

生成式人工智能外部特征对信任与使用意愿的影响

周 晨^{1*}, 袁熙庆², 王 鹏^{3#}

¹温州大学教育学院, 浙江 温州

²山东师范大学心理学院, 山东 济南

³小明生涯(济南)教育信息服务有限公司, 山东 济南

收稿日期: 2025年4月24日; 录用日期: 2025年5月16日; 发布日期: 2025年5月29日

摘 要

本研究聚焦生成式人工智能, 以ChatGPT和文心一言为例, 探讨头像类型、面孔吸引力、性别一致性等外部特征对用户信任度与使用意愿的影响机制。结合社会存在理论、性别认同一致性理论及技术接受模型分析发现, 头像类型在不同场景影响各异, 面孔吸引力通过“美即好”效应起作用, 性别一致性可缩小心理距离促进信任, 且吸引力可能放大性别一致性效应。研究为优化生成式人工智能设计提供理论参考, 也指出情境多样性不足等局限及未来研究方向。

关键词

生成式人工智能, 外部特征, 信任度, 使用意愿, 影响机制

The Influence of Generative Artificial Intelligence External Features on Trust and Willingness to Use

Chen Zhou^{1*}, Xiqing Yuan², Peng Wang^{3#}

¹School of Education, Wenzhou University, Wenzhou Zhejiang

²School of Psychology, Shandong Normal University, Jinan Shandong

³Xiaoming Career (Jinan) Education Information Service Ltd., Jinan Shandong

Received: Apr. 24th, 2025; accepted: May 16th, 2025; published: May 29th, 2025

*第一作者。

#通讯作者。

文章引用: 周晨, 袁熙庆, 王鹏(2025). 生成式人工智能外部特征对信任与使用意愿的影响. *心理学进展*, 15(5), 537-547. DOI: 10.12677/ap.2025.155329

Abstract

This study focuses on Generative Artificial Intelligence, taking ChatGPT and Wenxin Yiyan as examples, to explore the influence mechanisms of external features, such as avatar type, facial attractiveness, and gender congruence on users' trust and willingness to use. Analyses based on the Social Presence Theory, Gender Identity Congruence Theory, and Technology Acceptance Model reveal that avatar types have varying effects in different scenarios, facial attractiveness works through the "what is beautiful is good" effect, gender congruence can narrow psychological distance to promote trust, and attractiveness may amplify the gender congruence effect. The study provides theoretical references for optimizing Generative Artificial Intelligence design, while also pointing out limitations, such as insufficient situational diversity and future research directions.

Keywords

Generative Artificial Intelligence, External Features, Trust, Willingness to Use, Influence Mechanism

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 生成式人工智能

聊天机器人(Chatbot)也称为对话代理(Conversational Agent)。**Vaidyam 等人(2019)**将生成式人工智能定义为以硬件(如亚马逊的 Echo)或软件(如苹果的 Siri)形式存在的数字工具,它们利用机器学习和人工智能方法来模仿人类行为并能够参与对话。近年来,随着技术的发展和庞大语料库的支持,大语言模型生成式人工智能(下文简称:生成式人工智能)技术迅速发展并被广泛应用。截至 2024 年 2 月,OpenAI 的人工智能 ChatGPT,全球用户数已超 18 亿;百度旗下的人工智能“文心一言”用户数也已超 1 亿,模型调用位居国内第一(**郑新钰, 2024**)。并且由于生成式人工智能强大的性能以及广泛的适用性,人们用它来进行翻译、文案编写、邮件编辑、代码生成等活动。甚至专业测验结果显示,ChatGPT 可以在研究生入学考试、律师资格考试、专业水平测试甚至品酒师资质测试中取得中等或者优秀的成绩(**朱永新, 杨帆, 2023**)。文心一言更擅长中文任务情境的处理(**赵广立, 2023**)。二者为国内外大语言模型生成式人工智能(下文简称生成式人工智能)技术的领导者,故本研究使用 ChatGPT 和文心一言代表生成式人工智能开展具体研究。

以 ChatGPT 为代表的生成式人工智能具有极高的性能。这来自于大量数据、强大计算能力以及新颖的处理技术的结合,即数据、算法和算力(**顾小清等, 2023**)。以 ChatGPT 为例,在数据方面,相关资料显示,GPT-3 拥有 1750 亿个模型参数,接受了 570 GB 的文本训练,具体的训练数据包括维基百科、书籍、杂志期刊、网络内容以及代码库等数据源,涵盖了丰富多样的文本类型和领域(**张治, 2023; Qadir, 2023; Tamkin et al., 2021**)。在算法方面,ChatGPT 依托于人类反馈强化学习技术,具有突出的自然语言处理能力,在算力方面,ChatGPT 的总算力消耗约为 3640PF-days (即假如每秒计算一千万亿次,需要计算 3640 天;**陈静远等, 2023**)。同理,文心一言则基于百度庞大的数据集,也具有相似的能力。此外,由于生成式人工智能是以网页端或 APP 的形式呈现,在当前互联网普及的背景下,人们可以便捷地获得生成式人工智能的帮助。因此,生成式人工智能被广泛应用于各个领域。

2. 生成式人工智能的外部特征

在用户使用聊天机器人的过程中,二者的交互往往不仅是信息的传递,还包括情感表达、社交互动等多种因素(Rheu et al., 2021)。而聊天的外部特征(如头像、界面设计等)对用户的心理和行为具有重要影响作用,一个外观友好且符合人类视觉认知的聊天机器人可能会更容易被用户接受,甚至研究发现外部特征的吸引力比可靠性更重要(Yuksel et al., 2017)。这包括拟人化的头像、头像所表现出的吸引力以及性别特征等,这些特征有助于建立用户对机器人的信任,增加他们的使用意愿(Li & Suh, 2022),并且提升整体的用户体验。

2.1. 头像类型

2.1.1. 概念界定

头像(Avatar)是指在线交互中作为用户和机器人代表的由网站和计算机媒体提供的计算机生成的视觉图像(Nowak & Rauh, 2005)。在聊天机器人方面,开发者可以根据自身偏好或特定的目的上传动物、卡通或真人等图片作为其对外展示的形象。头像一般具有拥有眼睛、发色类似于真人的特征。社会反应理论指出,个体会对计算机或机器做出如同计算机是另一个人类的反应(Reeves & Nass, 1996)。虽然个体明白计算机并非真正的人类(Nass & Steuer, 1993),但是人们也通过社交方式自发地与机器进行互动和交流。

关于头像的类型,有研究按照其物理维度(如二维、三维等),拟人化程度(如高、低)以及形象类型(如动物、拟人和真人等)等方面进行分类。已有研究发现,不同的头像类型会影响人们对聊天机器人感知的吸引力、可信度以及使用意愿等因素(Mull et al., 2015)。

2.1.2. 相关理论

聊天机器人的头像呈现对教育用户的影响可以用社会存在感来解释,拟人化的设计线索促进了在互动过程中用户对客体的感知,进而加强了对社会存在的感知。同时,人们倾向于将品牌或数字服务等实体人性化,然后将自我与聊天机器人的概念进行比较,当一个人与另一个体或物体的自我概念之间相似性较高时,其心理距离会更近。研究发现感知到的人类相似性会增加人们对机器人的偏好、使用和参与(Al-Natour et al., 2011)。因此,当聊天机器人的头像呈现不同类型,如卡通和真人时,由于在不同头像类型下二者之间的自我概念相似性存在差异,用户使用聊天机器人时的使用、偏好和参与也会存在一定的差异。

2.2. 面孔吸引力

2.2.1. 概念界定

面孔吸引力(Facial attractiveness)是面孔特征的重要指标,Rhodes (2006)认为面孔吸引力是指目标人物面孔所产生的一种积极愉悦的情绪体验,并促使个体产生接近的意愿。该定义被相关研究者广泛使用。通常,面孔吸引力受面孔生理因素、面孔特征和情景三方面的影响。吸引力对人们的第一印象影响巨大,拥有吸引力的人会得到其他人更多积极特质上的期待(Adams & Crane, 1980)。且吸引力对人类之间的影响与人机交互之间的影响相似(Calhoun et al., 2019),人们也通过聊天机器人的面部特征来评估其可靠性或真实性等(Kim et al., 2020)。由此可见,生成式人工智能的面孔吸引力可能对用户的信任度和使用意愿起重要作用。

2.2.2. 相关理论

社会反应理论(Social response theory)是由 Reeves 和 Nass (1996)提出,该理论指出,人们与计算机等媒体设备的互动就像人与人之间的互动一样,人们会将这些媒体设备视为现实生活中的人类,并且无意

识地使用一些与人交往的社会规则与这些媒体设备进行互动和交流。这些社会规则包括信任、互惠等内容。人们对待媒体设备的这种反应是针对设备本身而非背后的程序设计人员。即使人们意识到自己正在与计算机进行交互,他们依然会使用一些与人交往的社会规则,这表明人们对于媒介设备的社会反应是无意识的、自发的。该理论被广泛应用于人机交互领域,用来探究人们与机器互动时的心理和行为反应。

2.3. 性别一致性

2.3.1. 概念界定

性别一致性(Gender Congruence)是指用户的性别与他人性别的匹配(Pitardi et al., 2023)。其中,性别是一个综合意义的词。包括两个方面,一是由染色体决定的生物解剖学性征差异,二是由社会的性别角色划分、身体规范等文化产生的性别差异。心理学家斯托勒(Robert Stoller)区分了二者之间的差异。斯托勒认为,社会性别指基于生理性别的一系列性别差异,这些差异是由社会文化所建构的(Stoller, 2020)。本文涉及到的用户性别为社会上的大多数群体,即性别认同与自身生理性别一致的个体。聊天机器人的性别是其呈现出的性别特征倾向,如聊天机器人呈现的是女性头像,则就判定为女性聊天机器人。此外,社会心理学研究表明,是性别一致性而非性别本身(即男性或女性)对于社会认同和关系的发展至关重要(Alabed et al., 2022; Bakar & McCann, 2014; Grohmann, 2009)。

2.3.2. 相关理论

性别认同一致性理论(Gender-identity Congruency Theory)是由 Grubb 和 Grathwohl (1967)基于社会认同理论(Social Identity Theory)提出的。社会认同理论认为,个体通过社会分类、社会比较和积极区分这三个过程形成对某一群体的社会认同。Grubb 和 Grathwohl (1967)认为,对于个体而言,性别是一种自我分类的方式。他们将形象认知、价值观和情感融入到人机交互过程中,并在性别认同的驱动下更倾向于选择与自身性别相符的聊天机器人或产品。

3. 生成式人工智能的外部特征与信任度

3.1. 信任度

3.1.1. 概念界定

信任(Trust)作为人机交互中的重要因素,是用户体验的一个组成部分。尤其是聊天机器人的反应对用户具有重要意义时,考虑人机交互中的信任因素十分必要(Zamora, 2017)。信任(Trust)是一个动态的概念,涉及多个阶段,包括建立,对他人的稳定信任,再到受到破坏后的恢复。Mayer 等人(1995)把信任定义为一个人对另一个人的能力、仁慈和正直信念的综合体。信任同样存在于人机交互过程中,并发挥重要作用。

3.1.2. 相关研究

由于聊天机器人不仅用于娱乐或便利服务,还被用于各个专业、隐私或高风险领域,如教育、决策等(Zamora, 2017),因此人机交互过程中的信任问题需要十分重视。为了更好的建立人机交互时的用户信任度,已有研究分析了可能影响人机交互信任度的因素,例如 Følstad 等人(2018)通过使用半结构化访谈,让 13 位参与者反馈对聊天机器人客户体验的信任体验。结果表明,影响用户信任聊天机器人的关键因素包括:正确解释用户请求、提供有用的回应、类人的回复方式以及清楚表达自我的能力。也有研究(Rheu et al., 2021)分析了聊天机器人影响用户信任度的五大特征,分别为社交智力、沟通风格、物理外观、非语言的沟通特征以及表现性能。并且这些影响特征通常也受其他因素影响,如聊天机器人的任务背景、当前任务、代理的角色、使用者的特征以及需要等。

3.2. 生成式人工智能的外部特征对信任度的影响

3.2.1. 头像类型对信任度的影响

目前研究表明, 聊天机器人的头像类型会影响用户的信任度。通常认为, 一个互动的动画形象或化身是强大的社会存在的来源, 呈现人物形象可以增强人与人之间的联系(Keeling et al., 2010)。类人界面可以通过提高感知能力来增加用户对技术的信任(Li & Suh, 2022), 并能缓解信任度的下降(de Visser et al., 2016)。已有研究发现, 在消费背景下, 呈现头像的聊天机器人可以增加人们的购买意向(Yen & Chiang, 2021), 包括提高购买行为中的信任度。但呈现不同类型头像的聊天机器人对用户使用行为的影响可能存在差异, 如 Salminen 等人(2021)的研究发现, 在用户感知头像角色的任务背景下, 拥有更逼真图片的角色也被认为是更合群、更开放、情绪更稳定的, 用户对这些图片的信心也更高。Toure-Tillery 和 McGill (2015)发现在广告代言任务背景下, 拟人化聊天机器人比人类代言人对用户具有更高的说服力。甚至 Moussawi 等人(2021)还发现在日常使用任务背景下, 聊天机器人头像拟人化不会显著影响用户信任。由此看来, 聊天机器人头像类型对用户信任度存在影响作用, 且这种影响还可能在不同情境下存在差异。

3.2.2. 面孔吸引力对信任度的影响

除了头像类型对人们的心理和行为存在影响, 在生涯规划教育背景下, 由于往往是真人之间进行的, 而真人面孔的吸引力往往对人们的感知存在重要影响。通常来说, 更有吸引力外观的人会在其他方面得到更多的积极品质的期待, 即“吸引力光环”。如陌生人归因的文献已证实“美的就是好的”的刻板印象的存在(Adams & Crane, 1980; Berscheid & Walster, 1974)。吸引力会引起个体对有吸引力和没有吸引力的目标的行为和特征的不同预期。面孔吸引力是吸引力的重要维度, 也是聊天机器人拟人化的重要方面之一。聊天机器人的面孔吸引力也会对个体的心理或行为产生一定的影响, 如研究发现视觉呈现增加了与聊天机器人进行对话的用户的社会存在感, 有助于促进用户与聊天机器人的社交行为(Kim & Sundar, 2012)。感知的面部吸引力与感知的可信度之间存在显著正相关, 早期的吸引力印象可能会倾向于可信度推断(Gutierrez-Garcia et al., 2019)。Zogaj et al. (2023)进一步发现, 与那些没有吸引力的代理人相比, 有身体吸引力的生成式人工智能更受信任, 而不受其专业知识高低的影响。由此可见, 生成式人工智能的面孔吸引力会影响用户对生成式人工智能的信任, 并且由于“吸引力光环”具有广泛的适用性。

3.2.3. 性别一致性对信任度的影响

研究普遍认为聊天机器人的性别有助于用户和机器人之间建立共同点, 进而促进人机交互(Eyssel & Hegel, 2012)。换言之, 当用户与聊天机器人呈现的性别特征一致时, 用户往往会感知到更多的共同点, 心理距离进一步缩小, 进而产生更多的信任行为。如 Kim 等人(2021)发现, 人们具有更喜欢与自己同性别的聊天机器人的趋势, 但这种趋势在不同任务背景下可能存在一定差异(Rheu et al., 2021), 如 Elkins 和 Derrick (2013)检验了在模拟机场筛选面试背景下, 将聊天代理的性别与用户的性别相匹配是否会增加信任, 结果显示代理人与被试性别是否相同在信任度上没有显著差异; 但 Nunamaker 等人(2011)发现, 在聊天机器人征集捐款的背景下, 性别效应取决于参与者的性别, 即男性参与者更信任女性声音的代理, 在女性参与者中并不存在这种跨性别效应。聚焦于教育背景, Zhang (2019)等人对国内 28 个县级单位的初中教师进行调查, 研究发现女数学教师显著提高了女学生的学习成绩。的研究也发现, 学生和导师的性别一致性对学业薄弱和没有 STEM (Science, Technology, Engineering, Mathematics)导向的学生特别有帮助。可以看出, 在教育背景下, 师生性别一致性提高了学生学习成绩, 并产生了较好的积极影响。

3.2.4. 面孔吸引力和性别一致性的交互作用对信任度的影响

根据上述研究,本研究得知面孔吸引力和性别一致性可以影响用户在使用聊天机器人时的信任度,即人们更愿意信任吸引力更高的、与自身性别一致的聊天机器人(Adams & Crane, 1980; Berscheid & Walster, 1974; Gutierrez-Garcia et al., 2019; Zogaj et al., 2023)。现有研究发现,吸引力会强化个体的性别特征,从而加强人们对其性别属性的感知(Ahuja & Pundir, 2019; Desrumaux et al., 2009)。有吸引力的女性和男性被认为比没有吸引力的同性别人士更具有其性别特征(Ahuja & Pundir, 2019)。同时,这种吸引力效应取决于个体所属职业类型是否具有性别特征,即人们对从事该职业的个体是否存在性别要求(Nault et al., 2020)。当个体所属职业类型存在性别要求时,吸引力效应会增强该性别效应(如性别一致性)的作用;但若相应职位不存在性别要求,个体的吸引力仅同等程度地提高自身的魅力。

在人机交互过程中,人们会对计算机或机器做出如同计算机是另一个人类的反应(Reeves & Nass, 1996)。由于,个体往往在性别认同的驱动下更倾向于选择与自身性别相符的生成式人工智能(Alabed et al., 2022),即聊天机器人所代表的职位类型存在性别特征。因此,人机交互过程中,当聊天机器人呈现出较高的面孔吸引力时,不仅会提高聊天机器人自身优势(Ahuja & Pundir, 2019),还会加强人们对其性别的感知,增强性别一致性效应,即人们会更愿意信任和使用与自身性别一致性的聊天机器人。而聊天机器人呈现出较低面孔吸引力时,人们对聊天机器人性别属性的感知会稍弱,性别一致性效应的影响也相对较小。

4. 生成式人工智能的外部特征与使用意愿

4.1. 使用意愿

4.1.1. 概念界定

使用意愿(Intention to Use)是指个体对某种行为期待实施的主观意识强度(Fishbein & Ajzen, 1977)。创新技术得到广泛应用,其关键在于如何让用户了解并接受它。在人机交互研究中,使用意愿是需要着重考虑的因素(Lu et al., 2019)。目前,聊天机器人已被广泛应用于个人或商业服务(Benlian et al., 2020; Hasan et al., 2021),这些生成式人工智能在推广或部署时,都需要重点考虑被用户使用的愿意以获得更好的应用效果(Gopinath & Kasilingam, 2023)。

4.1.2. 相关理论

技术接受模型(Technology Acceptance Model, TAM)是用户接受和使用行为研究中广泛使用的模型之一,由 Davis (1986)提出。该模型(见图 1)提出了影响用户使用行为的两个主要的决定因素,即感知有用性(Perceived Usefulness)和感知易用性(Perceived Ease of Use)。感知有用性指用户认为使用特定系统能提高工作绩效的程度,而感知易用性则指用户认为系统使用起来的难易程度(Davis, 1989)。此外,模型还指出用户对信息技术的使用行为(Behavior)受行为意向(Behavior Intention)的影响,而行为意向由感知有用性和使用态度(Attitude)共同影响决定。使用态度受感知有用性和感知易用性的共同影响,而感知易用性还可以直接影响感知有用性。该模型假设外部变量(External Variables)会直接影响感知有用性和感知易用性。

4.2. 生成式人工智能外部特征对使用意愿的影响

4.2.1. 头像类型对使用意愿的影响

聊天机器人的头像类型会影响用户的使用意愿。具体来说,具有类人特征的机器人更智能和可靠,也会被人们认为是值得信赖的,也更容易被接受和使用(Han & Yang, 2018)。这是由于机器人在被赋予类人特征后,人们会对其产生温暖的感知(Kim et al., 2019)。研究发现,用户通常认为人形(与非人形相比)

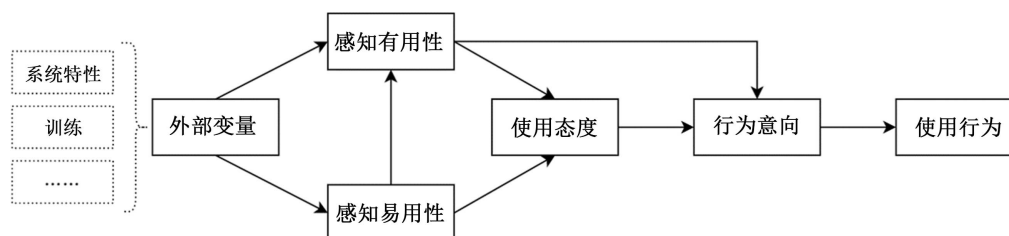


Figure 1. Technology acceptance model

图 1. 技术接受模型

服务机器人具有更多令人温暖的属性,例如友好、善于交际、讨人喜欢和善良(Tussyadiah & Park, 2018),而对能力感知上没有差异(Choi et al., 2021)。Mull 等人(2015)探究了多种类型头像对消费者感知可信度、同质性、吸引力和互动意向的影响。结果发现,与想象、动物和人形头像相比,人类头像被认为在可信度、同质性、吸引力方面的感知最强,被使用意愿也最高。但也有研究发现 Nowak & Rauh (2005)发现,在社交判断任务背景下,与不太像人类的聊天机器人相比,更像人类的聊天机器人被认为不那么可信和被人喜欢;由此可见,聊天机器人的头像类型可以影响用户的使用意愿等因素。但考虑到人们在不同情境下的偏好可能存在差异。根据 Prebensen & Rosengren (2016)对聊天机器人使用情境的分类,即功利主义和享乐主义,功利主义的任务背景特征是功能质量和物有所值,目的是实现便捷、务实、经济、快捷、节约(Liu et al., 2022),与此对应,生涯教育属于教育领域,其特征是知识、严谨、务实等,二者相匹配。同时由于在功利主义背景任务下,人们对有能力的机器人表现出更高的使用意愿,呈现真人头像更契合有能力的聊天机器人形象(Glocker et al., 2009; Keating et al., 2003)。

4.2.2. 面孔吸引力对使用意愿的影响

生成式人工智能的面孔吸引力会影响用户的使用意愿。具体来说,聊天机器人头像所呈现的面孔形象是用户最直观感知的特征。对用户来说,更具有吸引力外观的人会在其他方面得到更多的积极期待。如研究发现,吸引力和可信度之间的显著相关性在四种强度的六种情绪表达和中性面孔中始终存在(Gutierrez-Garcia et al., 2019)。在购买背景下,与那些没有吸引力的聊天机器人相比,有身体吸引力的生成式人工智能更受信任(Zogaj et al., 2023)。并且由于用户的信任度可以正向影响人们的使用意愿(Arifi et al., 2021; Hooda et al., 2022)。

4.2.3. 性别一致性对使用意愿的影响

人机性别一致性会影响用户的使用意愿。普遍观点认为,性别的社会线索可以减少人们在人机互动过程中的努力程度(Tay et al., 2014)。因此,聊天机器人的性别有助于用户和机器人之间建立共同点,进而促进人机交互的行为(Eyssel & Hegel, 2012)。早期研究也表明,聊天机器人的性别可以影响用户行为(Beldad et al., 2016; McDonnell & Baxter, 2019),并且用户对表现出人类性别的社交机器人持积极态度,人们更容易接受与自身性别和个性类型一致的机器人(Tay et al., 2014)。研究还发现,用户与生成式人工智能的性别匹配会促进用户对自我一致性的感知,进而促进用户的使用意愿。在消费任务中,男性用户更喜欢男性生成式人工智能,女性用户更喜欢女性生成式人工智能,非二元性别的用户更喜欢中立的生成式人工智能(Zogaj et al., 2023)。

4.2.4. 面孔吸引力和性别一致性的交互作用对使用意愿的影响

根据上述研究发现,面孔吸引力和性别一致性可以影响用户在使用聊天机器人时的使用意愿,即人们更愿意使用吸引力更高、与自身性别一致的聊天机器人(Gutierrez-Garcia et al., 2019; Zogaj et al., 2023)。

此外,前人研究还发现,吸引力会强化个体的性别特征,从而加强人们对其性别属性的感知(Ahuja & Pundir, 2019; Desrumaux et al., 2009)。同时,这种吸引力效应取决于人们对从事该职业的个体是否存在性别要求(Nault et al., 2020)。当个体所属职业类型存在性别要求时,吸引力效应会增强该性别效应(如性别一致性)的作用。

在人机交互过程中,人们会对计算机或机器做出如同计算机是另一个人类的反应(Reeves & Nass, 1996)。同时,个体往往在性别认同的驱动下更倾向于选择与自身性别相符的生成式人工智能(Alabed et al., 2022),当聊天机器人所代表的职位类型存在性别要求时,吸引力会增强用户对聊天机器人的性别特征感知(Nault et al., 2020),因此,当生成式人工智能呈现出较高的面孔吸引力时,不仅会提高生成式人工智能的吸引力(Ahuja & Pundir, 2019),还会增强人们选择使用与自身性别一致的生成式人工智能的倾向;而生成式人工智能呈现出较低面孔吸引力时,人们对生成式人工智能性别属性的感知会削弱,选择同性别生成式人工智能的倾向会相对较小。

5. 小结

生成式人工智能作为人工智能技术的重要应用场景,其外部特征对用户信任度与使用意愿的影响机制逐渐成为研究热点。本文系统梳理了头像类型、面孔吸引力、性别一致性等关键特征的作用路径,并结合社会存在理论、性别认同一致性理论及技术接受模型(TAM)展开分析。

头像类型在享乐主义与功利主义场景中呈现差异化影响:卡通或真人头像通过增强社会存在感提升信任,但功利导向任务中人类头像更易被信任,因其传递能力属性。面孔吸引力通过“美即好”效应显著影响信任与使用意愿,吸引力越高,用户越易感知可靠性,尤其在教育等专业场景中,“吸引力光环”可能强化职业能力联想。性别一致性则通过缩小心理距离促进信任,用户更倾向选择与自身性别匹配的机器人,尤其在性别角色鲜明的任务(如教育辅导)中,性别一致性可提升自我认同感。

研究还揭示了特征间的交互作用:吸引力可能放大性别一致性效应,高吸引力个体更易强化性别标签认知,从而影响信任判断。此外,技术接受模型显示,外部特征通过感知有用性(如效率提升)和易用性(如交互流畅)间接作用于使用意愿,信任在其中起到中介作用。

当前研究仍存在局限,如情境多样性不足、跨文化差异探讨较少。未来需结合神经科学技术深化认知机制研究,并关注伦理问题(如性别刻板印象强化)。总体而言,优化生成式人工智能外部特征需兼顾任务需求与用户心理,以实现人机交互的效能最大化。

参考文献

- 陈静远, 胡丽雅, 吴飞(2023). ChatGPT/生成式人工智能促进以知识点为核心的教学模式变革研究. *华东师范大学学报(教育科学版)*, 41(7), 177-186.
- 顾小清, 胡艺龄, 郝祥军(2023). AGI 临近了吗: ChatGPT 热潮之下再看人工智能与未来教育发展. *华东师范大学学报(教育科学版)*, 41(7), 117-130.
- 张治(2023). ChatGPT/生成式人工智能重塑教育的底层逻辑和可能路径. *华东师范大学学报(教育科学版)*, 41(7), 131-142.
- 赵广立(2023-03-23). 文心一言是如何炼成的? *中国科学报*, 3.
- 郑新钰(2024-01-15). 百度文心一言用户数破 1 亿专业版用户开测智能体模式. *中国城市报*, A10.
- 朱永新, 杨帆(2023). ChatGPT/生成式人工智能与教育创新: 机遇、挑战以及未来. *华东师范大学学报(教育科学版)*, 41(7), 1-14.
- Adams, G. R., & Crane, P. (1980). An Assessment of Parents' and Teachers' Expectations of Preschool Children's Social Preference for Attractive or Unattractive Children and Adults. *Child Development*, 51, 224-231.
<https://doi.org/10.2307/1129610>

- Ahuja, K. K., & Pundir, T. (2019). No 'Pretty' Please at Work: Studying Effect of Gender and Physical. *International Journal of Human Resource Management and Research*, 9, 83-94.
- Alabed, A., Javornik, A., & Gregory-Smith, D. (2022). AI Anthropomorphism and Its Effect on Users' Self-Congruence and Self-AI Integration: A Theoretical Framework and Research Agenda. *Technological Forecasting and Social Change*, 182, Article 121786. <https://doi.org/10.1016/j.techfore.2022.121786>
- Al-Natour, S., Benbasat, I., & Cenfetelli, R. (2011). The Adoption of Online Shopping Assistants: Perceived Similarity as an Antecedent to Evaluative Beliefs. *Journal of the Association for Information Systems*, 12, 347-374. <https://doi.org/10.17705/1jais.00267>
- Arfi, W. B., Nasr, I. B., Kondrateva, G., & Hikkerova, L. (2021). The Role of Trust in Intention to Use the IoT in E Health: Application of the Modified UTAUT in a Consumer Context. *Technological Forecasting and Social Change*, 167, Article 120688. <https://doi.org/10.1016/j.techfore.2021.120688>
- Bakar, H. A., & McCann, R. M. (2014). Matters of Demographic Similarity and Dissimilarity in Supervisor-Subordinate Relationships and Workplace Attitudes. *International Journal of Intercultural Relations*, 41, 1-16. <https://doi.org/10.1016/j.ijintrel.2014.04.004>
- Beldad, A., Hegner, S., & Hoppen, J. (2016). The Effect of Virtual Sales Agent (VSA) Gender—Product Gender Congruence on Product Advice Credibility, Trust in VSA and Online Vendor, and Purchase Intention. *Computers in Human Behavior*, 60, 62-72. <https://doi.org/10.1016/j.chb.2016.02.046>
- Benlian, A., Klumpe, J., & Hinz, O. (2020). Mitigating the Intrusive Effects of Smart Home Assistants by Using Anthropomorphic Design Features: A Multimethod Investigation. *Information Systems Journal*, 30, 1010-1042. <https://doi.org/10.1111/isj.12243>
- Berscheid, E., & Walster, E. (1974). Physical Attractiveness. In *Advances in Experimental Social Psychology* (pp. 157-215). Elsevier. [https://doi.org/10.1016/s0065-2601\(08\)60037-4](https://doi.org/10.1016/s0065-2601(08)60037-4)
- Calhoun, C. S., Bobko, P., Gallimore, J. J., & Lyons, J. B. (2019). Linking Precursors of Interpersonal Trust to Human-Automation Trust: An Expanded Typology and Exploratory Experiment. *Journal of Trust Research*, 9, 28-46. <https://doi.org/10.1080/21515581.2019.1579730>
- Choi, S., Mattila, A. S., & Bolton, L. E. (2021). To Err Is Human(-Oid): How Do Consumers React to Robot Service Failure and Recovery? *Journal of Service Research*, 24, 354-371. <https://doi.org/10.1177/1094670520978798>
- Davis, F. D. (1986). *A Technology Acceptance Model for Empirically Testing New End-User Information Systems*. Massachusetts Institute of Technology.
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13, Article 319. <https://doi.org/10.2307/249008>
- de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F. et al. (2016). Almost Human: Anthropomorphism Increases Trust Resilience in Cognitive Agents. *Journal of Experimental Psychology: Applied*, 22, 331-349. <https://doi.org/10.1037/xap0000092>
- Desrumaux, P., De Bosscher, S., & Léoni, V. (2009). Effects of Facial Attractiveness, Gender, and Competence of Applicants on Job Recruitment. *Swiss Journal of Psychology*, 68, 33-42. <https://doi.org/10.1024/1421-0185.68.1.33>
- Elkins, A. C., & Derrick, D. C. (2013). The Sound of Trust: Voice as a Measurement of Trust during Interactions with Embodied Conversational Agents. *Group Decision and Negotiation*, 22, 897-913. <https://doi.org/10.1007/s10726-012-9339-x>
- Eyssel, F., & Hegel, F. (2012). (S)he's Got the Look: Gender Stereotyping of Robots. *Journal of Applied Social Psychology*, 42, 2213-2230. <https://doi.org/10.1111/j.1559-1816.2012.00937.x>
- Fishbein, M., & Ajzen, I. (1977). Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. *Contemporary Sociology*, 6, Article 244.
- Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What Makes Users Trust a Chatbot for Customer Service? An Exploratory Interview Study. In *Lecture Notes in Computer Science* (pp. 194-208). Springer International Publishing. https://doi.org/10.1007/978-3-030-01437-7_16
- Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Valdez, J. N., Griffin, M. D. et al. (2009). Baby Schema Modulates the Brain Reward System in Nulliparous Women. *Proceedings of the National Academy of Sciences*, 106, 9115-9119. <https://doi.org/10.1073/pnas.0811620106>
- Gopinath, K., & Kasilingam, D. (2023). Antecedents of Intention to Use Chatbots in Service Encounters: A Meta-Analytic Review. *International Journal of Consumer Studies*, 47, 2367-2395. <https://doi.org/10.1111/ijcs.12933>
- Grohmann, B. (2009). Gender Dimensions of Brand Personality. *Journal of Marketing Research*, 46, 105-119. <https://doi.org/10.1509/jmkr.46.1.105>
- Grubb, E. L., & Grathwohl, H. L. (1967). Consumer Self-Concept, Symbolism and Market Behavior: A Theoretical Approach.

- Journal of Marketing*, 31, 22-27. <https://doi.org/10.1177/002224296703100405>
- Gutiérrez-García, A., Beltrán, D., & Calvo, M. G. (2019). Facial Attractiveness Impressions Precede Trustworthiness Inferences: Lower Detection Thresholds and Faster Decision Latencies. *Cognition and Emotion*, 33, 378-385. <https://doi.org/10.1080/02699931.2018.1444583>
- Han, S., & Yang, H. (2018). Understanding Adoption of Intelligent Personal Assistants. *Industrial Management & Data Systems*, 118, 618-636. <https://doi.org/10.1108/imds-05-2017-0214>
- Hasan, R., Koles, B., Zaman, M., & Paul, J. (2021). The Potential of Chatbots in Travel and Tourism Services in the Context of Social Distancing. *International Journal of Technology Intelligence and Planning*, 13, 63.
- Hooda, A., Gupta, P., Jeyaraj, A., Giannakis, M., & Dwivedi, Y. K. (2022). The Effects of Trust on Behavioral Intention and Use Behavior within E-Government Contexts. *International Journal of Information Management*, 67, Article 102553. <https://doi.org/10.1016/j.ijinfomgt.2022.102553>
- Keating, C. F., Randall, D. W., Kendrick, T., & Gutshall, K. A. (2003). Do Baby-Faced Adults Receive More Help? The (Cross-Cultural) Case of the Lost Resume. *Journal of Nonverbal Behavior*, 27, 89-109.
- Keeling, K., McGoldrick, P., & Beatty, S. (2010). Avatars as Salespeople: Communication Style, Trust, and Intentions. *Journal of Business Research*, 63, 793-800. <https://doi.org/10.1016/j.jbusres.2008.12.015>
- Kim, S. Y., Schmitt, B. H., & Thalmann, N. M. (2019). Eliza in the Uncanny Valley: Anthropomorphizing Consumer Robots Increases Their Perceived Warmth but Decreases Liking. *Marketing Letters*, 30, 1-12. <https://doi.org/10.1007/s11002-019-09485-9>
- Kim, S., Lee, S. Y., & Lee, J. (2021). Male, Female, or Robot: Effects of Task Type and User Gender on Expected Gender of Chatbots. *Journal of Korea Multimedia Society*, 24, 320-327.
- Kim, W., Kim, N., Lyons, J. B., & Nam, C. S. (2020). Factors Affecting Trust in High-Vulnerability Human-Robot Interaction Contexts: A Structural Equation Modelling Approach. *Applied Ergonomics*, 85, Article 103056. <https://doi.org/10.1016/j.apergo.2020.103056>
- Kim, Y., & Sundar, S. S. (2012). Anthropomorphism of Computers: Is It Mindful or Mindless? *Computers in Human Behavior*, 28, 241-250. <https://doi.org/10.1016/j.chb.2011.09.006>
- Li, M., & Suh, A. (2022). Anthropomorphism in Ai-Enabled Technology: A Literature Review. *Electronic Markets*, 32, 2245-2275. <https://doi.org/10.1007/s12525-022-00591-7>
- Liu, X., Yi, X., & Wan, L. C. (2022). Friendly or Competent? The Effects of Perception of Robot Appearance and Service Context on Usage Intention. *Annals of Tourism Research*, 92, Article 103324. <https://doi.org/10.1016/j.annals.2021.103324>
- Lu, L., Cai, R., & Gursay, D. (2019). Developing and Validating Service Robot Integration Willingness Scale. *International Journal of Hospitality Management*, 80, 36-51. <https://doi.org/10.1016/j.ijhm.2019.01.005>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, 20, 709-734. <https://doi.org/10.2307/258792>
- McDonnell, M., & Baxter, D. (2019). Chatbots and Gender Stereotyping. *Interacting with Computers*, 31, 116-121. <https://doi.org/10.1093/iwc/iwz007>
- Moussawi, S., Koufaris, M., & Benbunan-Fich, R. (2021). How Perceptions of Intelligence and Anthropomorphism Affect Adoption of Personal Intelligent Agents. *Electronic Markets*, 31, 343-364. <https://doi.org/10.1007/s12525-020-00411-w>
- Mull, I., Wyss, J., Moon, E., & Lee, S. (2015). An Exploratory Study of Using 3D Avatars as Online Salespeople. *Journal of Fashion Marketing and Management*, 19, 154-168. <https://doi.org/10.1108/jfmm-05-2014-0033>
- Nass, C., & Steuer, J. (1993). Voices, Boxes, and Sources of Messages. *Human Communication Research*, 19, 504-527. <https://doi.org/10.1111/j.1468-2958.1993.tb00311.x>
- Nault, K. A., Pitesa, M., & Thau, S. (2020). The Attractiveness Advantage at Work: A Cross-Disciplinary Integrative Review. *Academy of Management Annals*, 14, 1103-1139. <https://doi.org/10.5465/annals.2018.0134>
- Nowak, K. L., & Rauh, C. (2005). The Influence of the Avatar on Online Perceptions of Anthropomorphism, Androgyny, Credibility, Homophily, and Attraction. *Journal of Computer-Mediated Communication*, 11, 153-178. <https://doi.org/10.1111/j.1083-6101.2006.tb00308.x>
- Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., & Patton, M. W. (2011). Embodied Conversational Agent-Based Kiosk for Automated Interviewing. *Journal of Management Information Systems*, 28, 17-48. <https://doi.org/10.2753/mis0742-1222280102>
- Pitardi, V., Bartikowski, B., Osburg, V., & Yoganathan, V. (2023). Effects of Gender Congruity in Human-Robot Service Interactions: The Moderating Role of Masculinity. *International Journal of Information Management*, 70, Article 102489. <https://doi.org/10.1016/j.ijinfomgt.2022.102489>
- Prebensen, N. K., & Rosengren, S. (2016). Experience Value as a Function of Hedonic and Utilitarian Dominant Services.

- International Journal of Contemporary Hospitality Management*, 28, 113-135. <https://doi.org/10.1108/ijchm-02-2014-0073>
- Qadir, J. (2023). Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE. <https://doi.org/10.1109/educn54358.2023.10125121>
- Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People*. Cambridge University Press.
- Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic Review: Trust-Building Factors and Implications for Conversational Agent Design. *International Journal of Human-Computer Interaction*, 37, 81-96. <https://doi.org/10.1080/10447318.2020.1807710>
- Rhodes, G. (2006). The Evolutionary Psychology of Facial Beauty. *Annual Review of Psychology*, 57, 199-226. <https://doi.org/10.1146/annurev.psych.57.102904.190208>
- Salminen, J., Jung, S., M. Santos, J., Mohamed Sayed Kamel, A., & J. Jansen, B. (2021). Picturing It: The Effect of Image Styles on User Perceptions of Personas. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). ACM. <https://doi.org/10.1145/3411764.3445360>
- Stoller, R. J. (2020). *Sex and Gender: The Development of Masculinity and Femininity*. <https://doi.org/10.4324/9780429479915>
- Tamkin, A., Brundage, M., Clark, J., & Ganguli, D. (2021). *Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models*. <https://doi.org/10.48550/arXiv.2102.02503>
- Tay, B., Jung, Y., & Park, T. (2014). When Stereotypes Meet Robots: The Double-Edge Sword of Robot Gender and Personality in Human-robot Interaction. *Computers in Human Behavior*, 38, 75-84. <https://doi.org/10.1016/j.chb.2014.05.014>
- Toure-Tillery, M., & McGill, A. L. (2015). Who or What to Believe: Trust and the Differential Persuasiveness of Human and Anthropomorphized Messengers. *Journal of Marketing*, 79, 94-110.
- Tussyadiah, I. P., & Park, S. (2018). Consumer Evaluation of Hotel Service Robots. In *Information and Communication Technologies in Tourism 2018* (pp. 308-320). Springer. https://doi.org/10.1007/978-3-319-72923-7_24
- Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *The Canadian Journal of Psychiatry*, 64, 456-464. <https://doi.org/10.1177/0706743719828977>
- Yen, C., & Chiang, M. (2021). Trust Me, If You Can: A Study on the Factors That Influence Consumers' Purchase Intention Triggered by Chatbots Based on Brain Image Evidence and Self-Reported Assessments. *Behaviour & Information Technology*, 40, 1177-1194. <https://doi.org/10.1080/0144929x.2020.1743362>
- Yuksel, B. F., Collisson, P., & Czerwinski, M. (2017). Brains or Beauty. *ACM Transactions on Internet Technology*, 17, 1-20. <https://doi.org/10.1145/2998572>
- Zamora, J. (2017). I'm Sorry, Dave, I'm Afraid I Can't Do That. In *Proceedings of the 5th International Conference on Human Agent Interaction* (pp. 253-260). ACM. <https://doi.org/10.1145/3125739.3125766>
- Zhang, Y. (2019). Representative Bureaucracy, Gender Congruence, and Student Performance in China. *International Public Management Journal*, 22, 321-342. <https://doi.org/10.1080/10967494.2018.1428702>
- Zogaj, A., Mähner, P. M., Yang, L., & Tscheulin, D. K. (2023). It's a Match! The Effects of Chatbot Anthropomorphization and Chatbot Gender on Consumer Behavior. *Journal of Business Research*, 155, Article 113412. <https://doi.org/10.1016/j.jbusres.2022.113412>