nature machine intelligence

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s42256-019-0113-5

In the format provided by the authors and unedited.

Behavioural evidence for a transparency–efficiency tradeoff in human–machine cooperation

Fatimah Ishowo-Oloko[1], Jean-François Bonnefon[2,3], Zakariyah Soroye[1], Jacob Crandall[4], Iyad Rahwan[3,5]* and Talal Rahwan[6]*

[1]Department of Computer Science, Khalifa University, Abu Dhabi, United Arab Emirates. [2]Toulouse School of Economics (TSM-R), CNRS, University Toulouse Capitole, Toulouse, France. [3]Center for Humans and Machines, Max-Planck Institute for Human Development, Berlin, Germany. [4]Department of Computer Science, Brigham Young University, Provo, UT, USA. [5]The Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA. [6]Department of Computer Science, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates. *e-mail: irahwan@mit.edu; talal.rahwan@nyu.edu

Supplementary Materials for

Behavioural Evidence of a Transparency-Efficiency Tradeoff in Human-Machine Cooperation

Fatimah Ishowo-Oloko, Jean-François Bonnefon, Zakariyah Soroye,
Jacob Crandall, Iyad Rahwan*, Talal Rahwan*

*Joint corresponding authors. E-mail: irahwan@mit.edu; talal.rahwan@nyu.edu

Supplementary Note 1. Recruitment

We conducted 5 different sessions on Amazon Mechanical Turk from December 2017 to January 2018, whereby eligibility was restricted to only participants from the US or Canada. Each participant was placed in a virtual waiting room and then matched with an associate in a round-robin manner. In particular, the first two participants in the room were paired with each other, whereas the third participant was paired with a bot; this pattern was repeated for all subsequent participants. In addition, every odd-numbered participant that entered the room was informed that his/her associate is a bot, whereas every even-numbered participant was informed that his/her associate is human.

A total of 774 participants entered the virtual room, 718 of which completed the game. As stated in the game tutorials, participants who dropped out before the end of the game were not paid. On the other hand, those who completed the game were paid a show-up fee of $2, as well as a bonus based on the total points accrued over the course of the game. After completion, participants were asked to recall the type of associate they were informed at the beginning of the experiment. 10 participants failed this manipulation test and, consequently, their results as well as those of their associates were excluded from all analysis, meaning that the manipulation test resulted in the exclusion of 20 participants in total.

Out of the 698 participants who completed the game and also passed the manipulation test, 350 were in the purported-human condition, and the remaining 348 participants were in the purported-bot condition. More specifically, in the purported-human condition, 180 participants were paired with a bot (i.e., with the S++ algorithm), and the remaining 170 participants were paired with each other (i.e., each of them was paired with another person from the same pool of 170 participants). On the other hand, in the purported-bot condition, 188 participants were paired with a bot, and the remaining 160 participants were paired with each other. This is summarized in Supplementary Table 1.

To reduce end-game effects, participants were not told how long the game would last, rather each participant played a random number of rounds between 50 and 59. Hence, the number of rounds considered in our analysis was limited to 50 rounds. As the bot can make

Supplementary Table 1: The number of participants in each condition. The row represents the type of the associate, while the column represents the information given to the participants about the type of their associate. Thus, the number of participants in the purported-bot condition is $188 + 160$, while the number of those in the purported-human condition is $180 + 170$.

		Information	
		Bot	Human
Associate	Bot	188	180
	Human	160	170

decisions almost instantaneously, we added delays to the bot in order to make it obscure to participants whether they were paired with a bot. In particular, as in prior studies [1], we used tit-for-tat delays, whereby the bot's response was delayed by the same amount of time that was taken by the participant to act in the previous round. The only restriction was that the delay for choosing the action does not exceed 20 seconds. For example, if the participant took 5 seconds to act in round 3, the bot would wait 5 seconds before acting in round 4. However, if the participant took, say, 30 seconds to act in round 3, the bot would only wait for 20 seconds before choosing its action in round 4.

Supplementary Note 2. Consent Form

At the beginning of the experiment, each participant had to read the following terms and conditions, and agree to them by clicking a button to give consent online:

> *Thank you for deciding to participate in the experiment. We are a research team at Masdar Institute and Massachusetts Institute of Technology (MIT), and we seek your participation in furthering our understanding of human behavior in games.*
>
> *You can make a reasonable amount of money from this game depending on the decisions you make. You are therefore advised to read the following instructions carefully in order to make informed decisions throughout the course of the game. Your participation in this game will take approximately 30 minutes.*
>
> *For your participation, you will be compensated with the task's standard reward of 2USD, as well as a significant bonus that depends on your performance in the experiment. These will be paid out at the end of the game, provided you follow all instructions, and play the game till the end.*
>
> *You will complete 2 stages before starting the game itself. In the first stage, you will be provided with a walkthrough of the game via screenshots of the actual game interface along with the instructions for the game. In the second stage, you will be asked 2 comprehension questions to test your understanding of the game.*
>
> - **For the purported-bot condition:** *After passing the comprehension test, you will be matched with a Bot who will be your associate throughout the course of the game. You may spend some time in a waiting room while we initialize the bot.*

- **For the purported-human condition:** *After passing the comprehension test, you will be placed in a waiting room. The game will start as soon as another MTurk worker joins the game, at which point a beep will alert you. If this does not happen within 300 seconds (there will be a timer counting down), you have the option of leaving with just the task's standard reward or you may continue waiting if you wish.*

Please confirm that you are willing to participate using the following checkbox. If you are not willing to do so, please return the assignment.

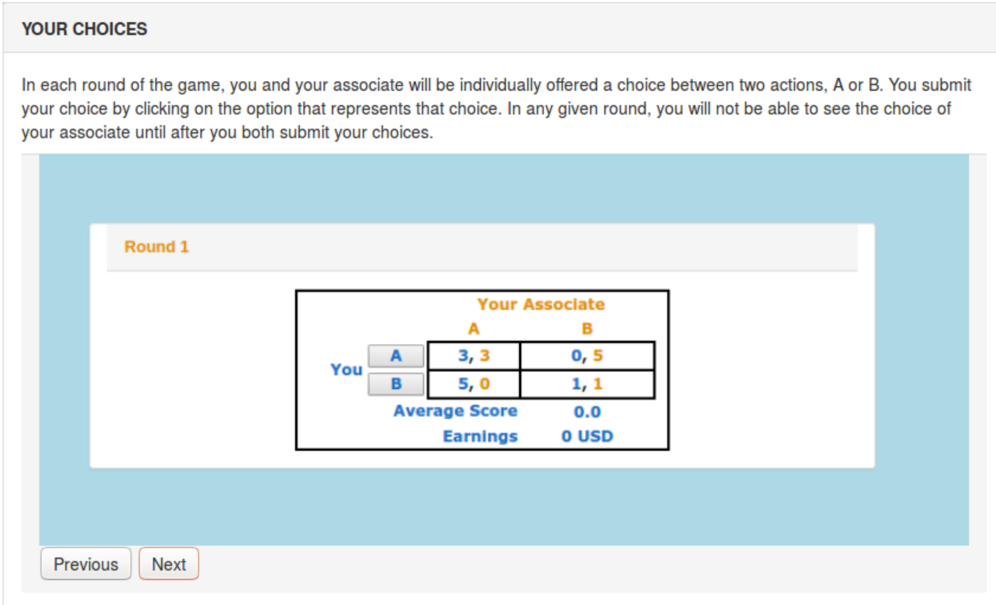
I understand that this game requires my undivided attention for up to 30 minutes, and I am willing to commit said time and attention until the game comes to an end.

Supplementary Note 3. Tutorial Interface

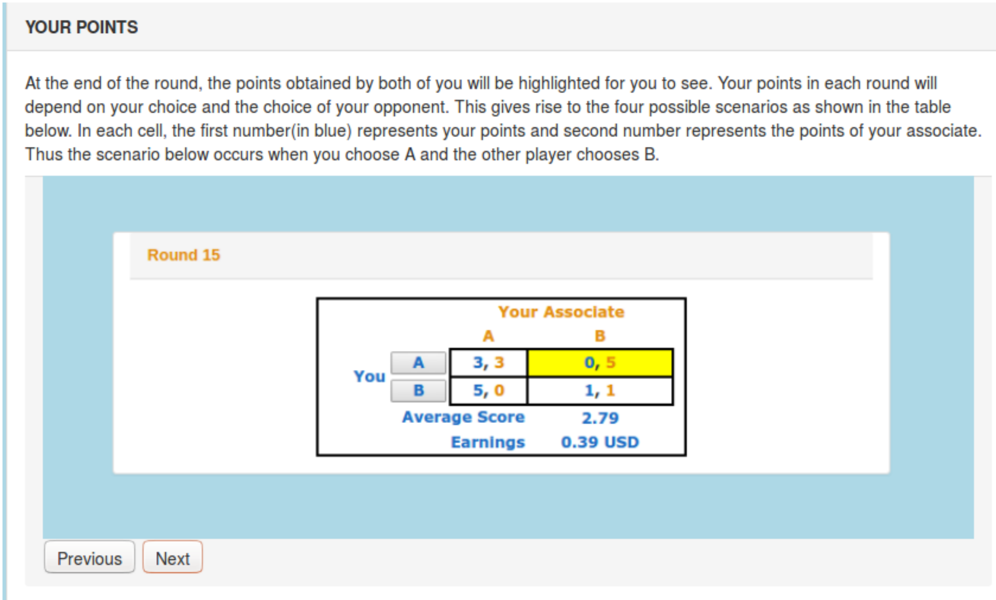
In the introduction to the game, participants in the purported-bot condition were informed that their associate was a bot, and a picture depicting a Human facing a Bot was also displayed to reinforce the information given. On the other hand, in the purported-human condition, participants were informed that their associate was a Human, and a picture depicting a Human facing another Human was displayed instead.

Screenshots of the tutorial can be found in Supplementary Figures 1 to 3. Supplementary Figure 4 presents a screenshot of the quiz, whereas Supplementary Figure 5 presents a screenshot of the manipulation test.

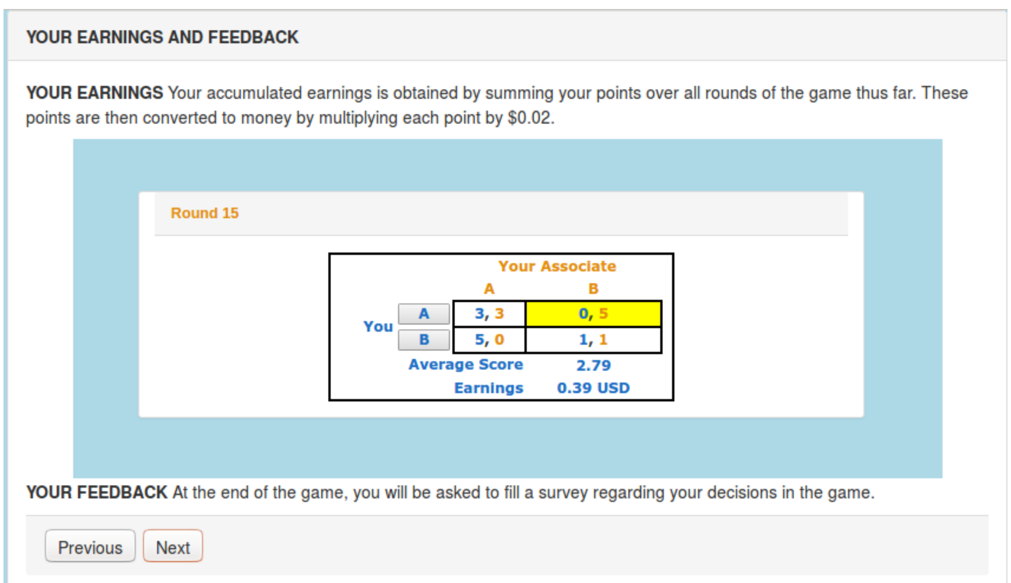
Supplementary Figures



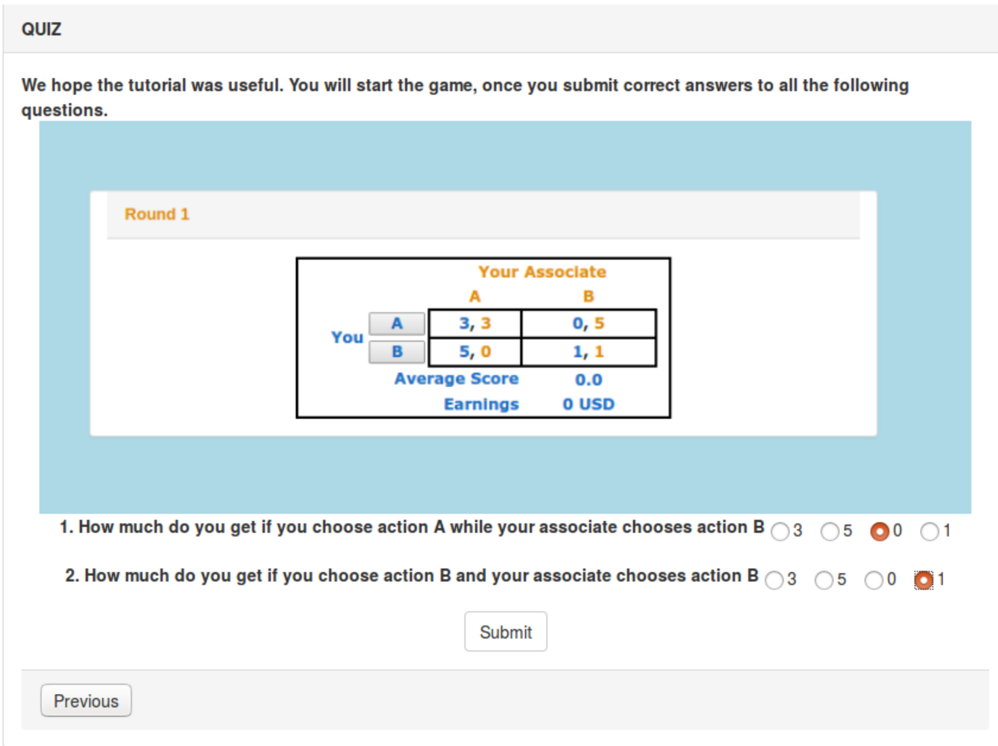
Supplementary Figure 1: Here, participants are shown their choices in the game; each participant has the same options.



Supplementary Figure 2: Here, participants are shown their choices in the previous round, as well as the choices of their associates. This is done by highlighting the corresponding cell.



Supplementary Figure 3: Here, participants are also shown their average score and the total earnings accumulated thus far in the game. The average score was obtained by dividing the total number of points obtained by the current round number, while the total earning was obtained by multiplying the total number of points by \$0.02.



Supplementary Figure 4: Participants are provided with two comprehension questions which they need to answer correctly in order to proceed to the game. This serves as a minimal check to see if they have properly understood their payoffs as a result of both their choices and their associates' choices. Participants who failed this test did not proceed to the game.

Supplementary Figure 5: At the end of the game, participants are given a manipulation test. This helps determine if participants remember the information about their associate, which was given at the beginning of the game. Participants who failed the manipulation test were excluded from all analysis of results, along with their associates.

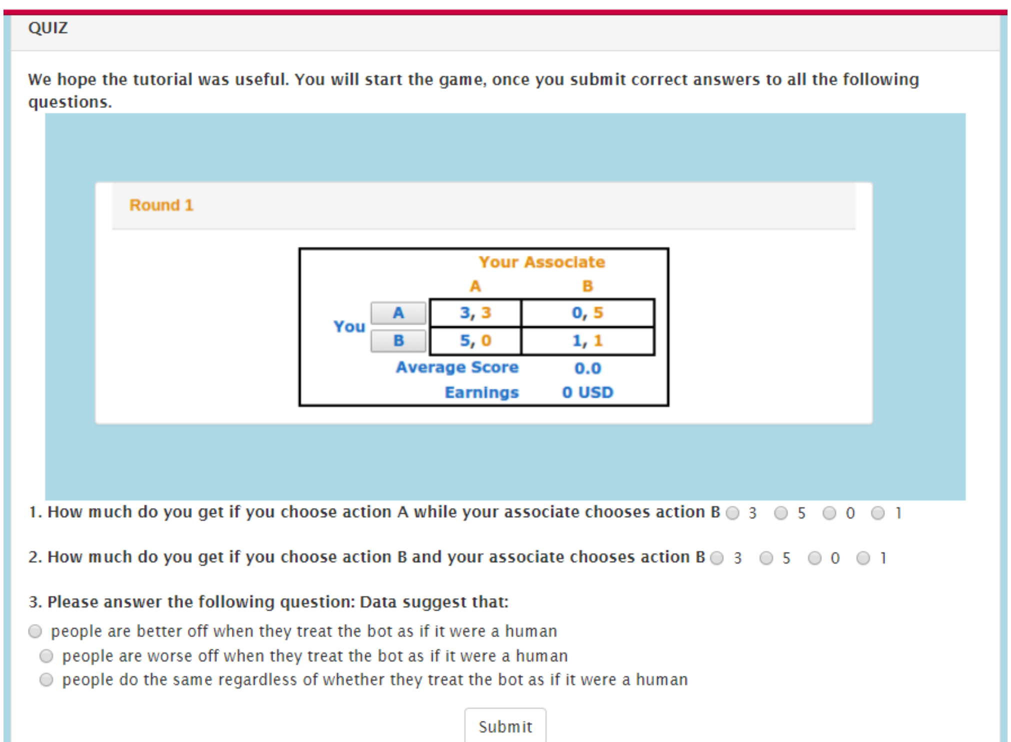
Supplementary Note 4. Debiasing treatment

We conducted an extra session on Amazon Mechanical Turk in July 2019, following the same specifications as in the main experiment. In this session, which we refer to as the debiasing condition in the main paper, participants were shown the same tutorial as participants in the purported-bot condition, with the following additional message:

> *"Data suggest that people are better off if they treat the bot as if it were a human."*

They were also asked a third comprehension question as in Supplementary Figure 6: *Please answer the following question: Data suggest that:*

- *People are better off when they treat the bot as if it were a human.*
- *People are worse off when they treat the bot as if it were a human.*
- *People do the same regardless of whether they treat the bot as if it were a human.*



Supplementary Figure 6: Participants were provided an additional comprehension question along with the two from the main experiment. Participants who failed this test did not proceed to the game.

Supplementary Note 5. S++ algorithm

In our studies, the bot followed S++ [2], a learning algorithm designed for repeated general-sum games as well as episodic general-sum stochastic games [3]. We selected this algorithm due to its success in previous studies, wherein it has been shown to be a top-performing algorithm across a wide variety of games and partners (including humans) [1]. A full description of the algorithm as it was implemented in our studies can be found in the original work by Crandall [2]. Here, we provide a brief overview of the algorithm to facilitate understanding of the results presented in this paper.

S++ is an expert algorithm. It computes or learns a set of different expert strategies, each of which behaves optimally given particular assumptions about it partners. It then learns over time which strategy from this set to follow based on its experience. Thus, S++ is defined in two parts: (1) Its set of expert strategies and (2) its mechanism for determining which strategy to follow in each round t of the repeated game.

The set of experts

Using the description of the game environment, S++ computes a diverse set of expert strategies $E = \{e_1, ..., e_k\}$. For the prisoner's dilemma game, this set of experts is summarized in Supplementary Table 2. These experts include both follower and leader strategies [4]. Each follower strategies is designed to play a best response to an associate assumed to behave according to a particular strategy. On the other hand, leader strategies play trigger-like strategies in which the bot plays its portion of a target solution (a sequence of joint outcomes) as long as its partner plays its part (each expert advocates a different target solution). If its partner deviates from this solution, it plays that strategy that minimizes its partner's maximum possible expected payoff until its partner's payoffs are lower than they would have been (on average) had it followed the target solution. It then returns to playing the target solution.

Expert Selection

S++ uses a form of aspiration learning [5] to select which expert to follow in each round t. That is, S++ encodes an aspiration level $\alpha(t)$ which it uses to (a) determine when to consider switching expert strategies and (b) to determine which expert to follow in each round. More specifically, S++ selects an expert strategy in round t, and then follows that strategy for m rounds. If that strategy produces an average payoff over those rounds that meets or exceeds the current aspiration level $\alpha(t)$, then it continues to use that expert. Otherwise, it randomly selects a new expert from the set of expert strategies whose *potential* (see Supplementary Table 2) exceeds its aspiration level. Formally, this set is defined as $E(t) = \{e_j \in E : \rho_j(t) \geq \alpha(t)\}$. Note that lower aspiration levels result in S++ selecting from among a larger set of expert strategies. Thus, in the prisoner's dilemma, S++ is less likely to select a strategy corresponding to mutual cooperation with an aspiration level of $\alpha(t) = 2.0$ (in which it would only consider selecting experts 1, 2, 3, 4, 7, 8, 9, 10, and 13; see Table 2) than an aspiration level of $\alpha(t) = 3.0$ (in which it would only consider selecting experts 1, 2, 7, 8, and 13; see Supplementary Table 2).

Supplementary Table 2: A summary of the set of experts computed by S++ [2] for the prisoner's dilemma studied in this work (see Figure 1). In Leader strategies (expert IDs 1-6), the bot plays its part of the action sequence as long as its partner plays its portion. For example, solution 1: $< (a, A), (b, A) >$ describes a target solution where player 1 alternates between cooperation a and defection b, while its associate always cooperates (A) resulting in a payoff of 4.0 to player 1. If its partner deviates from this solution, the bot punishes it until its partner's payoffs are lower than they would have been (on average) had it followed the target solution. Follower strategies play a best response to an associate assumed to behave according to a particular strategy. Experts 7-12 assume their partner is playing a leader strategy based on the corresponding target solution (hence the best response is to always play the target solution). Expert 13 assumes its partner is playing the static strategy $\hat{\pi}_{-i}$ and Expert 14 assumes its partner is an adversary trying to attack it (and therefore plans for the worse case). The *Potential* of each strategy denotes the expected payoff of the bot when its partner behaves as desired (in the case of leader strategies) or expected (in the case of follower strategies).

ID	**Type**	**Strategy Specification**	**Potential**
1	Leader	Target sequence: (a, A), (b, A)	4.0
2	Leader	Target sequence: (a, A)	3.0
3	Leader	Target sequence: (b, A), (a, B)	2.5
4	Leader	Target sequence: (a, A), (b, B)	2.0
5	Leader	Target sequence: (a, A), (a, B)	1.5
6	Leader	Target sequence: (b, B)	1.0
7	Follower	Target sequence: (a, A), (b, A)	4.0
8	Follower	Target sequence: (a, A)	3.0
9	Follower	Target sequence: (b, A), (a, B)	2.5
10	Follower	Target sequence: (a, A), (b, B)	2.0
11	Follower	Target sequence: (a, A), (a, B)	1.5
12	Follower	Target sequence: (b, B)	1.0
13	Follower	Model-based RL	$r_i\left(br(\hat{\pi}_{-i}), \hat{\pi}_{-i}\right)$
14	Follower	Maximin strategy	1.0

The aspiration level $\alpha(t)$ is updated over time. Initially it is set high, which causes S++ to only select experts that could potentially meet this aspiration level. In the implementation of S++ used in the study, $\alpha(0)$ is set to the payoff it receives in the Nash bargaining solution. The aspiration level is then updated after each m rounds as follows:

$$\alpha(t+m) \leftarrow \lambda^m \alpha(t) + (1 - \lambda^m)\mathrm{R}, \tag{1}$$

where R is the average payoff obtained in the last m rounds. We used $\lambda = 0.96$ in this study. In effect, S++ performs a relaxation search until it finds a strategy that consistently meets its aspiration level.

References

[1] Crandall, J. W. *et al.* Cooperating with machines. *Nature communications* **9**, 233 (2018).

[2] Crandall, J. W. Towards minimizing disappointment in repeated games. *Journal of Artificial Intelligence Research* **49**, 111–142 (2014).

[3] Crandall, J. W. Robust learning in repeated stochastic games using meta-gaming. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence* (2015).

[4] Littman, M. L. & Stone, P. Leading best-response strategies in repeated games. In *IJCAI workshop on Economic Agents, Models, and Mechanisms* (Seattle, WA, 2001).

[5] Karandikar, R., Mookherjee, D., R., D. & Vega-Redondo, F. Evolving aspirations and cooperation. *Journal of Economic Theory* **80**, 292–331 (1998).