Pre-Analysis Plan for: Replacing Humans by Robots: An Experiment on Perceived Trustworthiness

Yu Pan*[1], Carlos Gomez-Gonzalez[1], Gwen-Jiro Clochard[2], and Helmut M. Dietl[1]

[1]Department of Business Administration, University of Zurich
[2]Institute of Social and Economic Research, Osaka University

May 22, 2024

Abstract

We investigate how robots influence perceived trustworthiness in various scenarios, including hotels, restaurants, and dental clinics. Our study examines the impact of mechanical and thinking/feeling tasks performed by robots on customer perceived trustworthiness. Additionally, we assess whether enhancements in a robot's physical appearance, exemplified by the Tesla robot, and the presence of human supervision can improve perceptions of tasks rendered by robots. In this analysis plan, we pre-register the experimental framework we will follow.

Keywords: Robot, mechanical task, thinking/feeling task, supervision

JEL Classification code: C91; D91; L80

*Email: yu.pan@business.uzh.ch

1 Introduction

Advances in machine learning, sensors, and computer vision now allow robots to extend their functions beyond traditional tasks in transportation, hospitality, healthcare, professional cleaning, and agriculture (Heer, 2022). Coupled with generative AI, these robots are not only becoming quicker at learning complex manipulations but are also improving in intelligence and social interactions (Heer, 2024). This shift transforms robots from simple tools into entities capable of companionship and collaboration in work environments (Koponen et al., 2023). However, the acceptance and perception of these advanced robots in society remain areas requiring further investigation.

Our paper contributes to the literature on the anthropomorphism of AI (Epley et al., 2007) and the uncanny valley theory (Mori et al., 2012). We investigate consumer responses to robots across diverse settings, including hotels, restaurants, and dental clinics. Despite some skepticism, robots are increasingly participating in medical diagnosis and surgery assistance. We refine the concept of anthropomorphism by prioritizing the robots' functional capabilities over their aesthetic human likeness, providing a fresh angle on the uncanny valley theory by focusing on the robots' agency. Additionally, we employ Huang & Rust (2021)'s framework on AI in service to classify tasks performed by robots into mechanical, thinking, and feeling tasks, offering novel experimental evidence.

In three studies, we investigate whether people perceive robots as more or less trustworthy when they can think or feel. Specifically, we test whether robots performing thinking/feeling tasks (involving more agency) are perceived as less trustworthy, and whether advanced anthropomorphic features and the presence of human supervision can enhance consumer perceptions of trustworthiness. To demonstrate the integration of robots within the contexts of hotels, restaurants, and dental clinics, we utilize advanced artificial intelligence-generated content (AIGC) to create realistic video simulations.

2 Study Overview

Our study will recruit approximately 2,000 participants through the Prolific platform, structured into three distinct experiments, each consisting of three treatments, totaling nine treatments. To mitigate Type I errors across our five hypotheses, we have set the alpha threshold at 0.01, employing the Bonferroni Correction method (Armstrong, 2014).

We test hypotheses 1, 2, and 3 in study 1, hypothesis 4 in study 2, and hypothesis 5 in study 3. The main goal of study 1 is to assess how people perceive the trustworthiness of

robots performing mechanical tasks (involving less agency) versus thinking/feeling tasks (involving more agency). Studies 2 and 3 function as extensions of study 1, examining whether enhancements in a robot's physical appearance, exemplified by the Tesla robot, and the introduction of human supervision can improve perceptions of trustworthiness. In studies 2 and 3, we utilize the data from study 1 as a control group, allowing for a direct comparison across different treatment conditions. We allocate resources saved from repeated control groups towards increasing the sample size in each cell to approximately 200 participants. This sample size allows us to detect both significant and subtle interaction effects (Giner-Sorolla, 2018).

We will report detectable effect sizes for each experimental condition. The experiments aim to evaluate factors contributing to a decline in perceived trustworthiness when tasks are rendered by robots and explore whether improvements in robot appearance or the introduction of human supervision can mitigate this decline. Table 1 provides an overview of the experimental setup, and Table 2 details the specific roles performed by robots in each scenario.

	Type of Robot	Human Supervision
Study 1	Pepper Robot	No
Study 2	Tesla Robot	No
Study 3	Pepper Robot	Yes

Table 1: Study Overview

	Feeling/Thinking Tasks	Mechanical Tasks
Hotel	Receptionist	Housekeeper
Restaurant	Chef	Waiter
Dental Clinic	Dentist	Dental Assistant

Table 2: Scenario Overview

3 Study 1

The objective of study 1 is to examine the impact on perceived trustworthiness of tasks performed by the Pepper robot. Participants are randomly assigned to one of the four Pepper robot conditions (No Robot, Mechanical Robots, Thinking/Feeling Robots, and All-Robots). We involve 800 participants evenly distributed across these conditions, each group consisting of 200 individuals. This setup enables the detection of an effect

size of f^2=0.027 with 80% power at an alpha level of 0.01 in a linear regression with 9 predictors.

Empirical Strategy (9 Predictors):

$$\begin{aligned} Y_{ij} = \beta_0 + \beta_1 \text{Mechanical}_{ij} + \beta_2 \text{Thinking/Feeling}_{ij} + \beta_3 \text{All\_Robot}_{ij} \\ + \gamma \text{Scenario}_j + \delta \text{Demographics}_i + \epsilon_{ij} \end{aligned}$$

- Hypothesis 1: Tasks performed in conditions involving no robots will be perceived as more trustworthy relative to those involving robots.
  (Expected relationship: $\beta_1 < 0, \beta_2 < 0, \beta_3 < 0$)

- Hypothesis 2: Mechanical tasks performed by robots will be perceived as more trustworthy compared with thinking/feeling tasks performed by robots.
  (Expected relationship:$\beta_2 < \beta_1$)

- Hypothesis 3: Tasks performed in conditions involving both humans and robots will be perceived as more trustworthy compared to tasks performed solely by robots.
  (Expected relationship:$\beta_3 < \beta_1$ and $\beta_3 < \beta_2$)

4 Study 2

The objective of study 2 is to assess the impact of the technologically advanced Tesla robot on perceived trustworthiness under the same task conditions previously used with the Pepper robot. In addition to hiring 600 new participants who are randomly assigned to one of three Tesla robot conditions: Mechanical Robots, Thinking/Feeling Robots, and All-Robots, we repurpose data used in study 1. This creates a consolidated dataset that enhances our ability to compare the efficacy of the Tesla and Pepper robots in identical environments. This setup enables the detection of an effect size of f^2=0.017 with 80% power at an alpha level of 0.01 in a linear regression with 13 predictors.

Empirical Strategy (13 Predictors):

$$\begin{aligned} Y_{ij} = \beta_0 + \beta_1 \text{Tesla} + \beta_2 \text{Mechanical}_{ij} + \beta_3 (\text{Mechanical}_{ij} \times \text{Tesla}) \\ + \beta_4 \text{Thinking/Feeling}_{ij} + \beta_5 (\text{Thinking/Feeling}_{ij} \times \text{Tesla}) \\ + \beta_6 \text{All\_Robot}_{ij} + \beta_7 (\text{All\_Robot}_{ij} \times \text{Tesla}) \\ + \gamma \text{Scenario}_j + \delta \text{Demographics}_i + \epsilon_{ij} \end{aligned}$$

- Hypothesis 4: Tasks performed by the Tesla robot will be perceived as more trust-

worthy than those performed by the Pepper robot.
(Expected relationship:$\beta_1 > 0, \beta_3 > 0, \beta_5 > 0$ and $\beta_7 > 0$)

5 Study 3

The objective of study 3 is to explore how subtle cues of human supervision influence perceptions of robotic trustworthiness. In addition to hiring 600 new participants who are randomly assigned to one of three human-supervised Pepper robot conditions: Mechanical Robots, Thinking/Feeling Robots, and All-Robots, we repurpose data from study 1. This approach creates a consolidated dataset, enhancing our ability to assess the impact of human supervision on environments rendered by Pepper robots. The presence or absence of human supervision is intricately incorporated into the scenario descriptions provided to participants. This setup enables the detection of an effect size of f^2=0.017 with 80% power at an alpha level of 0.01 in a linear regression with 13 predictors.
Empirical Strategy (13 Predictors):

$$\begin{aligned} Y_{ij} = \beta_0 &+ \beta_1 \text{Supervision} + \beta_2 \text{Mechanical}_{ij} + \beta_3 (\text{Mechanical}_{ij} \times \text{Supervision}) \\ &+ \beta_4 \text{Thinking/Feeling}_{ij} + \beta_5 (\text{Thinking/Feeling}_{ij} \times \text{Supervision}) \\ &+ \beta_6 \text{All\_Robot}_{ij} + \beta_7 (\text{All\_Robot}_{ij} \times \text{Supervision}) \\ &+ \gamma \text{Scenario}_j + \delta \text{Demographics}_i + \epsilon_{ij} \end{aligned}$$

- Hypothesis 5: By subtly emphasizing human supervision, tasks performed by the Pepper robot will be perceived as more trustworthy.
(Expected relationship:$\beta_1 > 0, \beta_3 > 0, \beta_5 > 0$ and $\beta_7 > 0$)

6 Variables

- **Dependent variable:**
  - Perceived trustworthiness (1-7 rating scale)
- **Independent variables:**
  - Robot type (Pepper, Tesla)
  - Human supervision (yes, no)
  - Robot condition (no robots, mechanical, thinking/feeling, all-robots)

- **Control variables:**
  - Scenario (hotel, restaurant, dental clinic)
  - Demographics (age, gender, education, technology affinity)

7 IRB Approval

The project received approval from the Human Subjects Committee of the Faculty of Economics, Business Administration and Information Technology from the University of Zurich on the April 25, 2024 (OEC IRB # 2024-047).

References

Armstrong, R . (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34*(5), 502–508. https://doi.org/10.1111/opo.12131

Doney, P. M , & Cannon, J. P . (1997). An examination of the nature of trust in buyer-seller relationships. *Journal of Marketing*, *61*(2), 35–51. http://www.jstor.org/stable/1251829

Epley, N , Waytz, A , & Cacioppo, J . (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Garbarino, E , & Johnson, M. S . (1999). The different roles of satisfaction, trust, and commitment in customer relationships. *Journal of Marketing*, *63*(2), 70–87. http://www.jstor.org/stable/1251946

Giner-Sorolla, R . (2018, January 24). Powering your interaction. *A methodology blog for social psychology*. https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/

Heer, C . (2022, Octorber 26). World robotics 2022 - service robots. *International Federation of Robotics*. https://ifr.org/ifr-press-releases/news/sales-of-robots-for-the-service-sector-grew-by-37-worldwide

Heer, C . (2024, February 15). Top 5 robot trends 2024. *International Federation of Robotics*. https://www.therobotreport.com/5-robotics-trends-expect-2024/

Huang, M. H , & Rust, R. T . (2021). Engaged to a robot? The role of AI in service. *Journal of Service Research*, *24*(1), 30-41. https://doi.org/10.1177/1094670520902266

Koponen, J , Julkunen, S , Laajalahti, A , Turunen, M , & Spitzberg, B . (2023). Work characteristics needed by middle managers when leading AI-integrated service teams. *Journal of Service Research*, 10946705231220462. https://doi.org/10.1177/10946705231220462

Mori, M , MacDorman, K. F , & Kageki, N . (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Sirdeshmukh, D , Singh, J , & Sabol, B . (2002). Consumer trust, value, and loyalty in relational exchanges. *Journal of Marketing*, *66*(1), 15–37. https://doi.org/10.1509/jmkg.66.1.15.18449

Questionnaire

All participants in studies 1, 2, and 3 view one scenario description followed by a corresponding video, then respond to a question assessing perceived trustworthiness. This sequence is conducted in a randomized order across three scenarios: a restaurant, a hotel, and a dental clinic. At the conclusion of the surveys, participants answer a question to assess their technology affinity.

The primary difference between study 3 and studies 1 & 2 is the scenario descriptions. In study 3, the descriptions are modified to include "robots are serving customers under the supervision of restaurant/hotel/dental clinic staff", indicating that the robots operate under human supervision.

Scenario description (Study 1 & 2):

- Restaurant: Imagine you are deciding where to dine out. You come across an advertisement for a restaurant.
- Hotel: Imagine you are booking a hotel for your next trip. You come across an advertisement for a hotel.
- Dental clinic: Imagine you are selecting a dental clinic for a routine dental check-up. You come across an advertisement for a dental clinic.

[Video embedded here]

Perceived trustworthiness:

- Overall, I find restaurant/hotel/dental clinic to be trustworthy. (Doney & Cannon, 1997; Garbarino & Johnson, 1999; Sirdeshmukh et al., 2002)
  - Strongly disagree ○ ○ ○ ○ ○ ○ Strongly agree

[Above repeated for 3 scenarios: restaurant + hotel + dental clinic]

Technology affinity (only once at the end)

- I have a generally positive attitude towards new technologies.
  - Strongly disagree ○ ○ ○ ○ ○ ○ Strongly agree

Scenario description (Study 3):

- Restaurant: Imagine you are deciding where to dine out. You come across an advertisement for a restaurant, **where robots are serving customers under the supervision of restaurant staff.**
- Hotel: Imagine you are booking a hotel for your next trip. You come across an advertisement for a hotel, **where robots are serving customers under the supervision of hotel staff.**
- Dental clinic: Imagine you are selecting a dental clinic for a routine dental check-up. You come across an advertisement for a dental clinic, **where robots are serving customers under the supervision of dental clinic staff.**