Reference:

People Do Not Feel Guilty About Exploiting Machines

Introduction:

In this study, the authors attempt to show that by measuring both guilt and envy in various economic games (lottery, ultimatum, and dictator), you can show a lower amount of guilt towards taking advantage of machines and an equal amount of envy towards machines than if they were replaced by humans in the same situation. The studies sited by the authors are generally used to establish a baseline that A) humans can treat machines as humans in social settings, and that B) humans tend to treat other humans preferentially compared to machines. They also establish that emotion has a profound effect on social interactions, specifically those emotions of guilt and envy.

Method (Experiment 1):

Participants:

There were 81 people participating in this experiment. Participants were chosen randomly with no knowledge or identifiable features available to either the researchers or the participants via Amazon's Mechanical Turk. The paper does not specify how the conditions were assigned, however based on the results table I believe that each participant did a round where they were told they were against a human and a round where they were told they were against a machine. In both cases, they were against the same algorithm.

Conditions:

The independent variables were whether the participants were told they were against a human or a machine.

Measures:

The dependent variables were the participants' self-reported answers to the following questions on a scale of 1 to 7:

(1) Were you concerned with not taking advantage of the other party?
(2) Did you feel guilty when the outcome favored you?
(3) Were you comfortable with outcomes that favored you over the other party?
(4) Were you concerned with not being taken advantage of?
(5) Did you feel envious when the outcome favored the other party?
(6) Were you comfortable with outcomes that favored the other party over you?

Procedure:

Each participant and their counterpart was given 20 tickets and told that any tickets placed into the public pool would be given 1.5x their value then split between the two players. These tickets at the end of the game were entered into a $30. Machine players were also entered into the lottery. The two participants then took turns offering tickets to be placed into the pool for 8 rounds.

Method (Experiment 2):

Participants:

There were 165 people participating in this experiment. Participants were found from the Psychology student pool of the University of Southern California. Each participant was given one of a human or machine counterpart, and one of either the ultimatum game or the modified dictator game.

Conditions:

The independent variables were the game and counterpart to which the participant was assigned, and their choices based on those initial conditions.

Measures:

The dependent variables were the α (envy) and β (guilt) values found using Blanco et al.'s procedure.

Procedure:

The participants were asked to give their choice for every possible combination of tickets offered/tickets kept. They were then matched with a partner based on those answers. Their scores for α and β were then found using their performance in their round.

Results and Discussion:

These experiments found that, as the researchers hypothesized, humans feel less guilt and the same amount of envy towards machines as compared to humans. This has, at the least, a potential impact on upcoming fields such as artificial intelligence. I would suggest more research

along this line, e.g., telling people that they are competing against an AI as opposed to a "machine."

Conclusion:

I think a meaningful conclusion here is that people treat machines and other people *differently*. Important in that is that people still treat machines like people, just *lesser* people. That is, if they were treating the machine purely as a machine, the correct move to make in the ultimatum game is to offer one ticket. However, on average people offered the machine 6.71 tickets. I think this is interesting because it means that in the short term of developing a general AI it means people will be likely to accept them as sentient things. However, over the long term, we will probably be looking at a fight over AI rights if we reach that point.