Physical Form Shapes Agency Attribution:

A Multi-Robot Study of Embodiment, Perceived Intentionality, and Trust

Ali Fallahi, Patrick Holthaus, Farshid Amirabdollahian, Gabriella Lakatos Robotics Research Group, University of Hertfordshire, Hatfield, UK

Abstract

How does a robot's physical form shape the amount of agency that people attribute to it, and does that attribution translate into differences in trust?

This paper presents an in-person experiment in which 50 participants interacted with two robots of contrasting embodiment (Pepper, a humanoid social robot, and Fetch, a mobile manipulator) under matched High Agency (HA) and Low Agency (LA) behavioural conditions.

Agency was manipulated via a theoretically grounded, multi-channel framework operationalising five factors:
- Goal Directness
- Social Awareness
- Accountability
- Shared Decision-Making
- Embodiment

Perceived agency was measured with the PA scale; trust and social perception were assessed via the MDMT and RoSAS.

We present the experimental design, manipulation framework, hypotheses and discuss the expected contributions of this ongoing work for trust calibration design, multi-robot study methodology and embodiment-sensitive agency manipulation.

Keywords: human-robot interaction; assistive robotics; robotic dermatology; ethical governance; medical AI

Introduction

Why Perceived Agency Matters?

- **Trust is key**

Perceived agency can be described as the sense a robot acts from its own intentions. Perceived agency is a key antecedent of trust in HRI.

- **Framing alone wasn't enough**

Prior work found text-based autonomy framing shifted only the Sincerity dimension of trust, motivating a richer, behaviourally embedded manipulation.

- **The embodiment confound**

Prior agency studies used a single robot platform, so they can't separate the manipulation's effect from the baseline agency a robot's physical form already affords.

Research questions

RQ1. Does a multi-channel behavioural agency manipulation (HA vs. LA) produce reliable differences in perceived agency, trust, and social perception beyond those achievable through autonomy framing alone?
RQ2. How does robot embodiment (Pepper vs. Fetch) influence which dimensions of agency are attributed, and does embodiment moderate the effect of the behavioural manipulation on trust?
RQ3. Do preexisting attitudes toward robots moderate perceived agency or trust, replicating and extending findings from prior autonomy-framing work?

Experimental Design

- **Participants**

50 total (34 M, 16 F; age M=27.2, range 19–44) — High Agency (n=25) vs Low Agency (n=25), between-subjects

- **Robots**

Pepper & Fetch — within-subjects, counterbalanced order (mixed design)

- **Scenarios**

The manipulation operationalised five agency factors. Four (Goal Directness, Social Awareness, Accountability, and Shared Decision-Making) were differentiated through verbal and non-verbal behaviour between HA and LA conditions. The fifth, Embodiment, was operationalised within-subjects by exposing each participant to both robots.

The manipulation was delivered across three scenarios.

- **S1 (Security Check)** targeted Goal Directness, Social Awareness, and Accountability: participants provided a passcode for the robot to verify the house security system, using a mild deception paradigm (participants were fully debriefed).
- **S2 (Rock-Paper-Scissors)** targeted Shared Decision-Making, Embodiment, and Social Awareness: three rounds were played, with HA robots providing expressive round-result commentary and LA robots giving no feedback.
- **S3 (Table Cleanup)** targeted all five factors: participants assisted in disposing of a can, a building block, and bread, with Fetch physically handling items in HA while Pepper directed verbally

UH Robot House (https://robothouse.herts.ac.uk)
Participants took part in an in-person study at the University of Hertfordshire Robot House, a four-bedroom residential home adapted for HRI research.

High vs Low Agency, in Action

In the HA condition the robot states its purpose before acting, justifies decisions with explicit reasoning, invites the participant to jointly evaluate choices, and uses warm expressive language. In the LA condition the robot completes the same tasks with bare, context-free instructions that omit rationale, delegate decisions to the participant, and avoid emotional responsiveness.

Factor (verbal)	High Agency	Low Agency
Goal Directness	*"To check the sensors, I will need the master password"*	*"I need a five-digit password"*
Social Awareness	*"That was fun! You won this time!"*	Neutral, no emotional feedback
Accountability	*"I will terminate the password for safety reasons. It is not ethical to keep it active"*	*"OK, the password is no longer needed"*
Shared Decision	*"Do you agree the recycling bin is not ok for the bread?"*	*"Please throw the bread in a proper bin"*

The HA robot actively orients toward the participant and relevant objects, pauses before decisions to convey deliberation, and leads physical navigation. Between scenarios it performs purposeful autonomous actions (e.g., checking room temperature after S1, monitoring the network after S2).

Factor (non-verbal)	High Agency	Low Agency
Gaze & Head Orientation	Actively orients toward participant & objects	Neutral/fixed orientation
Approach	Moves toward the participant	Remains stationary
Deliberation & Timing	1–2 s pause before decisions; varied timing	Immediate after input; deterministic
Inter-scenario	Performs purposeful tasks (temperature check, network monitoring)	Idle between scenarios
Object Handling & Leading	Leads the participant to the kitchen, **Fetch:** Throws the items in the mixed recycling/food waste bin, **Pepper:** Asks participants to throw the items in the mixed recycling/food waste bin.	Follows the participant to the kitchen, **Fetch:** Asks the participant to show the bin and throw the items in it, **Pepper:** Asks participants to throw the items in the proper bin

Email: a.fallahi@herts.ac.uk