

Supplementary Material

1 SUPPLEMENTARY DATA

All the data presented here are based on our exclusion criteria of two standard deviations from the mean of the analysis being performed. If a participant is excluded for not being in a normal range from one analysis, they are excluded from all further analyses. In this way, we prevent outliers from skewing our results. We had a total of 62 participants for this study, and six participants were excluded for falling outside our exclusion criteria.

Inflation limit

The BART consisted of two blocks of 30 balloons each. This design was chosen because we hypothesised that participants might use the first block to familiarise themselves with the game and develop a strategy for the second block.

Statistic	Block 1 (Baseline)	Block 2 (Robot-Assisted)
Mean	12.53	12.01
Median	12.56	12.00
IQR	2.01	2.48
Standard Deviation	1.96	2.03

Table S1. Descriptive Statistics for inflation behaviour. The inflation behaviour in block 1 represents the number of inflates made before collecting points, and in block 2, the inflation behaviour represents the number of inflates made before requesting help.

This consistency in inflation behaviour aligns with our assumption that the first block served as a familiarisation phase. Participants chose a limit they thought was safe and then requested help from the robot.

Inflation after help request

The following descriptive statistics (Table S2) provide context for the bar graph in Figure S1, which shows that there was a significant difference in inflation behaviour within groups but not much between groups.

Statistic	Inflate Non-Cust	Inflate Cust	Collect Non-Cust	Collect Cust
Mean	2.04	2.34	0.51	0.55
Median	1.90	2.05	0.27	0.33
IQR	1.54	1.28	0.25	0.73
Standard Deviation	1.07	1.16	0.65	0.58

Table S2. Descriptive Statistics for Inflation Behavior after requesting help

Risk taking behaviour

Further analysis of risk-taking behaviour in Table S3 and Figure S2 showed that participants took more risk when playing with a robot. leading to them collecting more points on average.



Figure S1. Bar graph showing the average number of inflations following 'inflate' and 'collect' requests by robots across different study conditions. Error bars indicate the Standard Error of the Mean, reflecting the variability and precision of these averages.

Statistic	Baseline	Robot-Assisted
Mean	12.53	13.47
Median	12.56	13.33
IQR	2.01	2.407
Standard Deviation	1.96	2.07

Table S3. Descriptive Statistics for Baseline and Robot-Assisted Games

Balloon Burst Behavior

The BART is a game of chance, and the balloon's inflation limit is unpredictable. Given this randomness, the burst balloon results do not directly indicate the level of trust in the robot's suggestions. However, due to the readability and in the interest of maintaining transparency, we provide the descriptive statistics of balloon burst counts here in the supplementary material for readers who wish to explore the data further.

As shown in Table S4, there is a slight increase in the mean number of burst balloons in the second block. However, burst count results do not directly reflect the level of trust in the robot. They should be interpreted with caution and in the context of the study's overall design and objectives. To avoid a type 1 error, a much larger sample size would be necessary to be confident in the observed effect.

Statistic	Baseline	Robot-Assisted
Mean	7.65	8.75
Median	7.00	8.00
IQR	3.50	5.00
Standard Deviation	3.12	3.99

Table S4. Descriptive Statistics for balloons that burst during each block of the BART played by the participants



Figure S2. Distribution of average inflations per game comparing baseline and robot-assisted conditions.