

ロボットの援助行動が人間の信頼に及ぼす影響

Does resource cost matter? How robots' helping behavior influences human trust to them

杭陳琳*¹ 塩見昌裕*² 山田誠二*¹
Hang Chenlin Shiomi Masahiro Yamada Seiji

*¹神奈川大学 *²株) 国際電気通信基礎技術研究所
Kanagawa university ATR

ロボット工学技術の急速な発展に伴い、ロボットの援助行動は機能的支援としてだけでなく、社会的に意義のある支援としても認識されるようになってきている。しかしながら、人間同士の援助行動における重要な要因である援助提供者が負担する資源コストは、ヒューマン・ロボット・インタラクション (HRI) 研究においては未解明のままである。本研究では、援助行動によって生じる相対的な資源消費量が同等である場合でも、資源の観察可能な指標であるロボットの初期バッテリー残量が、人間の資源コスト認知およびロボットに対する信頼にどのような影響を与えるかを検証するため、オンライン実験を実施した。本研究は、より信頼性の高いロボットシステムの設計と、より良好な人間-ロボット関係の構築に向けた示唆を提供するものである。

1. はじめに

ロボティクスの進歩に伴い、HRIの研究は、ロボットの社会的行動と、それが個人、対人関係、および社会規範に与える影響にますます焦点を当てている [1]。個人レベルでは、ロボットは感情 [2]、意思決定 [3]、受容 [4] に影響を与える。対人レベルでは、ロボットは相互作用と信頼構築を形成し [5]、社会レベルでは、規範や集団行動を再定義する可能性がある [6]。様々な社会的行動の中で、向社会的行動は HRI 研究の中心であり、そのコストは低いもの (例: 情報提供) から高いもの (例: 自己犠牲) まで多岐にわたる [7]。ロボットの自己犠牲に関する従来の HRI 研究は、主に道徳的ジレンマのシナリオ (よく知られたトロッコ問題など [8]) に焦点を当ててきた。この研究パラダイムには価値があるものの、極端すぎるシナリオや日常的な適用可能性の欠如といった限界もある。先行研究では、自律型警備ロボットが自己防衛の意図と比較して善意的な意図を持つと説明された場合、参加者のロボットに対する信頼が増加したことが報告されている [9]。この研究はより実生活に近い状況を検討したものの、テキストを用いてロボットの特性を説明するのみで、これらの特性をロボットの実際の行動に反映させていない。このギャップに対処するため、我々は日常生活に近い研究シナリオを提案する: ロボットが自身の電力を使用する場合と外部バッテリーを使用してユーザーのデバイスを充電する場合とで、人々の社会的行動 (例: 援助や信頼) における反応を比較する。本研究は 2 つの問いを探究する:

ロボットの自己犠牲的行動は、人間がロボットをどのように認識するかに影響を与えるか?

ロボットの自己犠牲的行動は、人々のロボットに対する向社会的行動への関与に影響を与えるか?

2. 実験デザイン

参加者の募集は Yahoo!クラウドソーシングプラットフォームを通じて実施され、最終的な分析対象サンプルは 136 名 (男性 72 名、女性 64 名、年齢: $Mean = 50.24$, $SD = 11.33$, Range: 24-81) で構成された。

資源の重要性に関する実験デザインに基づき、2 つの条件を実施した: ロボットに対する資源ニーズの重要度が低い条件と高い条件である。高重要度条件では、Sota が自分のバッテリー



図 1: 実験条件

を共有する時、横にあるディスプレイ画面は最初にバッテリーレベル 25% を示し、電力共有後に 5% に低下した (すべての数値は実験者のみが見ることができた)。さらに、バッテリーレベルが減少するにつれて、残バッテリーの色が緑から赤に変化した。低重要度条件では、ディスプレイ画面は最初にバッテリーレベル 100% を示し、電力共有後に 80% に低下した。この条件では、残バッテリーの色はプロセス全体を通じて緑のままであった。

実験の手続きについて、まず、参加者に実験の紹介を読んでもらい、次に人口統計学的質問とロボット使用頻度に関する質問を行った。その後、参加者にはテキストと画像を含むシナリオの説明が提供された。ビデオ刺激を視聴した後、参加者にはビデオ内容に関する 1 つの理解度確認質問が行われた。その後、参加者には一連の事後質問紙への回答を求めた。最後に、報酬を受け取るために、最後にコードが記載されたフォームに記入した。

3. アンケート

実験の開始時に、参加者に個人情報 (性別など) とロボット使用に関する過去の経験についての質問を行った。実験刺激を見た後、ロボットの認識に関して Godspeed [11] と MDMT v2 [12] を実施し、ロボットに対する向社会的行動について独裁者ゲームを実施した [13]。

4. 結果

1 要因 2 水準の実験デザインには常に独立サンプルの t 検定が使用されるが、Shapiro-Wilk 検定を実施したところ、すべてが有意に非正規分布であることが示されたため、本研究ではノンパラメトリック検定である Mann-Whitney U 検定を使用した。2 つのグループ間には有意差があり、高重要度条件 ($Mean = 4.19, SD = 1.72$) は、低重要度条件 ($Mean = 3.33, SD = 1.35$) よりもロボットの援助行動のコストをより高く認識した。RoSAS の結果について、温かさ ($p = 0.739, r = 0.0033$) および不快感 ($p = 0.691, r = 0.039$) については、2 つのグループ間に有意差はなかった。

ポジティブおよびネガティブな感情の結果について、罪悪感と恥の観点で資源タイプ間に有意差があり ($p = 0.002, r = 0.298$)、高重要度条件の参加者 ($Mean = 2.21, SD = 1.11$) は、低重要度条件の参加者 ($Mean = 1.69, SD = 0.85$) よりもロボットに対してより多くの罪悪感と恥を感じた。しかし、感謝については有意差はなかった ($p = 0.532, r = 0.062$)。

MDMT-ver2 の結果について、まず 2 つの広範な次元を分析した。パフォーマンス信頼 ($p = 0.838, r = 0.021$) および道徳的信頼 ($p = 0.838, r = 0.021$) は、2 つのグループ間に有意差がなかった。より詳細には、各主要因子の下位次元についてもさらなる分析を実施した。パフォーマンス信頼については、信頼性 ($p = 0.652, r = 0.045$) が 2 つのグループ間で有意性を示し、有能性は有意性を示さなかった ($p < 0.001, r = 0.409$)。道徳的信頼については、倫理的有能性 ($p = 0.095, r = 0.166$)、透明性 ($p = 0.991, r = 0.001$)、および善意性 ($p = 0.637, r = 0.046$) は、2 つのグループ間で有意性を示さなかった。

バッテリー版独裁者ゲームの結果について、2 つのグループ間に有意差はなかった ($p = 0.462, r = 0.066$; 高重要度: $Mean = 47.7, SD = 15.0$; 低重要度: $Mean = 47.2, SD = 14.6$)。

5. 考察

人間同士の相互作用 (HHI) において、高重要度の援助は通常、希少性下での犠牲を反映するため、より道徳的に意味のあるものとして認識される [17]。我々の知見はこのパターンに密接に従っていた: ロボットがバッテリー残量が少ない状態で電力を共有した場合、参加者はより大きな懸念と罪悪感を示した。参加者はそのような援助をより信頼性の高い犠牲的行為と評価して、ロボットを「親切」または「人間らしい」とさえ表現した。対照的に、援助が高バッテリーレベル (低重要度) で行われた場合、それはより頻繁に便利で信頼できるものと解釈され、比較的中立的な感情反応が示された。これは、HRI においても、HHI と同様に、資源の希少性が道徳的感情と社会的評価を著しく増幅することを示唆している。

しかし、注目すべきことに、HHI からの相違が現れた: HRI では、参加者はロボットの援助が善意的または価値があるかどうかを評価するだけでなく、その決定の合理性と持続可能性についても疑問を呈した。例えば、ロボットが低バッテリーにもかかわらず援助を提供した場合、参加者はその後もロボットが適切に機能できるかどうかを心配したり、そのような「犠牲」が合理的な選択であるかどうかを疑ったりした。これは、ロボットが人工エージェントとしての二重のアイデンティティを持つことを反映している—社会的行為者と技術システムの両方として同時に認識され—これが合理性と機能的持続可能性に関する追加的な認知的ジレンマを引き起こす。この知見は、

ロボット行動における機能性と社会性のバランスの重要性を強調することで、社会的ロボットの設計に関する議論を拡張するものである。

参考文献

- [1] Lugin, B., Pelachaud, C., and Traum, D. (2022). The Handbook on Socially Interactive Agents: 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics, Volume 2: Interactivity, Platforms, Application. Morgan & Claypool.
- [2] Rosenthal-von der Pütten, A. M., Krämer, N. C., Hoffmann, L., Sobieraj, S., and Eimler, S. C. (2013). An experimental study on emotional reactions towards a robot. *Adv. Robot.* 5, 17–34.
- [3] Hou, Y. T.-Y., Lee, W.-Y., and Jung, M. (2023). “Should I Follow the Human, or Follow the Robot?” — Robots in Power Can Have More Influence Than Humans on Decision-Making, in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, (New York, NY, USA: Association for Computing Machinery), 1–13.
- [4] Savela, N., Turja, T., and Oksanen, A. (2018). Social acceptance of robots in different occupational fields: A systematic literature review. *Int. J. Soc. Robot.* 10, 493–502.
- [5] Naneva, S., Sarda Gou, M., Webb, T. L., and Prescott, T. J. (2020). A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots. *Int J of Soc Robotics* 12, 1179–1201.
- [6] Hang, C., Ono, T., and Yamada, S. (2021). Designing nudge agents that promote human altruism, November 10–13, 2021, Proceedings 13. Available at: https://link.springer.com/chapter/10.1007/978-3-030-90525-5_32
- [7] Oliveira, R., Arriaga, P., Santos, F. P., Mascarenhas, S., and Paiva, A. (2021). Towards prosocial design: A scoping review of the use of robots and virtual agents to trigger prosocial behaviour. *Comput. Human Behav.* 114, 106547.
- [8] Lee, M., Ruijten, P., Frank, L., de Kort, Y., and IJsselstein, W. (2021). “People May Punish, But Not Blame Robots,” in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 1–11.
- [9] Lyons, J. B., Jessup, S. A., and Vo, T. Q. (2024). The role of decision authority and stated social intent as predictors of trust in autonomous robots. *Top. Cogn. Sci.* 16, 430–449.
- [10] Dovidio, J. F., Piliavin, J. A., Schroeder, D. A., and Penner, L. A. (2017). *The Social Psychology of Prosocial Behavior*. Psychology Press.

-
- [11] Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* 1, 71–81.
- [12] Ullman, D. and Malle, B. F. (n.d.). MDMT: Multi-Dimensional Measure of Trust v2. Available at: [https://research.clps.brown.edu/SocCogSci/Measures/MDMT_v2_\(2023\)_Full_scale.pdf](https://research.clps.brown.edu/SocCogSci/Measures/MDMT_v2_(2023)_Full_scale.pdf)
- [13] H. Chenlin, O. Tetsuo, and S. Yamada, “Perspective-taking for promoting prosocial behaviors through robot-robot VR task,” 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Busan, South Korea, 2023, pp. 2100–2105, doi: 10.1109/RO-MAN57019.2023.10309610.