人与AI的协作能在基于文本的点对点心理健康支持中实现更具同理心的对话

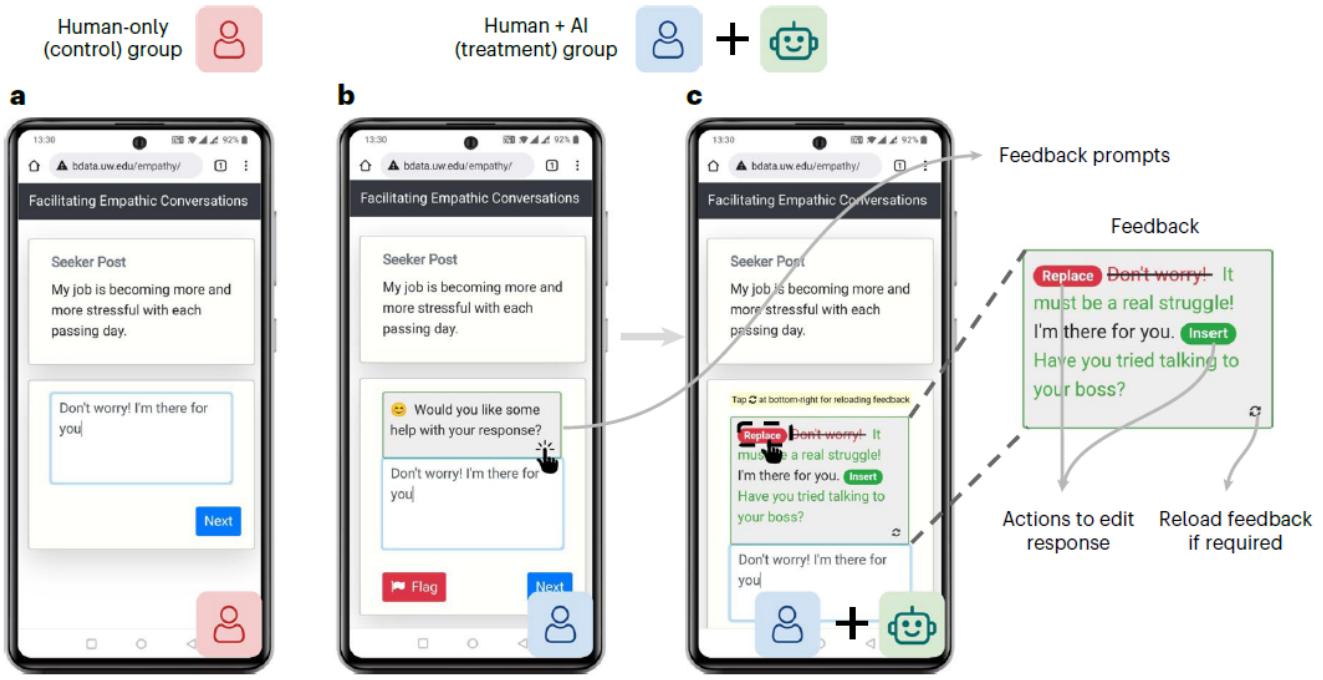
引言

随着人工智能(AI)技术的不断发展，AI系统已经开始在从电子商务到医疗保健的应用领域扩大和与人类合作。在许多情况下，尤其是在高风险的环境中，这种人类与人工智能的合作已被证明比用人工智能完全取代人类更加强大和有效。在这篇论文中，研究人员重点研究了基于文本的，对等的心理健康支持，并调查了人工智能系统如何与人类合作，以帮助在文本支持性对话中促进同理心的表达。

研究问题与假设

生活中需要心理健康服务的人很多，但是并不是每个人都能得到专业的服务，因此他们中的大多数人都会在论坛上寻求帮助，但论坛上提供帮助的人多数未经过专业的训练，缺乏具有同理心的表达。对此研究人员便开发了一种人工智能协作方法，在对话中提供即时的建议，以提供更具同理心的回复，但这种方法是否能真的提高回复的同理心水平呢?

实验方法

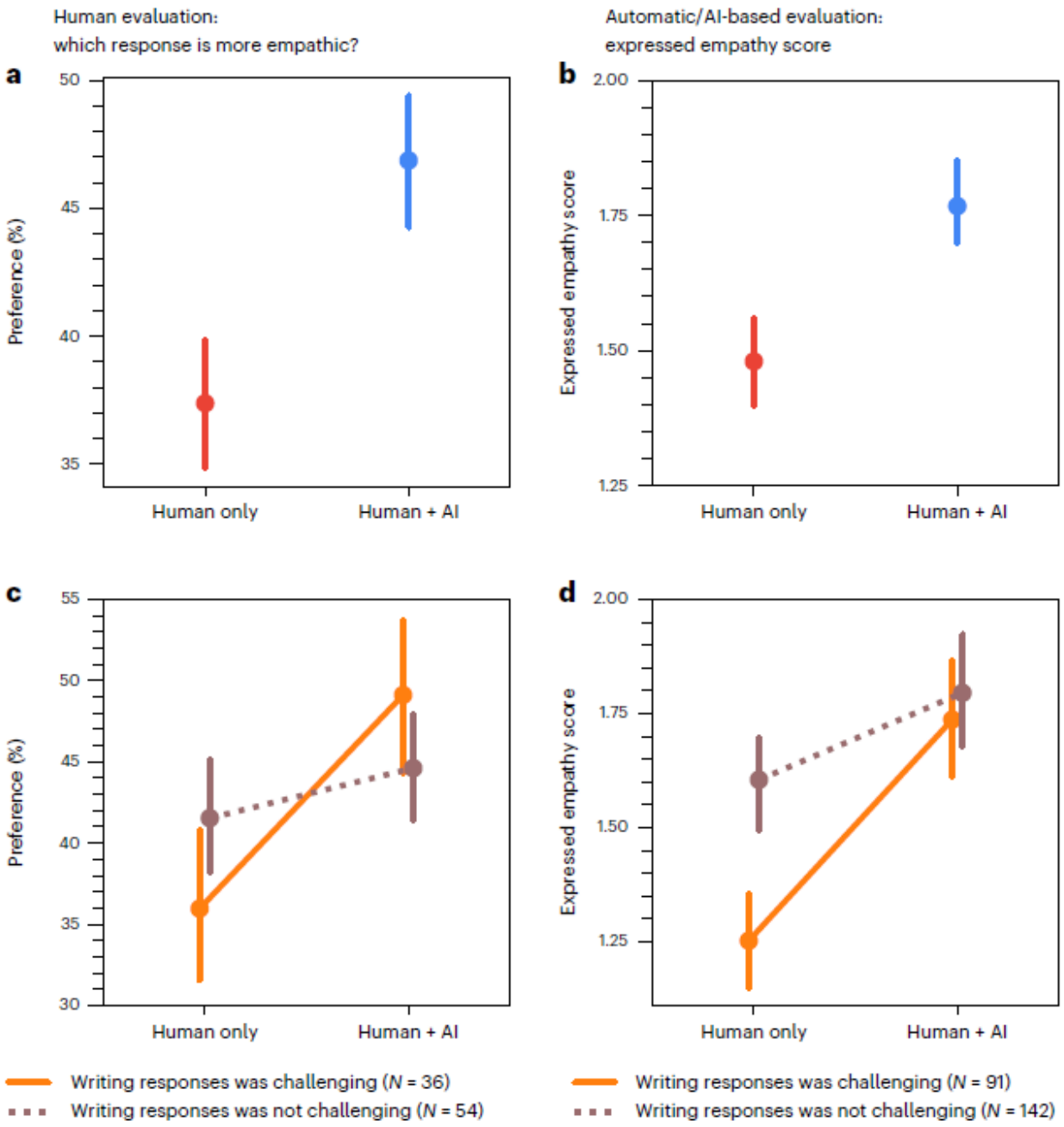


研究将参与者随机分为两组：仅人类（对照组）和人类+AI（治疗组），并要求他们分别在没有反馈和有AI反馈的情况下撰写帖子的支持性和有同理心的回复。

为了确定即时的人类-AI合作对同理心的提高，是否能够超越传统培训方法，两组参与者在研究开始前都接受了初步的同理心培训。在没有AI的情况下，人类回复者者面对一个空白的聊天框来撰写他们的回复，由于他们通常没有接受过关于同理心的培训，因此很少进行具有高度同理心的对话。而AI反馈助理则在同伴回复者撰写回应时提供即时的AI反馈。AI会提出可以对回复进行的修改建议，以使其更具同理心，这些建议包括可以插入的新句子、用更具同理心的对应句子替换当前句子。回复者可以通过点击“插入”和“替换”按钮接受这些建议，并继续编辑回复，或者根据需要获取更多反馈。

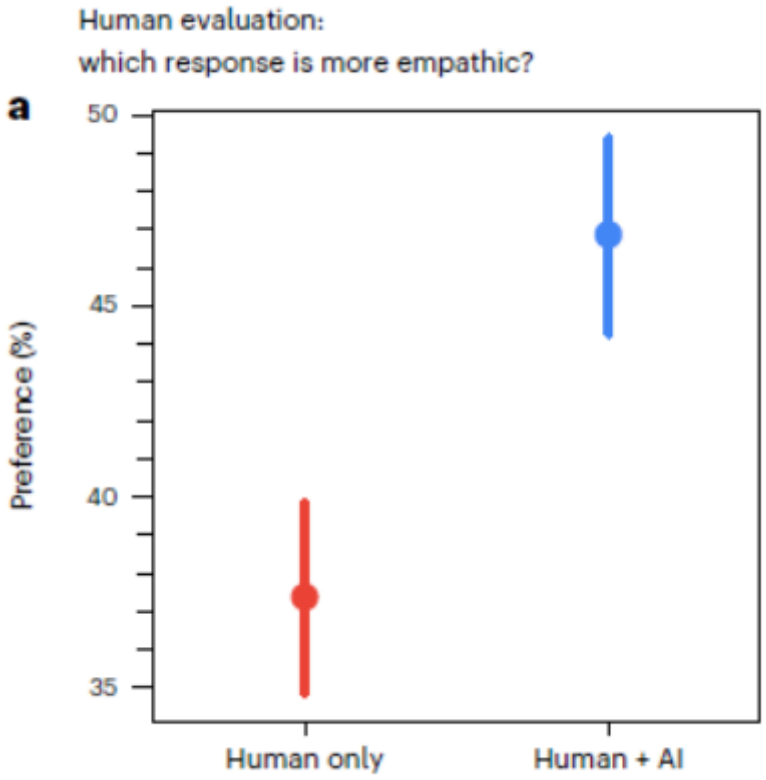
实验后，研究人员进行了事后评估，评估那些自己报告称在回复时遇到困难的参与者，是否从这个方法中获益更多，用以研究参与者与AI合作的差异，以及参与者对此方法的理解。

研究结果



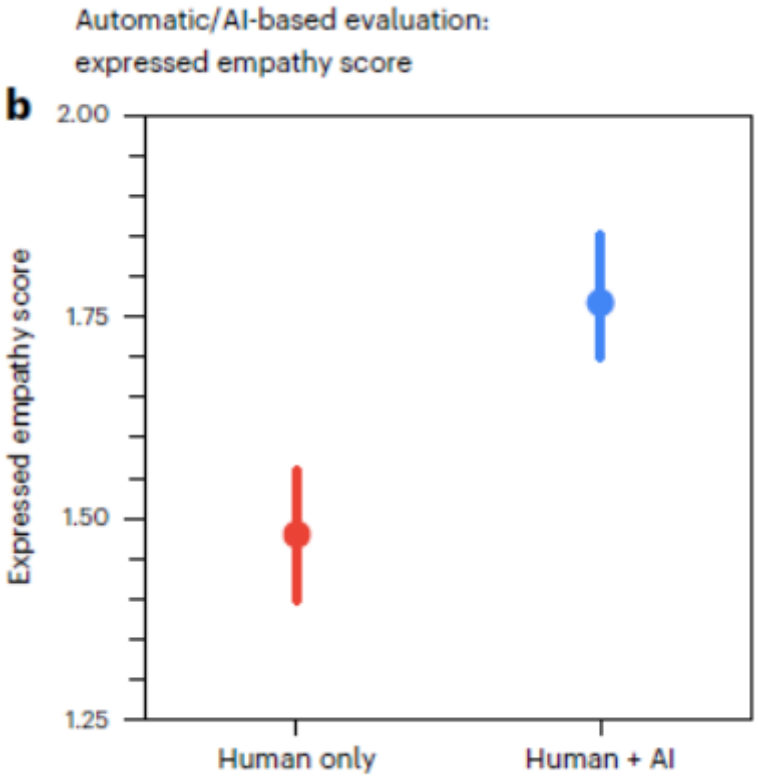
结果a,

随机对照实验证明人类与人工智能的合作能够产生更具同理心的对话。图a显示TalkLife(在线同伴支持社区)用户，对人类+AI的回复相对于仅人类回复的37.4%偏好具有更高的46.9%偏好。



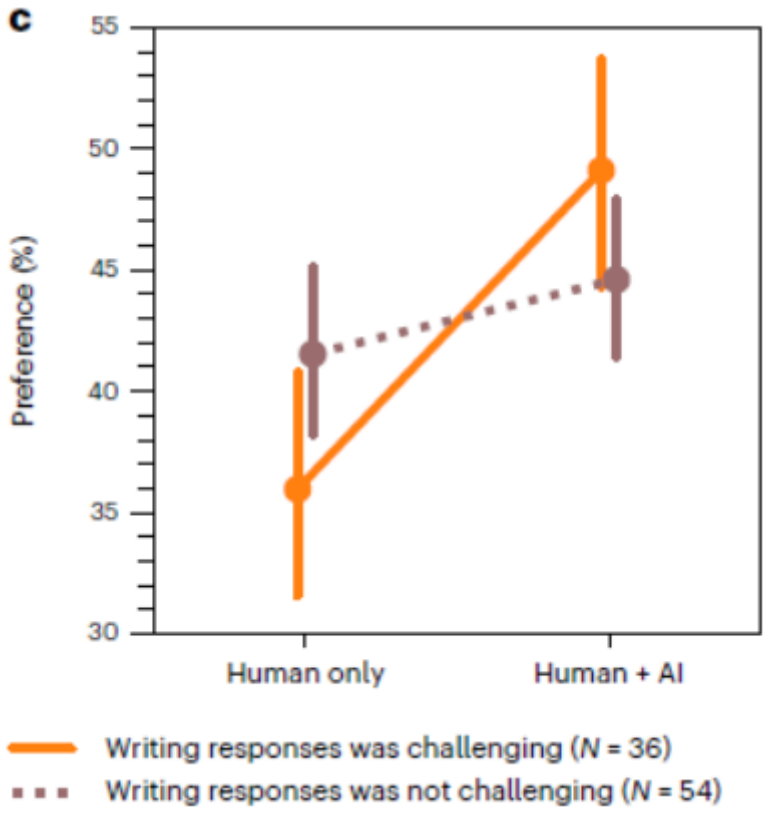
结果b

通过使用基于AI的同理心自动评估得分，我们发现人类+AI回复的得分比仅人单独回复高19.6% (1.77相比与1.48; Cohen's d = 0.24，P = 5.1 × 10-8，t = 5.46，d.f = 2,998，双侧t检验)。

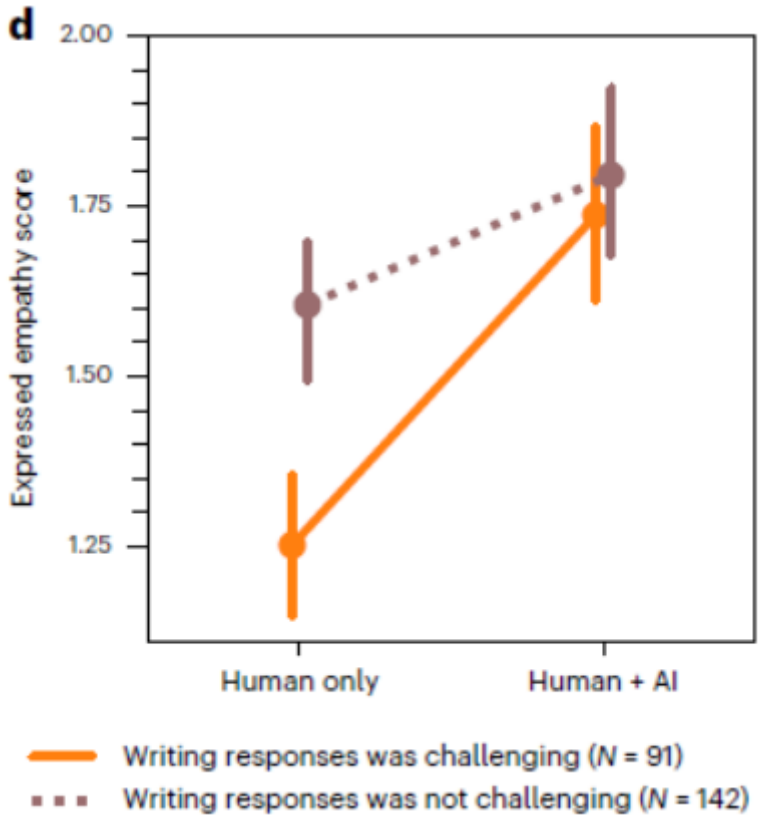


结果c

无论是回复时未遇到困难的参与者还是回复时遇到困难的参与者，TalkLife用户均更偏好人类+AI的回复



结果d
无论是回复时未遇到困难的参与者还是回复时遇到困难的参与者，人类+AI回复均比人单独回复具有更高的同理心得分。



讨论

本研究展示了人类和AI是如何在开放式、社会性和高风险的任务上进行合作，比如进行同理心的对话。同理心是复杂且微妙的，因此对AI来说比其他许多人类-AI协作的任务更具挑战性。
该研究对于解决精神卫生健康方面的障碍具有启示意义，因为现有的资源和干预措施不足以满足当前和新出现的需求。根据世界卫生组织报告，全球有超过4亿人患有精神疾病，其中大约3亿人患有抑郁症。总体而言，精神疾病和相关的行为健康问题占全球疾病负担的13%，高于心血管疾病和癌症。尽管心理治疗和社会支持可能是有效的治疗方法，但许多弱势个体获得治疗和咨询的机会有限。在全球范围内改善获得心理健康支持的一个可扩展的方法便是使用在线平台，如TalkLife (TalkLife.com)、 YourDost (YourDost)，将寻求支持者与同伴支持者联系起来。
然而，这样做的关键挑战在于，使未经训练的同伴支持者和有需要的人之间能够

进行有效且高质量的对话。而本篇研究表明，人类-AI协作可以大大增加同伴支持者的回复中的同理心，确保寻求支持者的理解和接受。
总而言之，尽管人工智能还不能完全取代心理治疗师进行心理治疗和社会支持，但通过适当的人工智能辅助技术来增强未经训练的同伴支持者是可行的。

参考文献

Sharma, A., Lin, I.W., Miner, A.S. *et al.* Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nat Mach Intell* **5**, 46–57 (2023). https://doi.org/10.1038/s42256-022-00593-2