

Evidence for strategic cooperation in humans

Maxwell N. Burton-Chellew,^{1,2} Claire El Mouden¹ & Stuart A. West^{1,2}

¹Department of Zoology, University of Oxford, OX1 3PS

²Calleva Research Centre for Evolution and Human Sciences, Magdalen College, Oxford, OX1 4AU

List of Electronic Supplementary Material Files

File 1

ESM Materials & Methods & Results

File 2

ESM Table 1

File 3

Fig. ESM 1. Overview of the experiment. The experiment was conducted in three stages. Stage one identified if players were rational or not. Diagnosis was based on if players cooperated with computers. Stage two was a repeated public-goods game with constant groups but no information between rounds. Only some players knew that their behavior was invisible to their groupmates. Stage three was another repeated public-goods game with new constant groups and information between rounds that either revealed a player's behavior to their groupmates fully, or partly (via payoffs), or not. Results show levels of initial cooperation.

File 4

Fig. ESM 2. Cooperation with computers. The figure shows the distribution of the level of contributions in the public goods game (cooperation) towards computerized groupmates. Players were told that their groupmates were computers that would be choosing contributions randomly. Contributions above 0 MU made a loss and therefore cannot be rationalized, in the context of this game, with either selfish or prosocial preferences.

File 5

Figure. ESM 3. Strategic cooperation when partly versus fully visible. The figures show the distribution of the level of contribution in a public goods game (cooperation), distinguishing between individuals who (a) had not cooperated with computers (rational players) or (b) had previously cooperated with computers (irrational players). The light grey bars are for individuals who knew their groupmates would be told their earnings and thus could infer the average cooperation of their groupmates (partly invisible). The dark grey bars are for individuals who knew their individual behavior would be revealed to their groupmates (fully visible).

File 6

Figure ESM 4. Analyzing players that pass the control questions gives the same results. The figures show the behavior of only those players that passed the control questions. Part (a) shows that among players that passed control

questions, the rational players cooperate at higher levels when their behavior is visible.

File 7

Fig. ESM 5. Rational players facing uncertainty are strategic cooperators. The figures show the distribution of the level of contribution in a public goods game (cooperation), distinguishing between individuals who (a) had not cooperated with computers (rational players) or (b) had previously cooperated with computers (irrational players). The light grey bars are for individuals who knew for certain their behavior would not be revealed to their groupmates (invisible), and the dark grey bars are for individuals whose visibility was ambiguous because they were not told what information their groupmates would receive (potentially visible). Part (a) shows that rational players cooperate at higher levels when their behavior is potentially visible.

File 8

Fig. ESM 6. Strategic cooperation over time. This figure examines the behavior of players in a repeated public goods games with no information between rounds. The figure distinguishes between individuals who (a) had not cooperated with computers (rational players) or (b) had previously cooperated with computers (irrational players). The light grey bars are for individuals who knew their behavior would not be revealed to their groupmates (invisible), and the dark grey bars are for individuals whose visibility was ambiguous because they were not told what information their groupmates would receive (potentially visible). Dashed lines are 95% confidence intervals.

Strategic cooperation in humans: Electronic Supplementary Material

Maxwell N. Burton-Chellew,^{1,2} Claire El Mouden¹ & Stuart A. West^{1,2}

¹Department of Zoology, University of Oxford, OX1 3PS

²Calleva Research Centre for Evolution and Human Sciences, Magdalen College, Oxford, OX1 4AU

ESM Materials & Methods

Participants

The data presented here were collected at the Centre for Experimental Social Science (CESS) Oxford as part of a larger experiment (24, 38). CESS recruited participants using ORSEE (39) and obtained participant consent forms. Experiments were conducted using z-Tree (40). We tested 288 participants (151 females, 137 males, mean self-reported age = 27 years, range = 18-74, S.D. = 11.97) across 12 sessions of 24 individuals.

For the purposes of best practice we report the following details (43). The overall sample size was as large as possible given the costs of the experiment and the value of a small-project grant from the John Fell Fund Oxford (maximum £7,500). No participants or sub-groups have been omitted from our analyses. All participants completed a questionnaire after the experiment recording their gender, age, and a short 10-item personality test (42). The primary purpose of this questionnaire was to occupy the participants while we arranged their cash payments. We have not analyzed any of these demographic variables, as they formed no part in our a priori hypotheses, although we may use them later for exploratory analyses.

Game parameters

In each round, we gave each individual 20 Monetary Units (MU) and made them contribute 0-20 MU to a group fund. We multiplied the group fund by an efficiency factor of 1.6 before sharing it among the four group members equally, regardless of differences in contribution. Therefore the marginal per capita return on each contributed MU was 0.4. The strategy that would maximize income in a single round (payoff dominant strategy) was therefore to contribute 0 MU. MUs from each round were banked and non-transferrable between rounds. The unique equilibrium for all rounds is also to contribute 0 MU. This can be deduced for a population of rational self-interested individuals that each assumes everyone to be rational, via backward induction from rational play in the final round (6). However, it can still be rational for a selfish individual to cooperate before the final round if individuals are uncertain about the rationality or selfishness of their groupmates and can be influenced in some way (6).

We gave our participants a copy of standard instructions and control questions routinely used in an attempt to ensure comprehension (5, 41). However, in contrast to other studies, we did not force our participants to correctly answer all questions before progressing to the experiment. We analyze all participants and then repeat our main analyses on the subset of participants that only answered all questions correctly. The results are qualitatively the same (ESM Results). Our experiment was conducted in three stages: first playing

computerized groupmates with no information, second playing human groupmates with no information, and third playing human groupmates with information between rounds that either allowed or prevented strategic cooperation.

No information game

Another advantage of the no information game is that by comparing the behavior of individuals in this no information setting, depending on how they are later on randomly assigned to either visible or invisible treatments, we can test if our treatment differences are due to different effects or just different participant pools. We avoided spillover effects because individuals cannot learn about either game or about other players in the no information treatment, and because all individuals play this no information treatment, it cannot differentially affect the later treatments.

Information games

In our invisible treatment, for the purposes of another study, we told 48 players the following:

Let us call this other group, Group A, and call your group, Group B.

All the players in both Group A and B will only receive information about the decisions in Group A.

Nobody in either group will learn anything about Group B.

Everybody in your group will only ever see information about the other group.

And we told 144 of these 192 participants:

The people in your group will receive information about 3 other groups. Let us call these Groups A, B and C.

Your group will receive information describing the decisions within these groups.

Your group will not receive any information about players in your own group.

Specifically, your group will be told

[and]

what the average decisions were in each group.

[or] what the decisions were of the player that made the most money in each of groups A, B, and C.

[or] what the decisions were of the three players that made the most money out of all the players in Groups A, B, and C combined.

How individuals responded to these different sources of social information over time is the focus of another study comparing how people respond to information on common or successful behaviours (38). We combined the data from these four sub-treatments into one 'invisible' treatment ($N=4 \times 48=192$, randomly assigned).

Statistical methods

We compared average levels of cooperation using the generalized linear model (GzLM) function in IBM SPSS version 22. We ran our models with a binary-logistic link set to 20 trials (because contributions were bounded between 0 and 20 MU). We set the parameter estimation method to hybrid; the scale parameter method to Pearson chi square, and the covariance matrix used a model-based

estimator. The chi square statistics used the likelihood ratio and the confidence interval type was profile likelihood. All other settings used the default. We compared levels of cooperation over time using the generalized linear mixed model (GzLMM) function to control for pseudo replication arising from repeated measures on the same individuals over time. We fitted random intercepts for each participant. We estimated degrees of freedom with the more conservative option to reduce false positives (Satterthwaite method), and all other options were set to the default. We compared the frequencies of cooperators using Fisher's exact test (FET) and all significance values are two tailed calculations.

ESM Results

Irrational players

For the distribution of cooperation towards computers see Fig. ESM 2.

Cooperation when partly visible versus fully visible

For the distribution of cooperation towards humans, when either partly or fully visible in the games with information, see Fig. ESM 3. We aimed to compare the cooperation of rational players across treatments. However the rarity of rational players meant that we only had limited samples to compare rational behaviour in the fully visible (N = 20) and partly visible (N = 23) treatments. Therefore for most analyses we combine the data from the partly visible and fully visible treatments, because in both treatments players are told their behavior will be visible, in some form, and have the potential to influence their groupmates' behavior.

For completeness, here we compare behavior in the partly and fully visible treatments anyway. Rational players significantly varied their cooperation across all three treatments (GzLM: $F_{2,100} = 9.0$, $P < 0.001$) whereas irrational players did not (GzLM: $F_{2,182} = 0.8$, $P = 0.430$). However, the difference in cooperation levels of both rational and irrational players in the partly and fully visible treatments was not significant. Rational players cooperated at 46% (9.2 ± 1.88 MU, N = 23) when partly visible and 54% (10.8 ± 1.96 MU, N = 20) when fully visible, a non-significant difference (GzLM: $F_{1,41} = 0.3$, $P = 0.568$). Irrational players cooperated at 46% (9.2 ± 1.47 MU, N = 25) when partly visible and 55% (11.1 ± 1.02 MU, N = 28) when fully visible, a non-significant difference (GzLM: $F_{1,51} = 1.2$, $P = 0.284$).

Strategic cooperation & robustness checks

The increased cooperation among rational players when visible was due to both more players cooperating (28/43, versus 18/60, FET: $P = 0.0006$) and significantly higher levels of cooperation from those that cooperated (mean cooperation of rational players that contributed > 0 when invisible = 55% [11.0 ± 1.5 MU]; when visible = 76% [15.2 ± 1.17 MU]; GzLM: $F_{1,44} = 4.6$, $P = 0.039$, $\beta = 1.0 \pm 0.45$, Fig. 2a). Among irrational players, the frequency of players that cooperated with humans (contributed > 0 MU) did not significantly differ between the visible and invisible treatments either (117/132 versus 50/53, Fisher's Exact Test: $P = 0.2852$).

For robustness we repeated the main analyses in the text on only those players that passed all the control questions ($N = 94$ of 288, 33%). We found the same qualitative results, with rational players nearly quadrupling their cooperating when visible (54% versus 14%), shifting their modal cooperation from 0% to 100%. In contrast, irrational players did not significantly adjust their behavior (48% versus 41% respectively). These differences lead to a significant interaction between how individuals behaved with the computer (rational versus irrational), and if their behavior was visible or not to their groupmates (GzLM rationality*visibility: $F_{1,90} = 5.4$, $P = 0.022$, ESM results Fig. ESM4).

We also confirmed that our rational players assigned to visible treatments were not merely of a more cooperative nature. We were able to do this, by examining their behavior in the prior game played with no information between rounds, before they were assigned to either the invisible or visible treatments in the later game with information. If the two groups of players had different levels of 'baseline' cooperation, then we would have detected it here, but we did not.

Across all six rounds the levels of cooperation were similar between those rational individuals that would later on be assigned to visible or invisible treatments (GzLM comparing cooperation in round 1 of the no-information game depending on what treatment player was assigned to later on in the game with information: R1: $F_{1,101} = 1.2$, $P = 0.266$; R2: $F_{1,101} = 1.3$, $P = 0.251$; R3: $F_{1,101} = 0.0$, $P = 0.926$; R4: $F_{1,101} = 0.5$, $P = 0.482$; R5: $F_{1,101} = 0.1$, $P = 0.734$; R6: $F_{1,101} = 0.9$, $P = 0.336$).

We tested how rational and irrational players changed their cooperation when moving from the no-information game to the game with information, depending on if they had been assigned to the visible or invisible treatments. This way we compare their response to learning that their behavior will be visible, whilst controlling for how they as individuals behaved previously when they were invisible in the no-information game. This basically replicates our main analysis but controls for an individual's baseline cooperation. Overall, the increase in cooperation from the end of the no-information game to the start of the games with information depending on visibility was significantly greater for rational players (GLM, testing interaction between individual change in cooperation level ~ rationality*visibility: $F_{1,284} = 7.7$, $P = 0.006$). Specifically, rational players assigned to a visible treatment increased their cooperation from 15% (3.0 ± 1.00 MU) to 50% (9.9 ± 1.35 MU) whereas those assigned to the invisible treatment only increased from 10% (1.9 ± 0.67 MU) to 17% (3.3 ± 0.79 MU). This meant there was a significant interaction for rational players between visibility and round (end of first game or start of second game) (GzLMM: $F_{1,112} = 9.6$, $P = 0.003$). In contrast, this same interaction was not significant for irrational players, who increased from 43% (8.6 ± 1.00 MU) to 51% (10.2 ± 0.88 MU) when restarting into the visible treatment, and similarly from 43% (8.6 ± 0.61 MU) to 47% (9.3 ± 0.58 MU) if assigned to the invisible treatment (GzLMM: $F_{1,144} = 0.8$, $P = 0.369$).

Finally, we tested if *how much* more than 0 MU a player cooperated with a computer inversely predicted how strategic they were (thank you to a reviewer

for this suggestion). For each player we calculated how much they increased their cooperation from the end of the no information game in stage 2 to the start of the information game in stage 3, depending on if assigned to a visible treatment or not. We then regressed this 'strategic response' against how much individuals cooperated with computers to test if there was a negative correlation, which may suggest that players that cooperate more with computers understand the game less or are less strategic than those that cooperate a little with computers. However there was no significant correlation for irrational players assigned to a visible treatment (linear model: $F_{1,51} = 1.5$, $P = 0.231$, $\beta = -0.13 \pm 0.108$).

Strategic decline

A perfect strategic cooperator would initially cooperate and then decrease their cooperation without ever increasing it. We therefore analyzed how individuals changed their cooperation over time (Table ESM1). We took an individual's cooperation for each of the six rounds in each of the no-information game and the information game, depending on the visibility of their behavior and if they were irrational or not (had previously cooperated with computers or not). We considered four mutually exclusive patterns of individual behavior: (1) a player's cooperation decreased at least once and never increased; (2) conversely, a player's cooperation increased at least once and never decreased; (3) a player's cooperation never changed; and (4) a player's cooperation both increased at least once and decreased at least once. The first pattern of behavior is most in keeping with ideas of strategic cooperation, as it requires a player to initially cooperate and then to decrease their cooperation without ever increasing it.

We found that the first behavioural pattern, whereby an individual initially cooperates and then decreases it without ever increasing it, was most common in rational players in the visible treatments. Overall, around 25% of rational players showed this behavior when visible (24% when potentially visible in the no-information game, and 26% when visible in the games with information). In contrast only 10% of rational players behaved this way when they were invisible. Although these differences are not significant on their own, possibly due to small sample sizes, they are when combined. Specifically, 24% of rational players (20/82) conformed to this pattern of initially cooperating and then reducing but never increasing their cooperation when potentially visible in the no information game, versus just 10% (2/21) when invisible (FET significance value = 0.2311). Likewise, in the games with information, 26% of rational players (11/43) behaved this way when visible and only 10% (6/60) when invisible (FET significance value = 0.0576). Pooling these together gives 25% (31/125) versus 10% (8/81, FET = 0.0064).

Consistent with the above, among rational players, we found they were significantly less constant when visible in the information game (16 v 58%, FET significance value = 0.0001) and less constant, although not significantly so, in the no-information game (50 v 62%, FET = 0.4633). Rational players, when visible or potentially visible, show more changes in behavior, suggesting strategic motivations. In contrast, we found that irrational players were very

volatile, with 69-74% of them both increasing and decreasing their cooperation at least once, even in the no-information versions of the game.

Finally we tested how many players could be classified as potential strategic cooperators, cooperating in the first round of the game (contribute > 0 MU) but then completely free riding in the final round (contribute 0 MU) (Table ESM1). Rational players were more likely to show this pattern of behavior in the visible treatments. In the no-information game, 32% (26/82) of rational players exhibited this behavior when partly visible, but only 10% (2/21) when invisible (FET: $P = 0.0539$). In the games with information, 49% (21/43) of rational players exhibited this behavior when visible, but only 18% (11/60) when invisible (FET: $P = 0.0013$).

Table ESM1: An individual level analysis of patterns of cooperation over six rounds of the public goods game depending on rationality and treatment visibility.

| Rationality | Information | Visibility | N | Only decrease¹ | Never change² | Only increase³ | Volatile⁴ | PSC⁵ |
|--------------------|--------------------|-------------------|----------|----------------------------------|---------------------------------|----------------------------------|-----------------------------|------------------------|
| Rational | No-information | Invisible | 21 | 10% | 62% | 5% | 24% | 10% ⁶ |
| | | Ambiguous | 82 | 24% | 50% | 1% | 24% | 32% |
| | Information | Invisible | 60 | 10% | 58% | 2% | 30% | 18% |
| | | Visible | 43 | 26% | 16% | 2% | 56% | 49% |
| Irrational | No-information | Invisible | 51 | 4% | 16% | 6% | 75% | 12% |
| | | Ambiguous | 134 | 10% | 17% | 4% | 69% | 16% |
| | Information | Invisible | 132 | 22% | 8% | 1% | 70% | 35% |
| | | Visible | 53 | 17% | 8% | 6% | 70% | 36% |

¹ A player's cooperation decreases at least once and never increases.

² A player's cooperation is constant.

³ A player's cooperation increases at least once and never decreases.

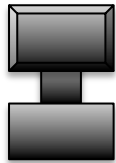
⁴ A player's cooperation both increases at least once and decreases at least once.

⁵ Potential strategic cooperation: a player's cooperation was greater than 0 in round 1 and equalled 0 in final round.

⁶ Values within a row may sum to more than 100% due to rounding approximations.

Stage 1

Play with computers



Did not cooperate

Cooperated

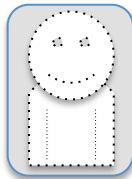
Rational players

Irrational players

Stage 2

Not visible

Maybe visible



Hypothesis

Cooperation →

Results: cooperation with humans

Rational
14% 32%

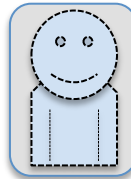
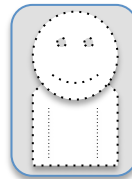
Irrational
53% 49%

Stage 3

Not visible

Partly visible

Fully visible



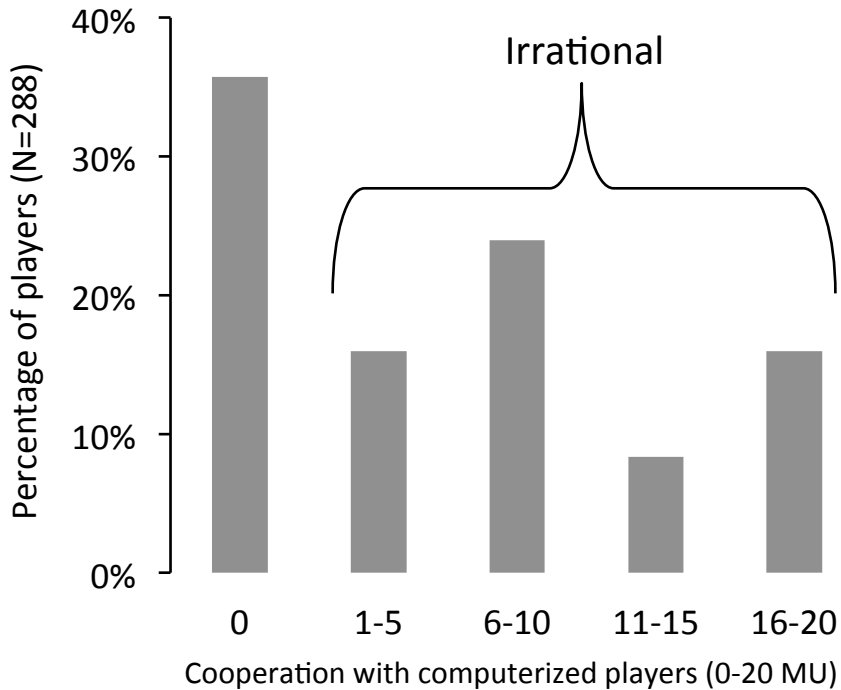
Hypothesis

Cooperation →

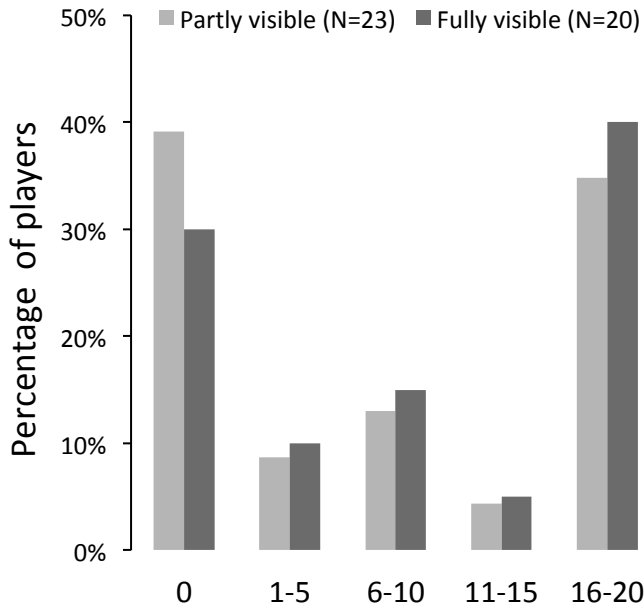
Results: cooperation with humans

Rational Players
17% 46% 54%

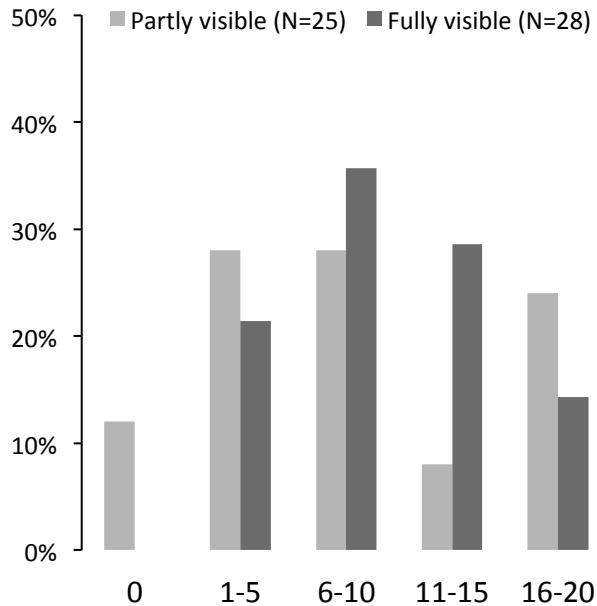
Irrational Players
47% 46% 55%



**(a) did not cooperate with computers
(Rational)**



**(b) cooperated with computers
(Irrational)**



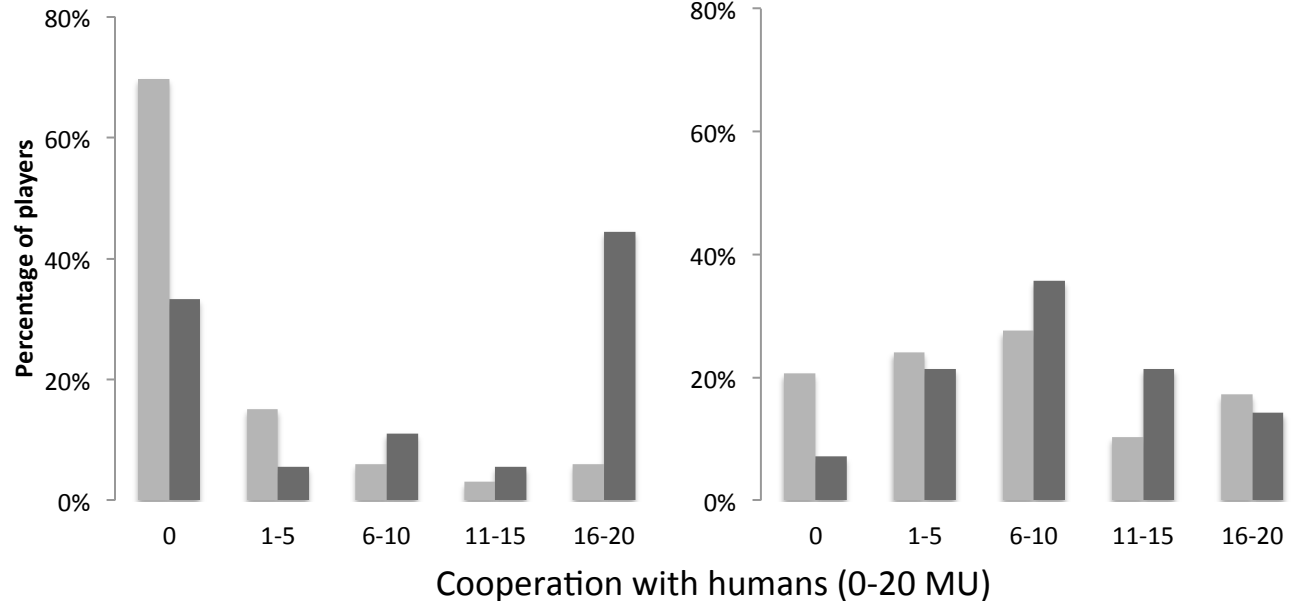
Cooperation with humans (0-20 MU)

**(a) did not cooperate with computers
(Rational)**

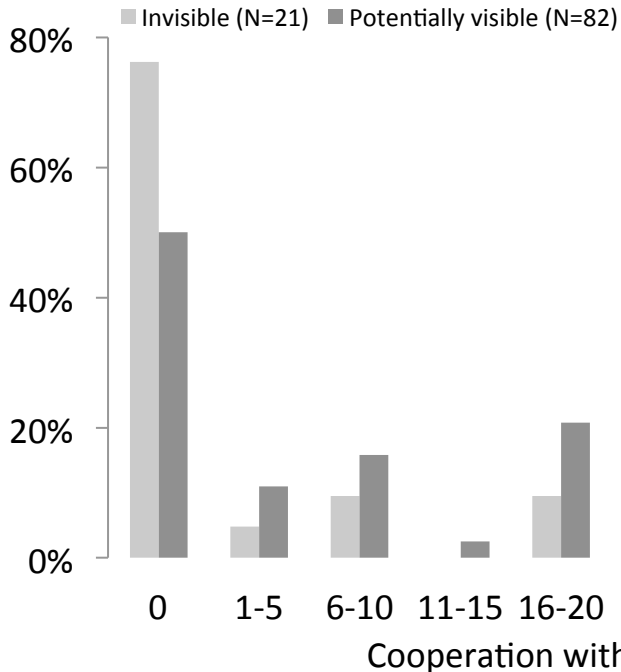
**(b) cooperated with computers
(Irrational)**

■ Invisible (N=33) ■ Visible (N=18)

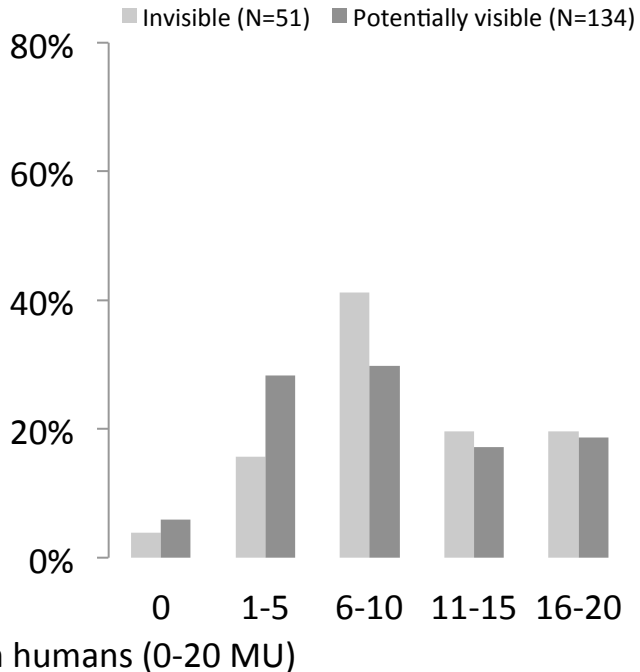
■ Invisible (N=29) ■ Visible (N=14)



**(a) did not cooperate with computers
(Rational)**



**(b) cooperated with computers
(Irrational)**



○ invisible (irrational, N = 51)

● potentially visible (irrational, N=134)

□ invisible (rational, N =21)

■ potentially visible (rational, N =82)

