 张婧婧, 刘善柔. (2025). 和而不同: 生成式人工智能凸显下人类的社会创造策略. *心理学报*, 57(11), .]
- Zhou, Z. S., Huang, Q., Tan, Z. H., Liu, R., Cao, Z. H., Mu, F. M., Fan, Y. C., & Qin, S. Z. (2025). Emotional capabilities evaluation of multimodal large language model in dynamic social interaction scenarios. *Acta Psychologica Sinica*, 57(11), 1988–2000.
- [周子森, 黄琪, 谭泽宏, 刘睿, 曹子亨, 母芳蔓, 樊亚春, 秦绍正. (2025). 多模态大语言模型动态社会互动情景下的情感能力测评. *心理学报*, 57(11), 1988–2000.]

Human advantages and psychological transformations in the era of artificial intelligence

WU Michael Shengtao¹, PENG Kaiping²

(¹ School of Philosophy and Sociology, Jilin University, Changchun 130012, China)

(² Department of Psychological and Cognitive Sciences, Tsinghua University, Beijing 10084, China)

Abstract

In the era of artificial intelligence (AI), the boundaries between humans and machines have become blurred, and re-understanding and developing humanity's unique advantages are increasingly prominent and urgent. Meanwhile, with the rapid development of technology and scientific paradigm, a broad psychology encompassing the minds and behaviors of humans, animals, and machines is emerging. Recent researchers have conducted a series of studies on the psychology and governance of AI, from the perspectives of impacts of AI, new human-machine relationships, AI methods, and interdisciplinary empowerment. Future psychology researchers should focus on human society and future development, and reflect on the status of humanity and human dignity under the impact of AI, especially the unique advantages derived from human evolution as well as the expansions of human nature and identity; truly master and utilize AI technologies to empower the development of psychology, making mind research on the black box of human consciousness and complex social behavior more precise and efficient, and promoting AI-based mind computation and intervention across time and space scales and personalized interventions. More important, they must consider how psychology (with strengths in studying human nature, social relations, and ethical values) could empower the development of AI, by exploring AI cognition and its comparison with humans and animals, which is critical for promoting the AI application and governance in a human-machine symbiotic society.

Keywords psychology of artificial intelligence, human advantage, intelligent cognition, personalization, technology for goodness.