

当“机器人”成为社会行动者： 人机交互关系中的刻板印象

申琦 王璐瑜

内容提要 人与社交机器人的互动模拟了人际交往的过程，了解人们如何看待社交机器人并与之展开交往互动十分重要。已有社交机器人刻板印象的研究运用传播学经典的媒体等同理论与计算机是社会行动者范式，发现人们会将社交机器人视为社会行动者，并依据人际交往中的刻板印象社会规则对其形成相应的情感态度和行为意愿，由此产生了两类研究取向：一是社交机器人如何顺应刻板印象以更好地服务人类；二是如何通过社交机器人的反刻板印象设计调节人类社会的偏见。爬梳已有研究，可为从媒体等同理论视角考察社交机器人社会行动者角色提供理论与方法上的借鉴，并为探索和反思人机交互关系提供启示。

关键词 社交机器人 媒体等同理论 刻板印象 社会偏见

随着人工智能技术的发展，社交机器人（social robot，如聊天机器人、家庭照料机器人，养老机器人等）在信息传播、日常陪伴和情感交流等方面发挥着越来越重要的作用，引发了我国新闻传播学者对其传播行为、传播效果与传播伦理等问题的深入探讨^①。然而，对于人们如何看待社交机器人并与之展开交往互动的研究较少。

社交机器人延伸了人的“主体意识”和“思维”，在人类传播中不再是简单的媒体“中介”，而是能够在人类的物理空间和符号空间自主参与社交互动的主体。“人+社交机器人”的传播格局已成为一种可想象的社会图景^②。人与社交机器人之间的互动交往，成为人工智能与传播学领域关注的重点问题。近20年来，依据传播学经典的媒体等同理论（media equation theory）与计算机是社会行动者范式

① 师文、陈昌凤 《社交机器人在新闻扩散中的角色和行为模式研究——基于〈纽约时报〉“修例”风波报道在 Twitter 上扩散的分析》，《新闻与传播研究》2020 年第 5 期；张洪忠、石韦颖、韩晓乔 《从传播方式到形态：人工智能对传播渠道内涵的改变》，《中国记者》2018 年第 3 期。

② 张洪忠、段泽宁、韩秀 《异类还是共生：社交媒体中的社交机器人研究路径探讨》，《新闻界》2019 年第 2 期；Skewes J, Amodio D M, Seibt J, “Social Robotics and the Modulation of Social Perception and Bias,” *Philosophical Transactions of The Royal Society B*, vol. 374, 2019.

(computers are social actors, 以下简称 casa), 社交机器人刻板印象 (stereotype) 研究验证了人机交互中人们会根据社交机器人的社会化线索 (如外观特征, 社会赋予的标签分类等), 产生人类社会固有的刻板印象与相应的态度和行为反应, 从而影响到双方的信息交往行为与互动关系。研究认为: 社交机器人的社会化线索与其传播的内容、提供的服务具有同样重要意义, 值得重视; 并且, 考察人机交互中人们如何看待、对待社交机器人能够帮助我们理解人类本身, 更好地定义人与人、人与机器人之间的社会关系。目前已有不少西方传播学者对此展开研究。鉴于此, 本文将在深入分析人机交互领域社交机器人刻板印象已有研究成果的基础上, 从主要解决的问题、使用的理论与方法、存在的不足等方面展开讨论, 为我国相关研究提供理论依据和经验参考。

一、社交机器人: 媒体等同理论构建下的社会行动者

为了理解人们如何看待媒体技术并与之互动, 斯坦福大学传播学系学者李维斯 (Reeves) 和纳斯 (Nass) 于 20 世纪 90 年代提出媒体等同理论^①, 认为“媒体等同于现实生活”、“人们会对媒体产生社会的、自然的反应, 即便他们知道这样做不合理, 甚至认为不可能做出这样的反应”^②。这一大胆的观点在当时引起了强烈反响, 学界和业界认为它“为媒体研究建立了一个精巧的理论框架”、“对媒体产品设计提出了全新的设计原则和评价标准”^③, 成为此后人与计算机、社交机器人交互研究的重要理论依据。学者们想要发现, 当面对并使用上述媒体时, 人们是否会像对待真实人类那样展开交流, 进而产生一定的社会关系。

媒体等同理论的重要研究范式 CASA^④ 认为, 人们会将计算机、电视等媒体视为社会中的人, 根据它们的社会化线索 (如, 输出的文本、语音、语言风格, 可以与人互动等), 形成一定的社会规则 (social rules, 如刻板印象、以礼相待、互惠交往等), 对其产生社会化反应 (social responses, 如信任, 喜爱等)。如, 用户和计算机对话时, 当计算机像人类一样不断改变话题并表现出较高的互惠性时, 用户更愿意自我披露^⑤。经过一系列的 CASA 实验, 纳斯等人把人们将计算机视为社会行动者的

① Reeves, B., & Nass, C., *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*, New York: Cambridge University Press, 1996.

② (美) 巴伦·李维斯、克利夫·纳斯 《媒体等同: 人们该如何像对待真人实景一样对待电脑、电视和新媒体》, 卢大川等译, 上海: 复旦大学出版社, 2001年, 第6页。

③ (美) 巴伦·李维斯、克利夫·纳斯 《媒体等同: 人们该如何像对待真人实景一样对待电脑、电视和新媒体》, 卢大川等译, 上海: 复旦大学出版社, 2001年, 第3页。

④ Nass, C., Steuer, J., & Tauber, E. R., “Computers are Social Actors,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1994, pp. 72–78.

⑤ Moon, Y., “Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers,” *Journal of Consumer Research*, vol. 26, 2000, pp. 323–339.

现象解释为一种无意识的行为 (mindless behavior)^①，指出，现代媒体通常表现出具有人性意义的社会化线索，人们往往会直觉地关注此类线索，忽略了媒体呈现的工具性线索，从而自然地将媒体视为人一样的社会行动者，并做出社会化反应。如，当计算机输出女性声音时，受试者会下意识地认为自己是在和女性交流，快速启动现实中对女性的交往规则对待它。

根据 CASA 范式，人们会依据人际交往中的社会规则与计算机互动。其中，刻板印象这一最为常见的人际交往社会规则，受到了研究者的关注。通过一系列计算机性别刻板印象研究，纳斯等人发现人们会根据计算机的声音特性来划分性别，并且人与计算机互动中形成的刻板印象社会规则可以影响人对计算机的情感态度^②。如人们认为男性声音的计算机在讲解“计算机专业知识”类话题时更具说服力，而女性声音的计算机在“爱与人际关系”等感性话题上更胜一筹^③。进一步地，学者们追问，作为社会行动者，“媒体需要（迎合用户的期盼）让刻板印象永恒吗？还是（作为社会行为者的）媒体可以在改变刻板印象的过程中起到积极的作用？”^④

媒体等同理论在媒介技术决定论和媒介环境论之外，用经验主义的方法，通过思考人和媒体之间的关系，为我们重新思考人类社会的关系做出了指引。其使用的社会心理学理论与研究方法也为后续研究提供了重要支持。然而，媒体等同理论“侧重于人的感性需要和媒介效果的微观影响”^⑤，从早期人与计算机互动到今天人与社交机器人互动，大部分研究多是检验了该理论在不同应用场景中的解释力，未能对理论本身拓展和增益。

近年来随着人工智能技术的勃兴与应用，人与社交机器人之间的互动成为研究热点。社交机器人是一种“拥有社会智能” (socially intelligence)， “生存和嵌入于人类社会的” (socially situated & socially embedded)，能够引发人类的情绪感受 (socially evocative)，并能和人类社交 (sociable) 的自动化、拟人化机器^⑥。根据媒体等同理论与 CASA 范式，与早期研究中的计算机、电视等媒体相比，社交机器人的社会化线索更加丰富复杂，无疑是更积极活跃的社会行动者。那么，与社交机器

① Nass, C., & Moon, Y., “Machines and Mindlessness,” *Journal of Social Issues*, vol. 56, no. 1, 2000, pp. 81 - 103; Langer, E. J., *Mindfulness*, MA: Addison-Wesley, 1989 [美] 巴伦·李维斯、克利夫·纳斯 《媒体等同：人们该如何像对待真人实景一样对待电脑、电视和新媒体》，卢大川等译，上海：复旦大学出版社，2001年，第6-7、21页。

② Lee E. J., Nass C., Brave S., “Can Computer-Generated Speech Have Gender? An Experimental Test of Gender Stereotype,” in *CHI’ 00 Extended Abstracts on Human Factors in Computing Systems*, 2000, pp. 289 - 290.

③ Nass C., Moon Y., & Green N., “Are Computers Gender-Neutral? Gender Stereotypic Responses to Computers,” *Journal of Applied Social Psychology*, vol. 27, 1997, pp. 864 - 876.

④ (美) 巴伦·李维斯、克利夫·纳斯 《媒体等同：人们该如何像对待真人实景一样对待电脑、电视和新媒体》，卢大川等译，上海：复旦大学出版社，2001年，第142页。

⑤ 梅琼林、张晓 《“媒体等同”——从效果研究到理论建构》，《社会科学研究》2006年第5期。

⑥ Fong, T., Nourbakhsh, I. & Dautenhahn, K., “A Survey of Socially Interactive Robots,” *Robotics and Autonomous Systems*, vol. 42, no. 3 - 4, 2003, pp. 143 - 166.

人的互动是否也会启发人们的刻板印象?在这一过程中,社交机器人的社会行动者角色为何,作用如何?大量的社交机器人刻板印象研究尝试对此做出回应。

二、社交机器人刻板印象:人机交互中的社会规则

(一) 刻板印象、刻板印象内容模型与偏差地图

心理学中的刻板印象研究经历了从早期的概念界定,到刻板印象的分类与成因,以及将刻板印象与群际情绪、行为反应相结合的过程,且仍在不断发展中。刻板印象是人们对某一群体成员相对固定的观念或期望的基础上,形成的某种特定认知结构和社会认知图式,对人们的认知和行为有着重要指导作用^①。

早期的刻板印象研究主要从心理动力学和社会文化视角,从微观(如个体生活经验)和宏观(如社会背景、群际关系)两个层面关注刻板印象的生成和保持^②。如卡茨(Katz)和布拉利(Braly)在20世纪30年代普林斯顿大学开展的一系列有关大学生的刻板印象研究。研究要求被试大学生用给定的84个有关心理特质的形容词(如聪明的、警觉等)描绘10个种族或民族的群体特征;发现,白人学生会对不同国籍的人产生类型化评价,比如,德国人都有科学头脑,黑人都很迷信。据此,得出在某一群体中出现最多的特质,就是人们对该群体的刻板印象这一结论^③。这一时期的研究偏重于种族(国家)领域的刻板印象。

20世纪70年代后,认知心理学的大规模介入促使研究转向对刻板印象成因和影响的探索。如,刻板印象如何影响不同群体之间的交往,预测群体或群体中个人对特定群体的态度和行为倾向。研究多将刻板印象分为性别刻板印象、年龄刻板印象、种族刻板印象等^④。然而,这些以性别、年龄、种族、地域等个体或群体被动接受的外在表征为分类标准的研究,忽略了群际交往过程中人的主动性,容易将刻板印象与社会偏见等负面情感直接关联。

20世纪90年代起,心理学家菲斯克(Fiske)陆续提出了刻板印象内容模型(stereotype content model,以下简称scm)和偏差地图(behaviors from intergroup affect and stereotypes map,以下简称bias map)(见图1),不仅从人的能动性方面深入考察了刻板印象的形成过程,还将刻板印象与群际情绪、行为反应倾向结合^⑤。SCM认为人们出于进化的压力必须快速判断他者意图的好坏,以及他者根据意图采取行动的

① Fiske, S. T., "Stereotyping, Prejudice, and Discrimination," in D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of Social Psychology* (4th ed., vol. 2), Boston: McGraw-Hill, 1998, pp. 357-411.

② Allport, G. W., "The Nature of Prejudice," *The Journal of Negro History*, vol. 6, no. 3, 1954.

③ Katz, D., & Braly, K., "Racial Stereotypes of One Hundred College Students," *The Journal of Abnormal and Social Psychology*, vol. 28, no. 3, 1933, p. 280.

④ Fiske, S. T., "Stereotyping, Prejudice, and Discrimination," in D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of Social Psychology* (4th ed., vol. 2), Boston: McGraw-Hill, 1998, pp. 357-411.

⑤ 管健《刻板印象从内容模型到系统模型的发展与应用》,《心理科学进展》2009年第4期。

能力。1999年，菲斯克首次提出感知温暖（warmth）和能力（competence）两个维度，认为感知温暖指的是一个人是否被认为有积极的意图，并倾向照顾他人的利益^①。在温暖维度中，友善、乐于助人、真诚、可信和道德等都有助于人们感知到他者积极的意图；感知能力是指一个人是否有与实现意图相关的能力特质，可以通过智力、技能、创造力、效能和独立性来体现^②。如，人们通常会喜欢感觉温暖的人，如慈善的年老耆；会更尊敬有能力的人，如权威专家。并且在第一印象中存在“温暖优先效应”，温暖感知通常优先于能力判断，这对人们的态度和情感产生很大影响。进一步地，SCM在温暖和能力的基础上又细分了四个象限，形成了一个2×2的双维四象限坐标体系，将社会群体按照一定偏见类型归为高温暖-高能力、低温暖-高能力、高温暖-低能力和低温暖-低能力群体，分别对应着人们由此产生的钦佩、嫉妒、同情和蔑视等情感。

为了解释刻板印象与人们的情感和行为之间的关系，菲斯克等人在SCM的基础上又提出了BIAS Map^③，将温暖、能力与人的情绪、行为反应倾向结合。根据人的能力和温暖高低引发的行为反应，主要包括主动助长（active facilitation）（如，帮助与保护）、主动伤害（active harm）（如，攻击与反抗）、被动助长（passive facilitation）（如，合作与关联）与被动伤害（passive harm）（如，忽略与漠视）等。BIAS Map与SCM一样，假定温暖是刻板印象的首要维度，如果他人或他群是温暖的，则会得到主动助长（如，帮助），否则会受到主动伤害（如，攻击）。如果他人或他群是高能力的会引发被动助长（如，合作），否则引发被动伤害（如，忽视）。

SCM和BIAS Map从社会关系角度丰富了人们对刻板印象的理解，也对干预和调节社会偏见有着重要意义。模型不仅打破了单一的群体内与群体外的划分取向，用不以性别、种族等这些外在表征为基础形成的温暖和能力维度重新建构群体的社会认知，还充分考虑到了社会因素，用社会地位和竞争性来代表群体社会处境，考察了社会关系对刻板印象形成的影响^④。

在人机交互中，社交机器人刻板印象研究试图拓展CASA范式，一方面通过考察人们对社交机器人社会化线索（如外貌、性别、声音等）的反应，了解对社交机器人的分类。另一方面，研究结合SCM和BIAS Map，深入考察了人们如何通过社交机器人的社会化线索（如，感知温暖和能力等）识别他们的意图，产生一定的情感

① Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P., “(Dis) Respecting Versus (Dis) Liking: Status and Interdependence Predict Ambivalent Stereotypes of Competence and Warmth,” *Journal of Social Issues*, vol. 55, no. 3, 1999, pp. 473–489.

② Fiske, S. T., Cuddy, A. J., & Glick, P., “Universal Dimensions of Social Cognition: Warmth and Competence,” *Trends in Cognitive Sciences*, vol. 11, 2007, pp. 77–83.

③ Cuddy, A. J. C., Fiske, S. T., & Glick, P., “The BIAS Map: Behaviors from Intergroup Affect and Stereotypes,” *Journal of Personality and Social Psychology*, vol. 92, no. 4, 2007, pp. 631–648.

④ 汪新建、程婕婷、管健 《解析群际偏见——基于刻板印象内容模型的认知神经研究》，《广东社会科学》2014年第3期。



图1 SCM与BIAS Map示意图^①

和行为，进而形成某种社会化反应与社会关系^②。这为从人的心理认知角度探索人机共生环境下，人类如何通过刻板印象社会规则与机器人交往互动，以及这种交互过程又会如何作用于人类社会本身提供了研究思路。

(二) 基于社会化线索的社交机器人刻板印象分类

大量的研究发现，社交机器人社会化线索可以直接引发人们的无意识反应，对社交机器人进行快速分类。

人们首先会对社交机器人进行简单的性别、种族等分类。如，从外观上，人们会认为长头发机器人为女性^③。当面对两个一模一样的社交机器人时，人们会想当然地认为名叫“Nero”的机器人是男性，而“Nera”是女性^④。社交机器人的外观及“肤色”也会引发人们对其进行种族分类。不管是观看照片还是实际使用中，人们会

① 管健 《刻板印象从内容模型到系统模型的发展与应用》，《心理科学进展》2009年第4期。

② Ćaić, M., Avelino, J., Mahr, D., & Bernardino, A., “Robotic Versus Human Coaches for Active Aging: an Automated Social Presence Perspective,” *International Journal of Social Robotics*, vol. 6, 2019, pp. 1 - 16; Oliveira, R., Arriaga, P., Correia, F., & Paiva, A., “The Stereotype Content Model Applied to Human-Robot Interactions in Groups,” in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 123 - 132; Mieczkowski, H., Liu, S. X., Hancock, J. & Reeves, B., “Helping Not Hurting: Applying the Stereotype Content Model and BIAS Map to Social Robotics,” in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 222 - 229.

③ Eyssel, F. & Hegel, F., “(S) he’s Got the Look: Gender Stereotyping of Robots,” *Journal of Applied Social Psychology*, vol. 42, no. 9, 2012, pp. 2213 - 2230.

④ Reich-Stiebert, Natalia, and Eyssel, F., “(It) Relevance of Gender? on the Influence of Gender Stereotypes on Learning with a Robot,” in *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, IEEE, 2017, pp. 166 - 176.

自然地将黑色社交机器人归为黑色人种，白色的为白种人^①。

社交机器人的社会化线索还能激发人们对其进行人格、性格等更深层次的分类。如，人们会认为男性机器人更健壮有力，女性更善交际^②。拥有高音调、宽音域的社交机器人更加外向^③。相较于话语风格直接的社交机器人，话语风格较委婉的通常被认为更具女性气质^④。也有研究发现，外观拟人化程度较高的社交机器人（如仿人机器人）比拟人化程度不高的（如有机械臂的机器人）更能让人形成温暖和有能力印象。不仅是外观，社交机器人的产地也能让人们形成温暖或者有能力的印象^⑤。社交机器人的外观、声音、肤色等社会化线索能够激发人们的刻板印象，对其进行归类。这是对已有 CASA 范式的进一步验证。然而，社交机器人相比计算机拥有更多的社会化线索（如表情、体形、手势、行走步态等）^⑥，它们如何激活人类的刻板印象分类，目前的研究尚且不足。并且来自外观、声音、表情、行动风格等各层次的社会化线索相互交错，研究有待深入考察线索的主次、类型等对社交机器人刻板印象分类的不同影响。

（三）基于刻板印象对社交机器人的情感和行为反应

分类之后，人们也会对社交机器人产生与人际交往中相似的情感和行为反应。研究发现，人际交往中的种族偏见也延续到人机交互中。当人们将社交机器人作为被射击对象时，射击黑色社交机器人的速度和准确率都高于白色^⑦。人类社会中性别与职业匹配产生的刻板印象也存在于人机交互中。人们会更欢迎女性社交机器人从事医疗护理工作，男性机器人从事安全保护工作^⑧。而拥有男性声音的社交机器人作

-
- ① Ogunyale, T., Bryant, D. A. & Howard, A., "Does Removing Stereotype Priming Remove Bias? A Pilot Human-Robot Interaction Study," *arXiv preprint arXiv: 1807.00948*, 2018; Bartneck, C., Yogeewaran, K., Ser, Q. M., Woodward, G., Sparrow, R., Wang, S., & Eyssel, F., "Robots and Racism," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 196 - 204.
- ② Eyssel, F. & Hegel, F., "(S) he's Got the Look: Gender Stereotyping of Robots," *Journal of Applied Social Psychology*, vol. 42, no. 9, 2012, pp. 2213 - 2230.
- ③ Chang, R. C. S., Lu, H. P., & Yang, P., "Stereotypes or Golden Rules? Exploring Likable Voice Traits of Social Robots as Active Aging Companions for Tech-Savvy Baby Boomers in Taiwan," *Computers in Human Behavior*, vol. 84, 2018, pp. 194 - 210.
- ④ Kraus, M., Kraus, J., Baumann, M., & Minker, W., "Effects of Gender Stereotypes on Trust and Likability in Spoken Human-Robot Interaction," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, 2018.
- ⑤ Spatola, N., Anier, N., Redersdorff, S., Ferrand, L., Belletier, C., Normand, A., & Huguet, P., "National Stereotypes and Robots' Perception: the 'Made in' Effect," *Frontiers in Robotics and AI*, vol. 6, 2019.
- ⑥ Hwang, J., Park, T., & Hwang, W., "The Effects of Overall Robot Shape on the Emotions Invoked in Users and the Perceived Personalities of Robot," *Applied Ergonomics*, vol. 44, no. 3, 2013, pp. 459 - 471; Craenen, B. G., Deshmukh, A., Foster, M. E., & Vinciarelli, A., "Shaping Gestures to Shape Personality: Big-Five Traits, Godspeed Scores and the Similarity-Attraction Effect," in *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, 2018, pp. 2221 - 2223.
- ⑦ Bartneck, C., Yogeewaran, K., Ser, Q. M., Woodward, G., Sparrow, R., Wang, S., & Eyssel, F., "Robots and Racism," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 196 - 204.
- ⑧ Tay, B., Jung, Y. & Park, T., "When Stereotypes Meet Robots: the Double-Edge Sword of Robot Gender and Personality in Human-Robot Interaction," *Computers in Human Behavior*, vol. 38, 2014, pp. 75 - 84.

为教练讲解汽车知识时，会让人觉得更专业和更值得信任^①。

为进一步了解刻板印象对人机交互社会规则的影响，大量的研究尝试运用 SCM 和 BIAS Map 考察人们对社交机器人温暖和能力的感知，会引发什么样的情感反应和行为倾向。根据 SCM，有研究分析了人们对社交机器人温暖和能力的感知，与情绪以及使用意愿之间的关系。研究设计了一组将社交机器人作为对手或合作者的四人纸牌游戏，发现人们更钦佩温暖程度高的社交机器人，而与能力高低无关；更轻视低温暖-低能力组合的社交机器人；相比高温暖-低能力的机器人，人们更倾向同情高温暖-高能力的机器人。总体而言，人们更喜欢感觉温暖的社交机器人，并且在温暖的基础上，他们才会选择高能力的社交机器人一起完成纸牌游戏^②。这与人际交往中，人们首先考虑对方的温暖特质，更喜欢与感觉温暖的人合作的情况一致。

更进一步地，有研究运用 BIAS Map 考察了通过对社交机器人外观刻板印象的感知，人们会产生什么样的情感和行为反应。在让受试者观看了 324 幅社交机器人的图像后，研究发现，外观感知高温暖-高能力的社交机器人更受被试者钦佩，高温暖-低能力者更受同情；钦佩情感更容易让人产生对社交机器人帮助（无论是主动还是被动帮助）的行为意愿，蔑视情感则带来伤害的行为意愿（无论是主动还是被动伤害），嫉妒情感带来主动伤害和被动帮助的行为意愿，而同情带来消极伤害和积极帮助的行为意愿^③。研究证实了运用 BIAS Map 观察人们通过刻板印象对社交机器人产生情感和行为反应的可行性。并且由于传统 CASA 范式专注社会化线索带来的刻板印象，缺少对于社会化线索引发人们的情感态度的解释力，上述研究也为进一步丰富 CASA 范式，为从情感角度细致探索人机交互社会规则提供了指导。但是，正如研究者自己所言，验证人与社交机器人之间的社会规则，不能仅凭直觉的刻板印象，交往的动机、以往的行为经验动机等因素也需纳入考虑。

三、刻板印象之后：社交机器人作为多角色的社会行动者

媒体等同理论通过考察人们对媒体的态度，提出符合人际交往社会规则的媒体设计最受欢迎，也留下了“挑战性别刻板印象的媒体设计或许可以调节社会偏见问题”^④的思索。社交机器人刻板印象研究，一方面结合了多样的应用场景、用户人

① Johnson, D., & Gardner, J., "Exploring Mindlessness as an Explanation for the Media Equation: a Study of Stereotyping in Computer Tutorials," *Personal and Ubiquitous Computing*, vol. 13, no. 2, 2009, pp. 151 - 163.

② Oliveira, R., Arriaga, P., Correia, F., & Paiva, A., "The Stereotype Content Model Applied to Human-Robot Interactions in Groups," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 123 - 132.

③ Mieczkowski, H., Liu, S. X., Hancock, J. & Reeves, B., "Helping Not Hurting: Applying the Stereotype Content Model and BIAS Map to Social Robotics," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 222 - 229.

④ Lee E. J., Nass C., Brave S., "Can Computer-Generated Speech Have Gender? An Experimental Test of Gender Stereotype," in *CHI'00 Extended Abstracts on Human Factors in Computing Systems*, 2000, pp. 289 - 290.

群，通过顺应人类的刻板印象完善社交机器人设计；另一方面也试图探索社交机器人这一更加活跃的社会行动者在调节社会偏见，促进社会公平方面的可能性。

（一）服务人类的社会行动者

迄今不少的人机交互刻板印象研究都是为了了解人机互动中的用户感受，进一步改进社交机器人的外观和功能设计：一是，从社交机器人的外观、声音、功能等设计上满足人们的需要。如，研究通过收集人类表演中的动作集，将更受欢迎的动作用于机器人设计^①。有研究则通过人机互动，不断让社交机器人学习，为算法增加新的刻板印象模型，提升社交机器人的社交互动行为。让机器人反复与人对话，训练机器人的反应，增加不同的刻板印象模型，完善其交互式问答的功能^②。二是，考察刻板印象与其他心理因素结合以提高人们的使用意愿。如，利用人际吸引理论的相似吸引、互补吸引、能力吸引等规则，研究模拟了一个在同伴型和专家型机器人教师教导下的学习场景，结果发现，同伴型的机器人教师能让学生更快乐地学习，而专家型的机器人教师让学生的信任感更高^③。

可见，正如媒体等同理论为计算机设计带来的灵感一样^④，社交机器人刻板印象研究从用户感官、心理等角度为提升社交机器人的外观与功能设计，使其更好地满足用户需求提供了现实指导。这些研究，有的从引发人们刻板印象的社会化线索中找到思路，有的从触发人们社会化反应的不同应用场景中获取灵感。值得思索的是，尽管我们将用户需求更多地考虑进了社交机器人的设计中，但是如果大量的参试者都是西方研究者选定的，未来在社交机器人设计领域会不会产生种族、地域、文化甚至意识形态的分歧与鸿沟？

（二）调节社会偏见的社会行动者

社交机器人需要（迎合用户的期盼）让刻板印象永恒，还是可以在改变刻板印象的过程中起到积极的作用？已有研究指出，社交机器人中的刻板印象很大程度上是人类社会偏见的镜像，往往以不为人注意的方式影响到社交机器人的设计，导致其“复制现有的歧视模式，继承先前决策者的偏见”^⑤，加剧社会不公。有研究注意

① Randhavane, T., Bera, A., Kapsaskis, K., Gray, K., & Manocha, D., "FVA: Modeling Perceived Friendliness of Virtual Agents Using Movement Characteristics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 11, 2019, pp. 3135 - 3145.

② Wagner, A. R., "Using Cluster-Based Stereotyping to Foster Human-Robot Cooperation," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 1615 - 1622.

③ Liew, T. W., Su-Mae, T. & Jayothisa, C., "The Effects of Peer-Like and Expert-Like Pedagogical Agents on Learners? Agent Perceptions, Task-Related Attitudes, and Learning Achievement," *Journal of Educational Technology & Society*, vol. 16, no. 4, 2013, p. 275.

④ (美) 巴伦·李维斯、克利夫·纳斯 《媒体等同：人们该如何像对待真人实景一样对待电脑、电视和新媒体》，卢大川等译，上海：复旦大学出版社，2001年，第2页。

⑤ Federal Trade Commission., "Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues," 2016, <https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf>, 2020. 09. 10.

到,养老社交机器人简单将老年人定义为病患或者残疾人,而这与人类对老年人群体的既有偏见相关^①。研究批评,这种做法忽视了老年人的个体经历、社会关系以及基本人权,将老年人面临的问题从社会关注降级、窄化到机器人可以处理的事务。如,一项针对老年人的访谈显示,在观看养老机器人工作的视频后,老年人对其工作能力评价积极,却反感养老机器人将自己视为残疾失能的老人,这也是他们不愿意接受机器人的主要原因^②。

认识到对社交机器人的刻板印象可能会加剧社会偏见,有研究尝试利用反刻板印象(counter-stereotype)设计来调节社会偏见。反刻板印象是指表现与常规刻板印象不一致的心理认知,如男性优柔寡断,女性强硬理性等^③。在一项实验中,男女学生参加了为期七周由社交机器人教授的工程课程(常规刻板印象认为男生更爱好和擅长工程学科),结果发现,男性机器人教导下的男生在课程任务上表现更优秀,而女性机器人教导下的男生和女生在课程任务的表现上没有显著差异。研究认为这可能是因为女性机器人为女生展示了积极的榜样作用^④。

除了反刻板印象的心理学实验,有研究还从社交机器人设计的伦理规范、选用规范等方面展开探索^⑤。如,呼吁社交机器人的设计和选用应充分鼓励公众参与,尤其是社会弱势群体参与。老年人可能由于残疾而需要“帮助”,但同样也需要维持他们作为独立人的自主和尊严。因此,应将社交机器人纳入如何帮助老人更加自主地生活,削弱社会对老年群体“衰老”的印象,促进积极老龄化等。研究指出,应有更多的学者进入这些领域,帮助企业 and 设计者完善社交机器人的设计规范,形成专业标准^⑥。

作为社会行动者,社交机器人将越来越多地进入人类社会生活。无论是复制人际交往中的刻板印象,还是通过反刻板印象调节人类固有偏见,其中都会嵌入设计者的价值观。那么,设计者的价值观通过社交机器人,人机交互活动又将如何影响

① Neven. L., "But Obviously Not for Me: Robots, Laboratories and the Defiant Identity of Elder Test Users," *Sociology of Health & Illness*, vol. 32, no. 2, 2010, pp. 335 - 347; Sparrow, R. & Sparrow, L., "In the Hands of Machines? The Future of Aged Care," *Minds and Machines*, vol. 16, no. 2, 2006, pp. 141 - 161; Vines, J., Pritchard, G., Wright, P., Olivier, P. & Brittain, K., "An Age-Old Problem: Examining the Discourses of Ageing in HCI and Strategies for Future Research," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 22, no. 1, 2015, pp. 1 - 27.

② Lee, H. R., Tan, H. & Šabanović, S., "That Robot Is Not for Me: Addressing Stereotypes of Aging in Assistive Robot Design," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN)*, 2016, pp. 312 - 317.

③ Plaks J E, Stroessner S J, Dweck C S, Sherman J W, et al., "Person Theories and Attention Allocation: Preferences for Stereotypic Versus Counterstereotypic Information," *Journal of Personality and Social Psychology*, vol. 80, no. 6, 2001, pp. 876 - 893.

④ Sullivan, A. A., "Breaking the STEM Stereotype: Investigating the Use of Robotics to Change Young Children's Gender Stereotypes About Technology and Engineering," Tufts University, 2016.

⑤ Howard, A. & Borenstein, J., "The Ugly Truth about Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity," *Science and Engineering Ethics*, vol. 24, no. 5, 2018, pp. 1521 - 1536; Borenstein, J., Howard, A., & Wagner, A., "Pediatric Robotics and Ethics: The Robot Is Ready to See You Now but Should It Be Trusted?" in P. Lin, K. Abney, & G. Bekey (Eds.), *Robot ethics 2.0*. Oxford: Oxford University Press. 2017.

⑥ Lee, H. R. & Riek, L. D., "Reframing Assistive Robots to Promote Successful Aging," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 7, no. 1, 2018, pp. 1 - 23.

人类的主体性（思维模式、价值观念等），乃至整个社会的公平正义，仅从反刻板印象设计角度思考恐怕还不够。

四、社交机器人刻板印象研究使用的方法

有关社交机器人刻板印象的研究，主要运用心理学实验法探讨了社交机器人的社会化线索与人们刻板印象之间的相关性或者因果关系。这也是人机交互研究的经典方法，能够为完善社交机器人设计提供经验数据，但对长期观察人机交互间的社会规则与互动关系的贡献不够。

（一）实验变量及测量

实验变量包括自变量、因变量和控制变量。其中自变量通常为社交机器人的社会化线索（外观、体形、声音等）以及印象分类。如，研究者利用计算机合成（图像、声音、动态画面）等构建不同特征的社交机器人，或者模拟实体社交机器人。被试可以通过观看图片、视频，聆听声音等作出判断，也可与实体机器人共同参与到一项任务中（一起玩游戏、参与教学等）。这几种方式都是为了通过提供不同的刺激使被试对社交机器人生成不同的印象。

因变量主要包括被试对社交机器人的情感、行为倾向和互动表现。其中，情感是对社交机器人的情绪反应（基于 SCM 的钦佩、嫉妒、同情和蔑视等）、情感偏好（喜欢还是不喜欢等）。行为倾向主要为使用意愿、采纳意愿等。互动表现，主要为与社交机器人互动或者合作时基于 BIAS Map 分类的助长或伤害等行为。在一些教育、说服类场景中人们的学习成绩、态度改变等也是因变量^①。因变量的测量常用已有量表，或者用成绩等数据间接反映^②，也有用心理学内隐态度测量实验来反映被试心中的真实态度等^③。控制变量除了人口统计学变量，还包括技术使用能力和既往的社交机器人接触经验等^④。

（二）实验设计

根据自变量的多少，实验设计可分为单因素设计和多因素设计。其中，只有一个自变量的称为单因素实验设计，有两个及以上自变量的为多因素实验设计。相较

① Siegel M, Breazeal C L, Norton M L, "Persuasive Robotics: the Influence of Robot Gender on Human Behavior," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 2563 - 2568.

② Reich-Stiebert, Natalia, and Eyssel, F., "(Ir) Relevance of Gender? on the Influence of Gender Stereotypes on Learning with a Robot," *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, IEEE, 2017, pp. 166 - 176; Li, J., Kizilec, R., Bailenson, J., & Ju, W., "Social Robots and Virtual Agents as Lecturers for Video Instruction," *Computers in Human Behavior*, vol. 55, 2016, pp. 1222 - 1230.

③ Bartneck, C., Yogeewaran, K., Ser, Q. M., Woodward, G., Sparrow, R., Wang, S., & Eyssel, F., "Robots and Racism," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 196 - 204.

④ Reich-Stiebert, N. & Eyssel, F., "Learning with Educational Companion Robots? Toward Attitudes on Education Robots, Predictors of Attitudes, and Application Potentials for Education Robots," *International Journal of Social Robotics*, vol. 7, no. 5, 2015, pp. 875 - 888.

于单因素实验，多因素实验能够更加真实地模拟复杂的社交机器人使用场景，比较不同影响因素作用的强弱及其交互作用。具体表现为：

在一些较为简单的刻板印象激活实验中常采用单因素设计，以观察社交机器人的某一社会化线索对人们情感态度的影响。如，考察颜色这一单一因素对社交机器人感知影响的实验，253名被试在观看了黑色、黄色和米色三类机器人的图片后做出评价^①。

单因素实验只涉及单一的影响因素，事实上影响人们对社交机器人刻板印象形成、接受意愿等的因素往往比较复杂。因而大部分研究采用多因素实验设计，考察社交机器人任务类型、性格和用户性别、性格等因素的影响。如，有研究为探讨社交机器人的性别、性格以及任务对被试使用意愿的影响^②，采用了2（机器人性别男或女）×2（机器人性格外向或内向）×2（护理任务或安保任务）的设计，共涉及8个场景；将164名被试随机分配到8个场景中，回答对社交机器人的情感评价、认知评价、感知控制和使用意愿等问题。实验首先改变社交机器人的声音频率和姓名以操控其性别，改变机器人的语速和动作幅度操控机器人为外向或内向性格。然后，将不同性别和性格的社交机器人分配到护理和安保两类不同应用场景中。在护理场景中的被试接受机器人测量体温等护理操作，在安保场景中的被试观看机器人响应警报的视频。结果显示，社交机器人的性格和性别不会单一地影响被试反应，而是与应用场景共同作用。人们更容易接受性格×任务匹配（如，内向的机器人从事护理工作），性别×任务匹配（如，女性机器人从事护理工作）的机器人，并且性格×任务匹配的刻板印象比性别×职业匹配的刻板印象影响更大。

除了将社交机器人的社会化线索作为自变量，亦有研究将人类用户的特征（年龄、性格、性别等）作为自变量展开多因素实验研究^③。如，一项探究年轻人和老年人对社交机器人接受意愿的研究中，采用了2（被试：年轻，年老）×2（机器人外观：年轻，年老）×2（任务：操作型，认知型）×2（机器人可靠性：低，高）的混合设计。实验分别招募一组年轻人和一组老年人，邀请他们观看8个视频。这些视频包括年轻社交机器人完成操作型任务（如，换灯泡和成功移动箱子），未完成操作型任务，完成认知型任务（如，垃圾分类和分类清洗衣物），未完成认知型任务；

① Louine, J., May, D. C., Carruth, D. W., Bethel, C. L., Strawderman, L., & Usher, J. M., "Are Black Robots Like Black People? Examining How Negative Stigmas about Race are Applied to Colored Robots," *Sociological Inquiry*, vol. 88, 2018, pp. 626 - 648.

② Tay, B., Jung, Y., & Park, T., "When Stereotypes Meet Robots: the Double-Edge Sword of Robot Gender and Personality in Human-Robot Interaction," *Computers in Human Behavior*, vol. 38, 2014, pp. 75 - 84.

③ Pak, R., Crumley-Branyon, J. J., Visser, E. J. D., & Rovira, E., "Factors that Affect Younger and Older Adults' Causal Attributions of Robot Behaviour," *Ergonomics*, vol. 63, no. 4, 2020, pp. 421 - 439; Branyon, J., & Pak, R., "Investigating Older Adults' Trust, Causal Attributions, and Perception of Capabilities in Robots as a Function of Robot Appearance, Task, and Reliability," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, no. 1, 2015, pp. 1550 - 1554; Nomura, T., & Kinoshita, Y., "Gender Stereotypes in Cultures: Experimental Investigation of a Possibility of Reproduction by Robots in Japan," *International Conference on Culture & Computing*, 2016, pp. 195 - 196.

年老机器人完成操作型任务，未完成操作型任务，完成认知型任务，未完成认知型任务等。结果发现，老年人对机器人的信任更容易受到机器人外观、任务类型的影响，年轻人则多受机器人实际完成任务可靠性的影响，说明老年人受刻板印象认知影响更显著。

实验法之外，也有少量社交机器人刻板印象研究使用访谈法，让受访者更加主动、全面地陈述自己的想法^①。这一方法更适用于实验开始前的探索性研究，或者实验后对结果的完善。如，研究让老人观看社交机器人服务老年人群体的视频，通过半结构化访谈详细询问他们最近的生活经历（如是否孤独）、对衰老的看法（什么是变老）、使用意愿等三个方面的开放式问题^②。结果发现，老年人认为，社交机器人在工作中表现出将老人视为残疾无能群体的刻板印象减少了他们的使用意愿。访谈法便于深入细致地理解特定人群对社交机器人的刻板印象和采纳意愿等，更适用于老年人、青少年等难以理解实验的群体。

五、问题与展望

已有社交机器人刻板印象研究运用媒体等同理论，将社交机器人视为社会行动者，为我们观察人与社交机器人这一新型媒介之间的互动关系提供了重要切口，也为我们思考传播学研究的未来提供了可想象的空间。当人与媒介技术之间不仅是具身的关系，更有可能是它异的关系时^③，传播学需要聚焦除了人和机器如何交流，还有这种交流将会形成何种社会关系，又会怎样反作用于人类社会的本身，比如经典的权力分配、技术可供性背后的政治经济约束因素等^④。有鉴于此，我们认为，未来在人机交互刻板印象研究领域，传播学研究应重视下述问题：

（一）发展媒体等同理论

媒体等同理论自2001年引入我国后，鲜有学者关注。不过，在2020年5月第70届国际传播学年会（ica, international communication association）上，首次介绍该理论的《媒体等同：人们该如何像对待真人实景一样对待电脑、电视和新媒体》^⑤一书获会士图书奖（fellows book award）^⑥。专家认为“这是一本极具开创性的书，

① Carpenter, J., Davis, J. M., Erwin-Stewart, N., Lee, T. R., Bransford, J. D. & Vye, N., "Gender Representation and Humanoid Robots Designed for Domestic Use," *International Journal of Social Robotics*, vol. 1, no. 3, 2009, pp. 261 - 265.

② Lee, H. R., Tan, H., & Šabanović, S., "That Robot Is Not for Me: Addressing Stereotypes of Aging in Assistive Robot Design," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN)*, 2016, pp. 312 - 317.

③ (美) 唐·伊德 《技术与生活世界：从伊甸园到尘世》，韩连庆译，北京：北京大学出版社，2012年。

④ 刘海龙 《传播中的身体问题与传播研究的未来》，《国际新闻界》，2018年第2期。

⑤ Reeves, B., & Nass, C., *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*, New York: Cambridge University Press, 1996.

⑥ 会士图书奖（fellows book award）向所有 ICA 会员开放，以表彰那些对传播学以及社会科学做出重大贡献，并经得起时间考验的书籍。见 ICA: ICA Awards, <https://www.icaqdq.org/page/Awards>, 2020年9月10日。

在出版的24年后仍然熠熠生辉。如果说本书对思考人机交互关系产生了直接影响,那么这种影响仍在不断扩大。最近,它被广泛应用在社交机器人、健康数字游戏、Alexa和Siri等虚拟助手、智能手机、自动驾驶、增强现实、技术辅助合作、假新闻的易感性以及人工智能技术社会影响的方方面面……它将不断对传播学领域产生巨大影响。”^①

媒体等同理论提出的背景是在人与计算机、电视机媒体的交互环境下。今天我们已经逐渐步入人工智能时代,面对更加类人化、智能化的社交机器人与人类建立的社交关系和规则,发展媒体等同理论与CASA范式,增强其解释力十分必要。经典的CASA研究多以声音、外观等作为媒体社会化分类的线索,社交机器人无疑拥有更多的社会化线索,如细微的表情、手势,行走的步态,说话的风格语气,互动反应的速度等都可以刺激人们对其分类与评价。已有学者尝试扩展CASA范式为MASA范式(media are social actors),提出可以进一步明确社会化线索的定义,并区分主要社会化线索与次要社会化线索,以全面地考察人们对媒体的态度^②。运用媒体等同理论展开人机交互研究应注意到这点,对社会化线索、社会规则、社会化反应等概念做更为细致的界定。此外,媒体等同理论认为人们是无意识地将计算机等媒体视为社会行动者,这无疑忽略了人类作为理性行为者的情况。如,有研究发现,当同时面对人类声音和机器人声音时,人们会更偏爱人类的而不是机器人的声音^③。实际上,面对机器人时人们会进行有意识的思维活动,而不会完全将其视作日常交往中的人。因此,人类究竟会如何定义社交机器人的社会行动者身份,如何看待与接受人与社交机器人的差异?这需要进一步发展媒体等同理论,对有意识情况下的人机交互关系展开深入研究。

(二) 发展适合社交机器人刻板印象的研究模型

尽管已有不少研究使用了刻板印象内容模型和偏差地图,但目前为止对这一模型的贡献不大。一是,因为该模型在认知心理学领域提出时间短,仍处于模型建构、检验和初步应用阶段。有关温暖与能力两个维度之间存在负向的“补偿效应”还是正向“晕轮效应”,温暖优先效应假设是否稳定存在,是否还可以纳入其他的基本维度(如道德)等问题尚未解决。二是,有研究发现刻板印象内容模型在人机交互中的不适用性,但尚未能作出可靠解释。如,嫉妒、同情和蔑视等不同情感的得分均

^① ICA: ICA 通讯《祝贺 ICA 奖项获得者》,2020年5月, <https://www.icahdq.org/blogpost/1523657/348560/Congratulations-ICA-Award-Winners>, 2020年9月10日。

^② Xu, K., & Lombard, M., "Media are Social Actors: Expanding the CASA Paradigm in the 21st Century," in *Annual Conference of the International Communication Association*, 2016.

^③ Mitchell, W. J., Ho, C. C., Patel, H. & MacDorman, K. F., "Does Social Desirability Bias Favor Humans? Explicit-Implicit Evaluations of Synthesized Speech Support a New HCI Model of Impression Management," *Computers in Human Behavior*, vol. 27, no. 1, 2011, pp. 402-412.

显著低于中值，这一结果是因研究设计所致，还是社交机器人的特殊性，未能解释^①。三是，社交机器人刻板印象除了温暖和力量之外，是否还有新的评价维度，对人机交互关系的影响如何？如，有研究将人机交互领域的经典量表 Godspeed^② 与 SCM 的温暖/力量量表结合起来，发展出社交机器人社会属性量表（robotic social attributes scale, rosas），发现与机器人交互时的舒适感也是人们对社交机器人的基本判断维度^③。且如前所述，在有意识的思维过程中，人们并不会将社交机器人等媒体视为同类，而是更有可能将其视为异类和“他者”。当机器人与人越相似时，人们反而更加抗拒，比起一个复杂的动态聊天机器人，人们对简单文字机器人的负面情绪更少^④。人们会更喜欢犯错的社交机器人，认为这样的机器人更真实，威胁感更低^⑤。因而，应当充分考虑到对社交机器人的认知是个复杂过程，未来我们可以考虑结合归因理论、技术采纳理论、计划行为理论、社会认同理论等进一步探讨刻板印象在人机交互中的影响与作用。

（三）丰富研究与测量方法

心理学实验法是人机交互领域经典的研究方法，通过实验方法模拟和观察人与计算机、社交机器人互动之间的情绪、行为变化，效果显著。但是也应该看到此方法的不足。如，已有研究多将被试召集到实验室中，以社交机器人的图片、实物等作为简单刺激。事实上，当人们与社交机器人进行短暂接触时，常会因为新奇而表现出不可靠的反应^⑥。同时，实验研究多是个案研究，同一项研究较难为其他学者复制和比较。这与研究中所采用的社交机器人来自不同的应用领域，其外观和能力存在较大差异，交互场景没有相似处等有很大关系。人机交互应该是一个长期且自然的互动状态，因此，未来研究应更多地考察刻板印象在人与社交机器人长期交往、真实场景以及不同场景互动过程中的影响^⑦，以及发生的变化。这对研究社交机器人刻板印象的产生、变化和干预手段等问题会有更多启示，也对理解人与社交机器人

-
- ① Oliveira, R., Arriaga, P., Correia, F., & Paiva, A., "The Stereotype Content Model Applied to Human-Robot Interactions in Groups," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 123 - 132.
- ② Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S., "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots," *International Journal of Social Robotics*, vol. 1, no. 1, 2009, pp. 71 - 81.
- ③ Carpinella, C. M., Wyman, A. B., Perez, M. A., & Stroessner, S. J., "The Robotic Social Attributes Scale (RoSAS) Development and Validation," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 254 - 262.
- ④ Ciecchanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P., "In the Shades of the Uncanny Valley: An Experimental Study of Human-Chatbot Interaction," *Future Generation Computer Systems*, vol. 92, 2019, pp. 539 - 548.
- ⑤ N. Mirnig, G. Stollnberger, M. Miksch, S. Stadler, M. Giuliani, & M. Tscheligi, "To Err is Robot: How Humans Assess and Act toward Anerroneous Social robot," *Frontiers in Robotics and AI*, vol. 4, 2017.
- ⑥ R. Aylett, "Games Robots Play: Once More, with Feeling," in *Emotion in Games*, Springer, 2016, pp. 289 - 302.
- ⑦ Bršćić, D., Kidokoro, H., Suehiro, Y., & Kanda, T., "Escaping from Children's Abuse of Social Robots," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 2015, pp. 59 - 66.

的关系意义重大。因而未来研究可以考虑通过纵贯研究，弥补实验研究的不足。此外，在测量部分，未来研究也应该纳入更多的生理指标测量，如肌电图、呼吸仪等，更加客观地反映实验对象的情感与情绪波动等。

（四）重视反思和修正人的偏见

已有社交机器人刻板印象研究在探索人机交互关系方面做出了重要努力，从经验主义角度验证了人机交互中，人的意志如何使媒介技术偏离了中立、客观的初衷，并逐步以服务人类之名义被操纵的结果。人类认知的“定型”和人工智能中机器学习的“建模”在很大程度上具有相似性。刻板印象形成与人从小的生活经历、文化背景等息息相关。而社交机器人的算法设计不仅从人类社会获取数据，构建模型，还在与人类交往中不断“成长”。但显然，人类是更强势的灌输者，机器是更智慧的学习者。对社交机器人言行感到愉悦，或许意味着我们已经将人类社会中的刻板印象乃至偏见在人工智能中合理化、标准化。虽然已有研究开始尝试运用反刻板印象调节社会偏见，但是仅限于教育类、养老社交机器人，缺少对其他应用场景下的研究。随着社交机器人种类的丰富，应用场景的增多，我们也应探讨在信息传播、医疗辅助、家庭服务等领域中的社交机器人反刻板印象设计，讨论如何利用社交机器人促进社会公平正义的可能。同时，我们还应考虑在社交机器人设计标准的基础上，从行业规制、价值观引导、法律、道德规范等方面展开更深入的研究。

社交机器人越来越人性化，当人类十分自然地将“他们”视为身边的社会行动者时，“他们”仅仅是服务于人类的社会行动者，还是调节人类社会关系的行动者，或是能够成为其他更多充满想象意味的社会行动者，这都需要我们更加审慎地思考。未来人机交互日益密切深入，刻板印象之外，人类社会中的哪些社会规则能在人机交互中继续演绎，又会如何改变人机交互的秩序与关系，甚至如何改变人类社会，这也亟待传播学者的想象与思量。

本文系上海浦江人才项目“人机交互中的社会行动者：社交机器人刻板印象研究”（项目编号：2020PJC038）的阶段性研究成果。

作者：南京大学新闻传播学院教授
华东师范大学传播学院硕士生

37 • When Robots Become Social Actors: Stereotypes in Human-Robot Interaction

•Shen Qi ,Wang Luyu

Human-robot interaction simulates the process of interpersonal communication. It is important to understand how people perceive social robots and interact with them. Based on the media equation theory ,and the computers-are-social-actors paradigm , the existing research on social robot stereotype found that people would treat social robots as social actors. People will generate corresponding emotions , attitudes , and behavior intentions to social robots according to the social rule of stereotype. Two types of research orientations have emerged from this: one is how to adapt social robots to stereotypes in human society to better serve humans; the other is how to intervene the prejudices in human society through counter-stereotype designs of social robots. Crawling and combing existing research , the present research provides theoretical and methodological references for investigating the social roles of social robots from the perspective of media equation theory , and provides inspiration for the exploration and reflection on the human-robot interaction relationship.

53 • How Did Chinese Internet Enterprises Participate in Internet Governance—A Case Study Based on Baidu , Alibaba and Tencent (BAT)

•Gao Haitao

China's Internet industry has undergone rapid development. The mainstream academic view is that Chinese Internet enterprises' participation in governance is obviously insufficient. This paper puts forward concepts of corporate Internet governance behavior and governance performance. Based on 930 news reports of Baidu , Alibaba and Tencent (BAT) from August 2016 to July 2019 , using the method of content analysis by NVivo12 Plus , this paper found that Chinese Internet enterprises are very active in governance. There were a large number of governance behaviors , including market service behavior , alliance and cooperation behavior , social propaganda behavior , political association behavior , environmental protection and public welfare behavior. They correspond to significant governance performance , mainly including technical performance , economic performance , political performance and ethical performance. The paper also explored four basic issues of Chinese Internet enterprises' participation in Internet governance: topics of concern and discussion , subjects involved and their relationships , principles or models of subject action , scope and area of action. The study offers advice on how to construct a three-dimensional collaborative governance framework led by the government and how Internet enterprises should participate in global Internet governance more effectively.

70 • Searching for a Balanced Approach: Practical Dilemmas and Revisions of the Principle of Informed Consent in Privacy Protection

•Fan Haichao ,Gu Liping

As one of the basic principles in the field of personal privacy protection , the principle of informed consent has a rich practical history and intrinsic value. However , with the rapid development of