

Human versus robot recipients in dictator games: Allocation of gains versus losses

Abstract

This thesis investigates behavioral economic interactions with robot recipients, using the Dictator Game (DG) to ask whether the social norms that govern human-human allocation extend to artificial agents such as robots. Drawing on the Computers Are Social Actors (CASA) framework and the Media Equation, it first tests whether humans treat robots as socially meaningful recipients once allocations carry real economic consequences, or whether sharing remains categorically reserved for human counterparts. Building on this, the thesis then asks how allocation behavior changes when the outcome domain shifts from gains to losses, and whether any such shift is shaped by loss aversion, by inequality aversion, or by a recipient-contingent harm avoidance motive. Across four studies and a pooled cross-study analysis, recipient type is varied from humans to robots, computers, and non-social objects, alongside outcome framing, stakes, and participant sample. Across the studies, the findings converge on a coherent picture of who is treated as a sharing partner and how outcome valence shapes that decision. Participants do not treat robots as social recipients once allocations carry real cost. The human/non-human boundary is sharp, the social presence gradient does not hold within the non-human side of that boundary, and robots are indistinguishable from rocks. A weaker reading of CASA remains compatible with hypothetical- stakes results and with the broader low-stakes literature, but the strong claim that surface cues automatically trigger default social treatment does not hold when decisions carry genuine cost. Allocations to non-human recipients cluster at zero and at the equal split, and the equivalence of robots and rocks indicates that participants have made a prior, deliberate judgment that the sharing norm does not apply, rather than reducing their offers in response to weak cues. A gain-loss asymmetry runs through all recipient types, including non-social objects. Formal modeling of fairness preferences in the loss domain places this effect in inequity aversion rather than loss aversion, and shows that the shift is driven more by participants' aversion to having suffered a smaller loss than the recipient than by their aversion to having suffered a larger one. A pooled cross-study analysis then separates the allocation into two decisions, whether to share at all and how much to give when sharing. The first depends on who the recipient is and is consistent with a harm avoidance motive; the second applies regardless of recipient and is consistent with inequity aversion. The two accounts operate at different decision levels rather than competing to explain the same effect. Taken together, the findings support two concrete conclusions. Strong readings of CASA do not hold once allocations carry real cost, and surface social cues alone do not produce social treatment of artificial agents. DG offers to non-human recipients are best read as a judgment about whether sharing norms apply, not as a graded measure of prosocial preference, a distinction that matters wherever the DG is used to compare human and non-human targets. These claims are bounded by the artificial nature of the paradigm itself, a one-shot anonymous laboratory game, and need to be tested in richer, repeated, or more natural interactions with artificial agents. The pattern, however, is consistent across Indian and UK/Western European samples, suggesting that the categorical treatment of artificial agents in the DG generalizes beyond a single cultural context.