

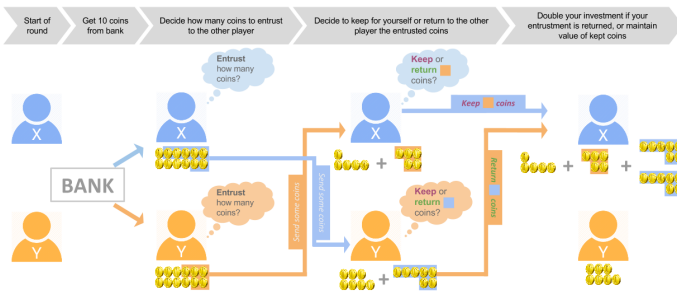
Motivation & Hypotheses

Trust and cooperation play a vital role in human interactions. We explored how trust and cooperation impact human-robot interactions by testing the following hypotheses using a coin entrustment game:

1. Humans will trust robots more completely than with other humans.
2. Humans will cooperate more frequently with robots than with other humans.

The Coin Entrustment Game

The Coin Entrustment (CE) game illustrated below is a variant of Iterative Prisoner's Dilemma (IPD) that attempts to separately measure trust and cooperation. In each of an undisclosed number of game rounds, both players begin with 10 coins and must make two decisions: (1) How many of their coins to entrust to the other player; and (2) whether to keep or return the coins that were entrusted to them by the other player. Coins that are returned double in value.



We measure trust and cooperation as:

- **Trust**—average number of coins entrusted in each round; and
- **Cooperation rate**—the rate at which players choose to return their opponents coins.

Experimental Procedure

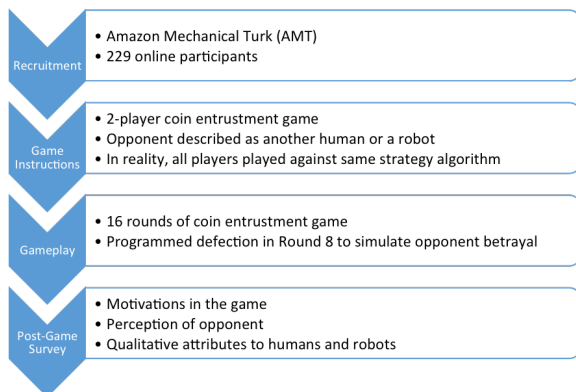
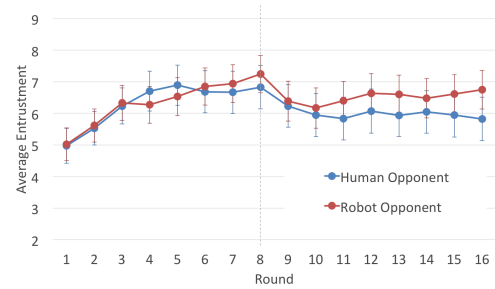


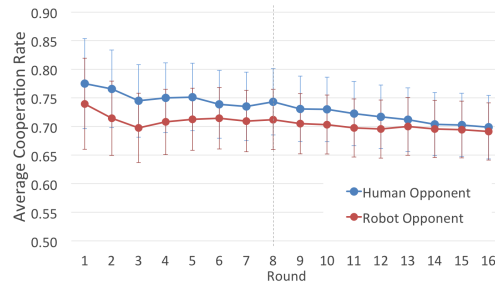
Figure 1: Gameplay process for experiment participants.

Our experiment was run on Amazon's Mechanical Turk, a crowd-sourcing site, with a total of 229 players. Participants navigated through three sections: a consent page with game instructions, 16 rounds of the CE game (the number of rounds is undisclosed to the player), and a post-game survey.

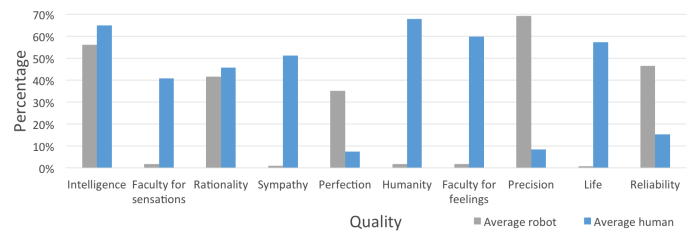
Results



Trust—Perceived opponent type leads to statistically significant differences in the number of coins entrusted across the 16 rounds (Mixed ANOVA: $F(15, 3405) = 1.804, p < .05$).



Cooperation rate—There is no significant difference in cooperation rates across the 16 rounds due to perceived opponent type.



Our post-game survey reveals participants perceive that humans possess greater sympathy and faculty for feelings and sensations, while robots are more precise and reliable.

Take-aways & Future Directions

The Coin Entrustment game offers insights into the nature of human-robot teamwork:

- Participants develop **trust faster** and to a **greater extent** with robots than humans
- Participants **cooperate** with robots and humans at **similar rates**
- Participants tend to respond to defection by lowering trust rather than cooperation
- Participants who play against a robot are primarily **motivated** by *maximizing profit*, whereas those who play against humans are primarily motivated by a desire to *win*.

Future Directions: We would like to extend our understand of trust and cooperation to embodied interaction domains. We would also like to explore possible dynamics in inter-personal relationships that are missing from human-robot interactions, e.g., the desire for social dominance.

For More Information

- Web: <https://www.cs.hmc.edu/HEAT/>
- Facebook: <https://www.facebook.com/muddheatlab>
- Twitter: <https://www.twitter.com/muddheatlab>