Gender Differences in the Cost of Corrections in Group Work

Yuki Takahashi*

Click here for the latest version

November 30, 2021

Abstract

Corrections among colleagues are an integral part of group work, but people may take corrections as personal criticism, especially corrections by women. I study whether people dislike collaborating with someone who corrects them and more so when that person is a woman. People, including those with high productivity, are less willing to collaborate with a person who has corrected them even if the correction improves group performance. Yet, people respond to corrections by women as negatively as by men and believe that women's ability is as good as men's. These findings suggest that although women do not receive a stronger negative reaction, correcting colleagues is costly both for individuals and groups and that negative response to corrections is irrational.

JEL codes: J16, M54, D91, C92
Keywords: correction, collaboration, group work, gender differences

---

*Department of Economics, University of Bologna. Email: yuki.takahashi2@unibo.it. I am grateful to Maria Bigoni, Siri Isaksson, Bertil Tungodden, Laura Anderlucci, and Natalia Montinari whose feedback was essential for this project. I am also grateful to participants of the experiment for their participation and cooperation. Sonia Bhalotra, Francesca Cassanelli, Alessandro Castagnetti, Mónica Costa-Dias, Seda Ertaç, Valeria Ferraro, Lenka Fiala, Ria Granzier-Nakajima, Silvia Griselda, Annalisa Loviglio, Yoko Okuyama, Monika Pompeo, Øivind Schøyen, Vincenzo Scrutinio, Erik Ø. Sørensen, Ludovica Spinola, Florian Zimmermann, and PhD students at the NHH and the University of Bologna all provided many helpful comments. This paper also benefited from participants' comments at the Annual Southern PhD Economics Conference, Brazilian Meeting in Family and Gender Economics, CSQIEP, EALE Conference, ESA Conference, ESA Job Market Seminar, FROGEE Workshop, Gender Gaps Conference, Irish Postgraduate and Early Career Economics Workshop, PhD-EVS, Stanford Institute for Theoretical Economics, TIBER Symposium, Warwick Economics PhD Conference, WEAI Conference, Webinar in Gender and Family Economics, Young Economists' Meeting, and seminars at Ca' Foscari University, Catholic University of Brasília, the NHH, Tilburg University, the University of Bologna, and the University of Copenhagen. Ceren Ay, Tommaso Batistoni, Philipp Chapkovski, Sebastian Fest, Christian König genannt Kersting, and oTree help & discussion group kindly answered my questions about oTree programming; in particular, my puzzle code was heavily based on Christian's code. Michela Boldrini and Boon Han Koh conducted the quasi-laboratory experiments ahead of me and kindly answered my questions about their implementations. Lorenzo Golinelli provided excellent technical and administrative assistance. This study was pre-registered with the OSF registry (https://osf.io/tgyc5) and approved by the IRB at the University of Bologna (#262643).

1 Introduction

Receiving corrections from colleagues is an integral part of group work. Consider academic research. From the development of ideas to the writing up of the final draft, researchers discuss their research project with their colleagues, receive criticisms, and refine the ideas and the analysis. However, people may take the corrections personally. Imagine a researcher presents their paper at a seminar/submit to a journal for which they spent several years, and someone/referee points out a possible flaw in their identification assumption or their experimental design. Since the validity of identification assumptions and experimental designs are debatable, they may take it as a personal criticism.

Those people may express their discomfort in some way, and a least aggressive way to do so is not to collaborate. However, not being invited for collaboration could be detrimental to one's career success because having collaborations is essential in academia, where people co-author the majority of papers (Jones 2021; Wuchty, Jones, and Uzzi 2007).

Women's corrections may receive stronger negative reactions because people often use double standards for women and men. Evidence suggests that men undervalue women when they criticize them (Sinclair and Kunda 2000) and that people punish women more harshly when they make mistakes (Sarsons 2019) and commit misconduct (Egan, Matvos, and Seru 2021). If women who correct their colleagues lose collaboration opportunities more than men, they face a higher hurdle in their career success. It is also detrimental to group efficiency as group members cannot fully benefit from female colleagues.

This paper studies whether people dislike collaborating with someone who corrects them and more so when that person is a woman. Answering this question using secondary data poses two challenges. First, group formation is not random and group corrections are endogenous. Second, different corrections are not necessarily comparable to each other.

Thus, I design a quasi-laboratory experiment, a hybrid of physical laboratory and online experiments, where group formation is randomized and define corrections such that researchers can track its quality mathematically. Specifically, participants are allocated to a group of eight and solve one joint task with each group member one by one. Each time participants finish the task, they state whether they would like to collaborate with the group member with whom they have just solved the task for the same task in the next stage, which is the main source of earnings. This gives a strong incentive for participants to select as good a collaborator as possible. The order of the group members with whom participants solve the task is randomized. As a joint task, I use Isaksson (2018)'s number-sliding puzzle, which allows me to calculate an objective measure of each participant's contribution to the joint task as well as to classify each move as good or bad. I define a correction as reversing a group member's move, which is comparable across different participants and can be classified as either good or bad.

I find that participants correctly understand the notion of good and bad moves; that is, the higher your contribution is to solving the puzzle, the more likely it is that you will be asked to join a team. This is in line with what one would expect and validates my experimental design.

Nonetheless, after controlling for the contribution, people are significantly less willing to collaborate with a person who has corrected their moves, even if the corrections move the puzzle closer to the solution. Although it may not be so costly to correct colleagues if only low productivity people respond negatively to corrections, high productivity people also respond negatively to corrections. High productivity people's negative response also suggests that the negative response is irrational: they should be able to identify better corrections that improve group performance.

Yet, people respond to corrections by women as negatively as by men; although I find suggestive evidence that men respond more negatively to women's good corrections, the finding is not robust. These findings are unlikely to be due to people's belief about differences in women's and men's abilities: women and men contribute equally to the puzzle, and neither women nor men underestimate women's contribution.

This paper's contribution is twofold. First, it introduces correction in the contribution of ideas and examines its cost on women and groups. Coffman (2014) finds that women are less likely to contribute their ideas to the group in a male task due to self-stereotyping and Gallus and Heikensten (2019) find that debiasing their self-stereotyping by giving an award for their high ability increases women's contribution of their ideas: they put women's idea further ahead of men without involving open correction of their group member. Indeed, women correct others less often than men: Isaksson (2018), who first uses this puzzle, finds that men are more likely to correct their group member's bad moves. However, on some occasions, the contribution of ideas has to be made openly, for example, in academic seminars and business meetings. In such cases, group members' response plays an important role in the effectiveness of the intervention, and indeed there is evidence that people respond to women's ideas less favorably. For example, Coffman, Flikkema, and Shurchkov (2021) find that group members are less likely to choose women's answers as a group answer in male-typed questions. Corroborating this, Guo and Recalde (2020) find that group members correct women's ideas more often than men's ideas. Further, Dupas et al. (2021) find that female economists receive more patronizing and hostile questions during seminars.

Second, this paper shows that people avoid collaborating with someone who corrected them – or who gave negative feedback – even if doing so reduces their expected payoff. This finding links my paper with studies on self-image concerns and information avoidance. For example, Kőszegi (2006) finds that people avoid a difficult task when it reveals their ability. Corroborating this, Castagnetti and Schmacker (2021) find people select information less informative about their ability, and Ewers and Zimmermann (2015) find people exaggerate their ability when others observe it even at the cost of reducing their payoff. People even distort information about others: Kennedy and Pronin (2008) find that people view others who disagree with them as biased, and Ronayne and Sgroi (2019) find that people often stick to their decisions even if others with higher abilities suggest otherwise. Regarding gender, Chakraborty and Serra (2021) find that people respond more aggressively to female managers' messages, a type of negative feedback. Thus, although investigating the mechanism of my finding is beyond the scope of this paper, these studies point that a likely mechanism is motivated reasoning (Kunda 1987): receiving good corrections is negative feedback and accepting them as correct damages people's self-image.[1]

1. Which means θ_i in the theoretical model in section 4 (equation 1) is not exogenous.

2 Experiment

Introducing a quasi-laboratory format I run the experiment in a quasi-laboratory format where we experimenters connect us to the participants via Zoom throughout the experiment (but turn off participants' camera and microphone except at the beginning of the experiment) and conduct it as we usually do in a physical laboratory, but participants participate remotely using their computers. Online Appendix section A discusses the pros and cons of the quasi-laboratory format relative to physical laboratory and standard online experiments.

Group task As the group task, I use Isaksson (2018)'s puzzle, a sliding puzzle with eight numbered tiles, which should be placed in numerical order within a 3x3 frame (see Figure 1 for an example). To achieve this goal, participants play in pairs, alternating their moves. This puzzle has nice mathematical properties: I can define the puzzle difficulty and classify a given move as either good or bad by the Breadth-First Search algorithm. From the number of good and bad moves one makes, I can calculate individual contributions to the group task; I measure it by net good moves, the number of good moves minus the number of bad moves an individual makes in a given puzzle.

I can also determine the quality of corrections of different participants objectively and comparably.[2] Further, the puzzle-solving captures an essential characteristic of teamwork in which two or more people work towards the same goal (Isaksson 2018), but the quality of each move and correction is only partially observable to participants (but fully observable to the experimenter).

At each stage of the puzzle, there is only one good strategy which is to make a good move and one bad strategy which is to make a bad move.[3] There can be more than one good and bad move, but different good/bad moves are equal. There is no path dependence either: the history of the puzzle moves does not matter.

At the beginning of each part, participants must answer a set of comprehension questions to make sure they understand the instructions.[4]

2.1 Design and procedure

Registration

Upon receiving an invitation email to the experiment, participants register for a session they want to participate in and upload their ID documents as well as a signed consent form.[5]

2. The difficulty is defined as the number of moves away from the solution, a good move is defined as a move that reduces the number of moves away from the solution, and a bad move is defined as a move that increases the number of moves away from the solution.

3. This is conditional on that both players are trying to solve the puzzle; I show in section 8 that the results are robust to exclusion of puzzles where either player might not be trying to solve the puzzle.

4. I do not tell participants that they can correct others to reduce experimenter demand effects.

5. I recruit a few more participants than I would need for a given session in case some participants would not show up to the session.

Pre-experiment

On the day and the time of the session they have registered for, participants enter the Zoom waiting room.[6] They receive a link to the virtual room for the experiment and enter their first name, last name, and their email they have used in the registration. They also draw a virtual coin numbered from 1 to 40 without replacement.

Then I admit participants to the Zoom meeting room one by one and rename them by the first name they have just entered. This information is necessary to match up their earnings in this experiment and their payment information stored in the laboratory database, so participants have a strong incentive to provide their true name and email address. If there is more than one participant with the same first name, I add a number after their first name (e.g., Giovanni2).

After admitting all the participants to the Zoom meeting room, I do roll call, a way to reveal participants' gender to other participants without making gender salient (Bordalo et al. 2019; Coffman, Flikkema, and Shurchkov 2021). Specifically, I take attendance by calling each participant's first name one by one and ask her or him to respond via microphone. This process ensures other participants that the called participant's first name corresponds to their gender. If there are more participants than I would need for the session (I need 16 participants), I draw random numbers from 1 to 40 and ask those who drew the coins with the same number to leave.[7] Those who leave the session receive the 2€ show-up fee. Figure 2 shows a Zoom screen participants would see during the roll call (the person whose camera is on is the experimenter; participants would see this screen throughout the experiment, but the experimenter's camera may be turned off).

I then read out the instructions about the rules of the experiment and take questions on Zoom. Once participants start the main part, they can communicate with the experimenter only via Zoom's private chat.

Part 1: Individual practice stage

Participants work on the puzzle individually with an incentive (0.2€ for each puzzle they solve). They can solve as many puzzles as possible with increasing difficulty (maximum 15 puzzles) in 4 minutes. This part familiarizes them with the puzzle and provides us with a measure of their ability given by the number of puzzles they solve. After the 4 minutes are over, they receive information on how many puzzles they have solved.

Part 2: Collaborator selection stage

Part 2 contains seven rounds, and participants learn the rules of part 3 before starting part 2. This part is based on Fisman et al. (2006, 2008)'s speed dating experiments and proceeds as follows: first, participants are allocated to a group of 8 based on their ability similarity as

6. Zoom link is sent with an invitation email; I check that they have indeed registered for a given session before admitting them to the Zoom meeting room.

7. I draw with replacement a number from 1 to 40 using Google's random number generator (https://www.google.com/search?q=random+number). If no participant has a coin with the drawn number, I draw next number until the number of participants is 16. I share my computer screen so that participants see the numbers are actually drawn randomly.

measured in part 1. This is done to reduce ability difference among participants, and participants do not know this grouping criterion.

Second, participants are paired with another randomly chosen participant in the same group and solve one puzzle together by alternating their moves. The participant who makes the first move is drawn at random and both participants know this first-mover selection criterion. If they cannot solve the puzzle within 2 minutes, they finish the puzzle without solving it. Participants are allowed to reverse the paired participant's move.[8] Reversing the partner's move is what I call correction in this paper. Each participant's contribution in a given puzzle is measured by net good moves. Figure 1 shows a sample puzzle screen where a participant is paired with another participant called Valeria and making their move.[9] The paired participant's first name is displayed on the computer screen throughout the puzzle and when participants select their collaborator to subtly inform the paired participant's gender.

Once they finish the puzzle, participants state whether they would like to collaborate with the same participant in part 3 (yes/no). At the end of the first round, new pairs are formed, with a perfect stranger matching procedure, so that every participant is paired with each of the other seven members of their group once and only once. In each round, participants solve another puzzle in a pair, then state whether they would like to collaborate with the same participant in part 3. The sequence of puzzles is the same for all pairs in all sessions. The puzzle difficulty is kept the same across the seven rounds. The minimum number of moves to solve the puzzles is set to 8 based on the pilot.

At the end of part 3, participants are paired according to the following algorithm:

1. For every participant, call it i, I count the number of matches; that is, the number of other participants in the group who were willing to be paired with i and with whom i is willing to collaborate in part 3.
2. I randomly choose one participant.
3. If the chosen participant has only one match, I pair them and let them work together in part 3.
4. If the chosen participant has more than one match, I randomly choose one of the matches.
5. I exclude two participants that have been paired and repeat (1)-(3) until no feasible match is left.
6. If some participants are still left unpaired, I pair them up randomly.

Part 3: Group work stage

The paired participants work together on the puzzles by alternating their moves for 12 minutes and earn 1€ for each puzzle solved. Which participant makes the first move is randomized at each puzzle, and this is told to both participants as in part 2. They can solve as many puzzles as possible with increasing difficulty (maximum 20 puzzles).

8. Solving the puzzle itself is not incentivized, and thus participants who do not want to collaborate with the paired participant or fear to receive a bad response may not reverse that participant's move even if they think the move is wrong. However, since I am interested in the effect of correction on collaborator selection, participants' *intention* to correct that does not end up as an actual correction does not confound the analysis.

9. All the texts are in Italian in the experiment; see Online Appendix section E for the original screen.

Post-experiment

Each participant answers a short questionnaire which consists of (i) the six hostile and benevolent sexism questions used in Stoddard, Karpowitz, and Preece (2020) with US college students and (ii) their basic demographic information and what they have thought about the experiment.[10] The answer to their demographic information is used to know participants' characteristics as well as casually check whether they have anticipated that the experiment is about gender, for which I do not find any evidence.

After participants answer all the questions, I tell them their earnings and let them leave the virtual room and Zoom. They receive their earnings via PayPal.

2.2 Implementation

The experiment was programmed with oTree (Chen, Schonger, and Wickens 2016) and conducted in Italian during November-December 2020. I recruited 464 participants (244 female and 220 male) registered on the Bologna Laboratory for Experiments in Social Science's ORSEE (Greiner 2015) who (i) were students, (ii) were born in Italy, and (iii) had not participated in gender-related experiments before (as far as I could check).[11] The first two conditions were to reduce noise coming from differences in socio-demographic backgrounds and race or/and ethnicity that may be inferred from participants' first name or/and voice, and the last condition was to reduce experimenter demand effects. The number of participants was determined by a power simulation in the pre-analysis plan to achieve 80% power.[12] The experiment is pre-registered with the OSF.[13]

I ran 29 sessions with 16 participants each. The average duration of a session was 70 minutes. The average total payment per participant was 11.55€ with the maximum 25€ and the minimum 2€, all including the 2€ show-up fee. Table 1 describes participants' characteristics. The table shows that female participants are more likely to major in humanities and male participants are more likely to major in natural sciences and engineering, a tendency observed in most OECD countries (see, for example, Carrell, Page, and West 2010).[14] Also, most female and male participants are either bachelor or master students (97% of female and 94% of male) and the rest are PhD students.

3 Data description

I use part 2 data in the analysis as part 2 is where we can observe collaborator selection decisions. I aggregate the move-level data at each puzzle so that we can associate behaviors in the puzzle to

10. I was planning to construct a gender bias measure from the hostile and benevolent sexism questions and use it to show those with higher gender bias respond more negatively to women's corrections. However, people do not respond more negatively to women's corrections and that I could not have enough variation in this gender bias measure, so decided not to report it. See the pre-analysis plan in the online Appendix section C.

11. The laboratory prohibits deception, so no participant has participated in an experiment with deception.

12. This number includes 16 participants from a pilot session run before the pre-registration where the experimental instructions were slightly different. The results are robust to exclusion of these 16 participants.

13. The pre-registration documents are available at the OSF registry: https://osf.io/tgyc5. The pre-analysis plan is also in the online Appendix C.

14. Individual fixed effects in the analysis control for one's major. However, I do not run heterogeneity analysis by major because major choice is endogenous to one's gender.

the collaborator selection decisions. As shown in Figure 3, both mixed gender and single-gender groups perform equally well (panel A), about 71% of the puzzles are solved within a minimum number of moves (panel B, the minimum number of moves is 8), and corrections happen across the moves.

Table 2 describes own (panel A) and partner's puzzle behaviors (panel B) and puzzle outcomes (panel C). Panel A shows that there are no gender differences in puzzle-solving ability: both contribution in part 2 and the number of puzzles solved in part 1, the difference between female and male participants are statistically insignificant at 5% and quantitatively insignificant.[15][16] This is consistent with Isaksson (2018), who also finds no gender difference in contribution or number of puzzles solved alone using the same puzzle, suggesting that any gender difference I would find is unlikely to come from their ability difference. Panel A also shows that there are no gender differences in propensity to correct partners, suggesting any gender differences I would find are not coming from either gender corrects more than the other gender.

Panel A of Figure 4 presents the distribution of contribution by participants gender to further elaborate panel A of Table 2 that women and men are equally good at puzzle solving: in about 70% of the puzzles, participants' contribution is 4 (total good moves minus total bad moves), and women's and men's distributions almost overlap.

Panel B shows that puzzle-solving ability as well as propensity to make corrections (both of a mistake and of a right move) of partners paired with female and male participants is the same, suggesting random pairing was successful and that any gender differences I would find are not coming from partners of either gender correct more often. Participants are corrected by their partner in 15-16% of the total puzzles, of which 12-13% are good corrections, and 5-6% are bad corrections.[17]

Panel C shows that participants state they want to collaborate with the partner 71-72% of the time. Participants spend on average 43-44 seconds for each puzzle (the maximum time a pair can spend is 120 seconds), and take 11 moves. 85-86% of the puzzles are solved and participants and the partner correct each other's move consecutively in 4% of the puzzles.[18] There is no gender difference in any of these outcomes, suggesting any gender differences cannot be attributed to the imbalance in these outcomes.[19]

4 Theoretical framework

I present a simple theoretical framework to provide a benchmark for rational agent's behaviors.

I consider a fully rational participant i who maximizes their expected payoff by selecting their collaborator j from a set of i's potential collaborators $J \equiv \{1, 2, 3, 4, 5, 6, 7\}$. The payoff is

15. The number of puzzles solved in part 1 is marginally significant but quantitatively insignificant.

16. The correlation coefficient between contribution and number of puzzles solved in part 1 is 0.1059 and the p-value is below 0.001 (with standard errors clustered at individual level).

17. The percentage of good corrections and bad corrections do not sum up to the percentage of any correction means there are puzzles where both good and bad corrections occurred. The results are robust to exclusion of these overlapping puzzles, as shown in Figure C4.

18. Indeed, in puzzles where consecutive correction happens, probability of selecting a paired participant as collaborator drops from 78.0% to 26.8%.

19. Note that time spent to solve a puzzle is endogenous to correction and not a good control. For example. if one corrects a mistake, then it takes fewer time to solve the puzzle. If one corrects a right move, on the other hand, then it takes more time to solve the puzzle.

increasing in i's belief about j's ability. I assume i can partially observe j's move quality, so i's belief about j's ability is increasing in j's ability perceived by i.

Thus, if i would select with whom to play in part 3, i would face the following problem:

$$\max_{j \in J} E_{\mu_j}[\pi(\mu_j(\tilde{a}_j, c_j, f_j))|\theta_i, \omega_i], \quad \partial\pi/\partial\mu_j > 0,\ \partial\mu_j/\partial\tilde{a}_j > 0 \tag{1}$$

where each term is defined as follows:

- μ_j: i's belief about j's ability
- $\tilde{a}_j$: j's ability perceived by i
- c_j: j's correction (=1 if j corrected i, =0 if j did not correct i)
- f_j: j's gender (=1 if female, =0 if male)
- θ_i: i's belief about their ability relative to other participants (>0 if higher, =0 if same, <0 if lower)
- ω_i: j's belief about women's ability relative to men (>0 if higher, =0 if same, <0 if lower)

If i can fully observe j's move quality and i is fully rational, then j's correction, c_j, and gender, f_j, do not convey any information about j's ability and is irrelevant for i's decision making. However, since i can only partially observe j's move quality, j's correction and gender convey information about j's ability even if i is fully rational.[20]

First, keeping j's ability perceived by i fixed, the information j's correction conveys depends on θ_i. If i believes they are good at the puzzle, they would consider a correction as a signal of low ability because i believes their move is correct. On the other hand, if i believes their ability is low, then they would consider a correction as a signal of high ability. If i believes their ability is the same as j's, then a correction would not convey any information.

However, since i can partially observe j's move quality, i considers a good correction as less negative/more positive signal than a bad corrections regardless of θ_i. Thus, we have the following proposition:

Proposition 1. *A fully rational participant responds more negatively to a bad correction than to a good correction regardless of their belief about their ability relative to other participants. That is:*

$$\partial\mu_j/\partial c_j|_{c_j \text{ is a bad correction}} < \partial\mu_j/\partial c_j|_{c_j \text{ is a good correction}}\ \forall\theta_i \tag{2}$$

Also, the more the i understands the puzzle, the more they can observe j's move quality, hence corrections, regardless of θ_i. Thus, we have the following proposition:

Proposition 2. *A fully rational participant with higher puzzle solving ability respond less negatively to good corrections and more negatively to bad corrections than another fully rational participant with lower puzzle solving ability. This is true regardless of their belief about their ability relative to other participants. That is:*

$$\begin{aligned} \partial\mu_j/\partial c_j|_{\text{i's ability is high \& } c_j \text{ is a good correction}} &> \partial\mu_j/\partial c_j|_{\text{i's ability is low \& } c_j \text{ is a good correction}}\ \forall\theta_i \\ \partial\mu_j/\partial c_j|_{\text{i's ability is high \& } c_j \text{ is a bad correction}} &< \partial\mu_j/\partial c_j|_{\text{i's ability is low \& } c_j \text{ is a bad correction}}\ \forall\theta_i \end{aligned} \tag{3}$$

20. I nonparametrically control for j's gender, but I also examine the effect of interaction term between j's correction and j's gender.

Similar to response to corrections, if i believes women is better at the puzzle, they would consider a correction from a woman as a signal of high ability relative to men's correction. On the other hand, if i believes women is worse at the puzzle, then they would consider a correction from a woman as a signal of low ability relative to men's correction. If i believes women and men are equally good at the puzzle, then a correction from a woman or man is irrelevant. Thus, we have the following proposition:

Proposition 3. *A fully rational participant's response to women's corrections relative to men's correction depends on their belief about women's ability relative to men's. This is true regardless of their belief about their ability relative to other participants. That is:*

$$\begin{aligned} \partial^2 \mu_j / \partial c_j \partial f_j > 0 \; \forall \theta_i \; \textit{if} \; \omega_i > 0 \\ \partial^2 \mu_j / \partial c_j \partial f_j < 0 \; \forall \theta_i \; \textit{if} \; \omega_i < 0 \end{aligned} \tag{4}$$

In particular, if they believe women and men have the same ability, then gender of the person who corrects them does not matter. That is:

$$\partial^2 \mu_j / \partial c_j \partial f_j = 0 \; \forall \theta_i \; \textit{if} \; \omega_i = 0 \tag{5}$$

I consider deviations from these propositions are evidence of non-rationality.

5 Response to corrections

In this section, I document evidence that participants – both women and men – understand the notion of good and bad moves. Thus, they are more willing to collaborate with someone who contributed more to the puzzle. This is in line with what one would expect and validates my experimental design.

However, after controlling for that person's contribution, participants are less willing to work with a person who corrected their move, even if that person makes good corrections, which is inefficient and seems to indicate deviation from the rational agent's benchmark in Proposition 1. However, it is unclear whether it is evidence of deviation from rationality; I will return to this point later in section 7.

5.1 Response to corrections: Estimating equation

I estimate the following model with OLS.

$$Select_{ij} = \beta_1 CorrectedGood_{ij} + \beta_2 CorrectedBad_{ij} + \beta_3 Female_j + \delta Contribution_j + \mu_i + \epsilon_{ij} \tag{6}$$

where each variable is defined as follows:

- $Select_{ij} \in \{0, 1\}$: an indicator variable equals 1 if i selects j as their collaborator, 0 otherwise.
- $CorrectedGood_{ij} \in \{0, 1\}$: an indicator variable equals 1 if j corrected i and moved the puzzle closer to the solution, 0 otherwise.

- $CorrectedBad_{ij} \in \{0, 1\}$: an indicator variable equals 1 if j corrected i and moved the puzzle far away from the solution, 0 otherwise.
- $Female_j \in \{0, 1\}$: an indicator variable equals 1 if j is female, 0 otherwise.
- $Contribution_j \in \mathbb{Z}$: j's contribution to a puzzle played with i.
- ϵ_{ij}: omitted factors that affect i's likelihood to select j as their collaborator.

and $\mu_i \equiv \sum_{k=1}^{N} \mu^k \mathbb{1}[i = k]$ is individual fixed effects, where N is the total number of participants in the sample and $\mathbb{1}$ is the indicator variable. Standard errors are clustered at the individual level.[21]

The key identification assumption is that $Contribution_j$ fully captures j's ability perceived by i through j's puzzle moves (not true ability).[22] This assumption is reasonable if we think participants' willingness to collaborate is increasing in the partner's contribution to the puzzle, which is consistent with that participants can partially observe their partners' ability and their expected utility is increasing in their payoff.

5.2 Response to corrections: Results

Table 3 presents the regression results of equation 6. Columns 1-3 include all participants' willingness to collaborate but column 1 excludes partner's contribution and individual fixed effects and column 2 partner's contribution. Columns 4-6 present the corresponding results for women and columns 7-9 for men.

Column 1 shows that when we do not control for between-participants variation, the coefficient estimate on good correction is underestimated. This is true for women (column 4) and men (column 7).

Column 2 shows that when we do not control for partner's contribution, the coefficient estimate on bad correction is negative and very large: the point estimate is -0.508 (p-value < 0.01); that is, participants are 50.8% less willing to collaborate with partners who made a bad correction, a correction that moved the puzzle far away from the solution. Indeed, these coefficient estimates are more negative than the coefficient estimates on good corrections: 0.271 more negative (p-value < 0.01). This is true when we separately examine women (column 5, 0.281 with p-value < 0.01) and men (column 8, 0.281 with p-value < 0.01).

Corroborating this, looking at column 3, the coefficient estimate on the partner's contribution is positive and quantitatively and statistically highly significant and is 0.084 (p-value < 0.01). This suggests that participants are 8.4% more willing to collaborate with partners who make one more good move. This is true for women (column 6, 0.089 with p-value < 0.01) and men (column 9, 0.080 with p-value < 0.01). This is evidence that my experimental design is valid: participants correctly understand the notion of good and bad moves and that participants are

21. This is because the treatment unit is i. Although the same participant appears twice (once as i and once as j), j is passive in collaborator selection.

22. By random pairing of participants, the paired participant's gender is exogenous to participant's unobservables. However, correction is not exogenous for two reasons: (i) correction can be correlated with the paired participant's ability and paired participant's ability can affect participant's collaborator selection; (ii) There is an effect similar to the reflection effect: participant's puzzle behavior affects the paired participant's behavior and vice versa; for example, a participant's meanness can increase the paired participant's correction and can also affect their collaborator selection. The identification assumption concerns the former point. To address the latter point, I add individual fixed effects.

more willing to collaborate with partners who contributed more.

The coefficient estimate on good correction in column 3 is negative and quantitatively and statistically highly significant and is -0.204 (p-value < 0.01). This suggests that people are 20.4% less willing to collaborate with those who made a good correction(s), which corresponds to an increase in the contribution by 0.87 standard deviation.[23] The corresponding coefficient estimate for women is -0.229 (column 6, p-value < 0.01) and -0.168 for men (column 9, p-value < 0.01).

The coefficient estimate on bad correction in columns 3 is also negative and quantitatively and statistically significant and is -0.100 (p-value < 0.01). However, the magnitude is smaller than the coefficient estimate on good correction: the difference is -0.104 (p-value < 0.05). This is mainly driven by men: the corresponding coefficient estimate for women is -0.172 (column 6, p-value < 0.01) but is -0.011 (p-value > 0.10) for men. These are inefficient. They also seem to indicate deviation from the rational agent's benchmark in Proposition 1, but we cannot say anything definitive because response to corrections depends on the belief about people's own ability relative to partners, and if they believe they are better at the puzzle than others, as people are in general overconfident, both women and men albeit that men are more overconfident (Croson and Gneezy 2009). I will come back to this point later in section 7.

6 Do women's corrections receive stronger negative reactions?

In this section, I document that people – either men or women – do not underestimate women's contribution, which suggests that their prior about women's ability to solve the puzzle is neither higher nor lower than men. I also documents that women's corrections do not receive stronger negative reactions, either by women or men. These are consistent with Proposition 3.

6.1 Do women's corrections receive stronger negative reactions? Estimating equation

I estimate the following model with OLS.

$$\begin{aligned} Select_{ij} =& \beta_1 CorrectedGood_{ij} + \beta_2 CorrectedBad_{ij} + \beta_3 Female_j \\ &+ \beta_4 CorrectedGood_{ij} \times Female_j + \beta_5 CorrectedBad_{ij} \times Female_j \\ &+ \delta_1 Contribution_j + \delta_2 Contribution_j \times Female_j + \mu_i + \epsilon_{ij} \end{aligned} \tag{7}$$

Where each variable is defined as in equation 6.

6.2 Do women's corrections receive stronger negative reactions? Results

Table 4 presents the regression results of equation 7. As Table 3, columns 1-3 include all participants' willingness to collaborate but column 1 excludes partner's contribution and individual fixed effects and column 2 partner's contribution. Columns 4-6 present the corresponding results for women and columns 7-9 for men. Excluding partner's contribution and individual fixed effects results in similar bias as Table 3: between-participants variation underestimates the

23. The standard deviation is taken from panel B of Table 2 and is an arithmetic average of 2.73 for partners faced by women 2.87 for and partners faced by men: (2.73+2.87)/2=2.80.

coefficient estimate of good correction and that participants correctly understand the notion of good and bad moves and that participants are more willing to collaborate with partners who contributed more.

Looking at column 3, the coefficient estimate on the interaction between partner's contribution and female partner is almost 0 and statistically insignificant even at 10%. Column 6 shows this is true for women and column 9 for men. These suggest that people – both women and men – do not underestimate women's contribution when selecting a collaborator. In other words, people believe women and men's puzzle solving ability is the same.

Again in column 3, the coefficient estimates on the interaction between good correction and female partner is negative although statistically insignificant. The coefficient estimates on the interaction between female partner and bad correction is positive and statistically significant at 5%.

The negative coefficient estimate on the interaction between good correction and female partner mainly comes from men: looking at column 9, the corresponding coefficient estimate for men is -0.119 although only marginally significant at 10%, while for women it is 0.035 although statistically insignificant (column 6).

The positive coefficient estimate on the interaction between female partner and bad correction comes from both women and men: the corresponding coefficient estimate is 0.090 for women (column 6) and 0.168 for men (column 9), although neither of them is statistically significant.

While men's negative reaction to women's good correction is bit worrying as men are still majority in top positions both in industry and academia. However, it is only marginally significant and as discussed later in section 8 it is not robust. Thus, women's corrections do not receive stronger negative reactions from either women or men. Together with the evidence that people believe women and men are equally good at solving the puzzle, this is consistent with Proposition 3 and gender of the person who makes correction does not matter.

7 Who respond negatively to corrections?

If only low ability people respond negatively to corrections, then correcting colleagues may not be very costly. Also, it was inconclusive whether people's negative reaction to corrections documented in section 5 is irrational. However, if it is rational, people with higher puzzle solving ability – as measured in the part 1 individual practice stage – should respond less negatively to good corrections and more negatively to bad corrections because they are better able to distinguish good and bad corrections as in 2.

In this section, I document that even high ability people respond negatively to good corrections with men responding more negatively, which indicates that (i) correcting colleagues is really costly both for individuals and groups and that (ii) negative response to corrections is irrational.

7.1 Who respond negatively to corrections? Estimating equation

I estimate the following model with OLS.

$$\begin{aligned} Select_{ij} =& \beta_1 CorrectedGood_{ij} + \beta_2 CorrectedBad_{ij} + \beta_3 Female_j \\ & + \beta_4 CorrectedGood_{ij} \times HighAbility_i + \beta_5 CorrectedBad_{ij} \times HighAbility_i \\ & + \delta_1 Contribution_j + \delta_2 Contribution_j \times HighAbility_i + \mu_i + \epsilon_{ij} \end{aligned} \tag{8}$$

where each variable is defined as follows:

- $HighAbility_i \in \{0, 1\}$: an indicator variable equals 1 if i solved the above-median number of puzzles in part 1 in a session they have participated, 0 otherwise.

Other variables are as defined in equations 8.

7.2 Who respond negatively to corrections? Results

Table 5 presents the regression results of equation 7. As Table 3, columns 1-3 include all participants' willingness to collaborate but column 1 excludes partner's contribution and individual fixed effects and column 2 partner's contribution. Columns 4-6 present the corresponding results for women and columns 7-9 for men. High ability dummy is added in columns 1, 4, and 7 because we need a different intercept for high and low ability.

Excluding partner's contribution and individual fixed effects results in similar bias as Table 3: between-participants variation underestimates the coefficient estimate of good correction and that participants correctly understand the notion of good and bad moves and that participants are more willing to collaborate with partners who contributed more.

The coefficient estimate on high ability on columns 1, 4, 7 are close to 0 and not statistically significant, which means high ability people are not more selective than low ability people. Some people may think high ability people must be more selective because they do not want to work with low ability people, but remember I matched up high ability people with high ability people – and low ability people with low ability people – in part 2 collaborator selection stage. Panel B of Figure 4 shows distribution of contribution of high ability people is indeed less dispersed.

In column 3, the coefficient estimate on the interaction between good correction and high ability is negative and statistically significant at 5%. This effect comes from both women and men, with the effect on men being stronger: the corresponding coefficient estimate for men (column 9) is more negative and statistically significant at 5%, and it is positive but not statistically significant for women (in column 6).

The coefficient estimate on the interaction between bad correction and high ability is almost zero. The corresponding coefficient estimate is positive for women (column 6) and negative for men (column 9), although they are both statistically insignificant.

Thus, even high ability people respond negatively to good corrections with men responding more negatively. Thus, correcting colleagues is really costly both for individuals and groups because even high ability people are less willing to collaborate with someone who corrects their mistakes.

Also, as discussed at the beginning of this section, high ability people should be able to distinguish good and bad corrections and should respond less negatively to good corrections and more negatively to bad corrections than low ability people, what I presented in Proposition 2 as a rational agent benchmark. However, what we see here is the opposite, at least inconsistent. So the negative reaction to corrections is irrational.

8 Robustness checks

In this section, I document evidence that my findings are robust to three concerns.

Concern 1: Unsolved puzzles Whether participants can solve a puzzle is an outcome of a particular pairing that is random. However, "a good move is only preferable if you are playing with a partner who is also trying to solve the puzzle" (Isaksson 2018, p. 25). If a participant is not trying to solve the puzzle, then a pair is unlikely to solve the puzzle and good and bad corrections may not be meaningful.

Concern 2: Rounds 6 and 7 Remember that each participant plays the puzzle for seven rounds and variables unaffected by treatment (interactions within a randomly-formed pair) must be balanced. Figure 5 plots average partner gender balance (fraction of female partners, panel A) and puzzle outcomes (panels B-H) across seven rounds along with their 95% confidence intervals, separately for female (blue) and male participants (green). We see that in rounds 6 and 7, participants are less willing to collaborate, experience more corrections, and less likely to solve the puzzle. Although they are all outcomes of a particular pair so can just be correlations, one may wonder whether rounds 6 and 7 are driving the results.

Concern 3: Puzzles where good and bad corrections occurred There are 495 puzzles in which at least one correction occurred, of which 325 puzzles experienced good corrections only, 110 puzzles bad corrections only, and 60 puzzles experienced both good and bad corrections. In these 60 puzzles, it is unclear which corrections – good or bad – dominated people's mind in determining whether to collaborate with a paired person.

To address these concerns, I re-estimate equations 6, 7, and 8 and plot the coefficient estimates and 95% confidence intervals of the main coefficients of interest in Figures 6, 7, and 8, respectively, with solved puzzles only (green dots and lines), with rounds 1-5 only (red dots and lines), and with puzzles where only good or bad corrections occurred (purple dots and lines). As a reference, I also plot the coefficient estimates and 95% confidence intervals with the main sample used in Tables 3, 4, and 5 (blue dots and lines). All estimates are from the full models (columns 3, 6, and 9 in all tables).

The main coefficients of interest for equation 6 are good and bad corrections. Looking at Figure 6, we see that most coefficient estimates are close to the main estimates. The estimates are more negative for good correction when the sample is limited to solved puzzles only, but they are more in line with the main findings.

The main coefficients of interest for equation 7 are the interaction between good correction and female partner and between bad correction and female partner. Looking at Figure 7, we again see most the coefficient estimates are close to the main estimates. Again, the estimates with solved puzzles only present somewhat different evidence; in particular, response to good corrections by female partners is negative although statistically insignificant for women and positive although very close to zero for men. But these are in line with the main findings.

The main coefficients of interest for equation 8 are the interaction between good correction and high ability and between bad correction and high ability. Looking at Figure 8, we again see most the coefficient estimates are close to the main estimates.

These robustness tests show that the findings are robust to these concerns.

9 Discussion and Conclusion

This paper demonstrates that people, including those with high productivity, are less willing to collaborate with a person who has corrected them even if the correction improves group performance. However, I do not find evidence that people respond more negatively to corrections by women. Thus, although people are not harsher against women's corrections, dislike to be corrected is costly for the individual and the group, as well as irrational.

While the laboratory setting is different from the real-world, my findings are likely to be a lower bound because of the following three reasons. First, being corrected is not observed by others in my experiment: those who have been corrected do not face any reputation cost, unlike in the real-world. Second, the emotional stake is much smaller: the puzzle ability is not informative of the ability relevant for their work or study; it is not something people have been devoting much of their time to, such as university exams, academic research, and corporate investment projects. Third, participants are equal in my experiment; in the real-world, on the other hand, there are sometimes senior-junior relationships, and corrections by junior people may induce stronger negative reactions.

But there are two caveats. The first is that participants are strangers to each other in my experiments while people know each other in the real-world. Thus, it is possible that repeated interactions would mitigate people's negative response to corrections (but they may also magnify the negative response due to rivalry, failure to build a good rapport, etc.). The second is that most participants are bachelor or master students who are supposed to have a weaker gender bias. Women's corrections may receive stronger negative reactions if participants are older.

References

Arechar, Antonio A., Simon Gächter, and Lucas Molleman. 2018. "Conducting interactive experiments online." *Experimental Economics* 21 (1): 99–131.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2019. "Beliefs about Gender." *American Economic Review* 109 (3): 739–773.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *The Quarterly Journal of Economics* 125 (3): 1101–1144.

Castagnetti, Alessandro, and Renke Schmacker. 2021. *Protecting the Ego: Motivated Information Selection and Updating.* Working Paper.

Chakraborty, Priyanka, and Danila Serra. 2021. *Gender and leadership in organizations: Promotions, demotions and angry workers.* Working Paper.

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree—An open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Experimental Finance* 9:88–97.

Coffman, Katherine, Clio Bryant Flikkema, and Olga Shurchkov. 2021. "Gender stereotypes in deliberation and team decisions." *Games and Economic Behavior* 129:329–349.

Coffman, Katherine Baldiga. 2014. "Evidence on Self-Stereotyping and the Contribution of Ideas." *The Quarterly Journal of Economics* 129 (4): 1625–1660.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47 (2): 448–474.

Dupas, Pascaline, Alicia Sasser Modestino, Muriel Niederle, Justin Wolfers, and Seminar Dynamics Collective. 2021. *Gender and the Dynamics of Economics Seminars.* Working Paper.

Egan, Mark, Gregor Matvos, and Amit Seru. 2021. "When Harry Fired Sally: The Double Standard in Punishing Misconduct." *Journal of Political Economy.*

Ewers, Mara, and Florian Zimmermann. 2015. "Image and Misreporting." *Journal of the European Economic Association* 13 (2): 363–380.

Fisman, Raymond, Sheena S. Iyengar, Emir Kamenica, and Itamar Simonson. 2006. "Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment." *The Quarterly Journal of Economics* 121 (2): 673–697.

———. 2008. "Racial Preferences in Dating." *The Review of Economic Studies* 75 (1): 117–132.

Gallus, Jana, and Emma Heikensten. 2019. *Shine a light on the bright: The effect of awards on confidence to speak up in gender-typed knowledge work.* Working Paper.

Goeschl, Timo, Marcel Oestreich, and Alice Soldà. 2021. *Competitive vs. Random Audit Mechanisms in Environmental Regulation: Emissions, Self-Reporting, and the Role of Peer Information.* Working Paper 0699. University of Heidelberg, Department of Economics.

Greiner, Ben. 2015. "Subject pool recruitment procedures: organizing experiments with ORSEE." *Journal of the Economic Science Association* 1 (1): 114–125.

Guo, Joyce, and María P. Recalde. 2020. *Overriding in teams: The role of beliefs, social image, and gender.* Working Paper.

Isaksson, Siri. 2018. *It Takes Two: Gender Differences in Group Work.* Working Paper.

Jones, Benjamin F. 2021. "The Rise of Research Teams: Benefits and Costs in Economics." *Journal of Economic Perspectives* 35 (2): 191–216.

Kennedy, Kathleen A., and Emily Pronin. 2008. "When Disagreement Gets Ugly: Perceptions of Bias and the Escalation of Conflict." *Personality and Social Psychology Bulletin* 34 (6): 833–848.

Kőszegi, Botond. 2006. "Ego Utility, Overconfidence, and Task Choice." *Journal of the European Economic Association* 4 (4): 673–707.

Kunda, Ziva. 1987. "Motivated inference: Self-serving generation and evaluation of causal theories." *Journal of Personality and Social Psychology* 53 (4): 636–647.

Ronayne, David, and Daniel Sgroi. 2019. *Ignoring Good Advice.* Working Paper.

Sarsons, Heather. 2019. *Interpreting Signals in the Labor Market: Evidence from Medical Referrals.* Working Paper.

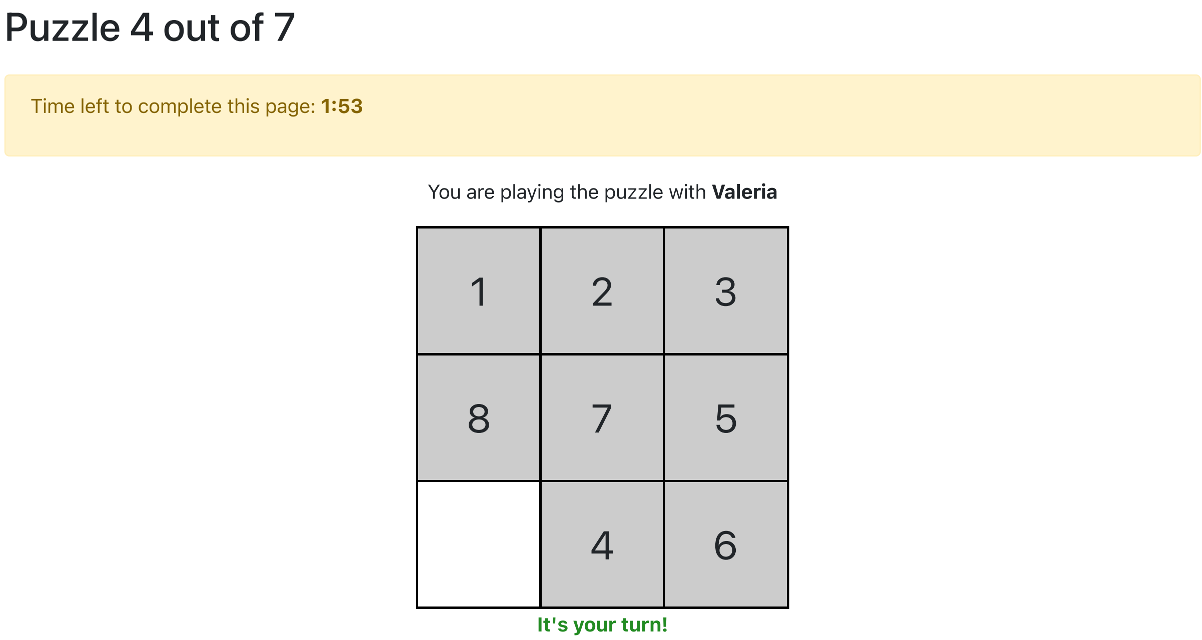
Sinclair, Lisa, and Ziva Kunda. 2000. "Motivated Stereotyping of Women: She's Fine if She Praised Me but Incompetent if She Criticized Me." *Personality and Social Psychology Bulletin* 26 (11): 1329–1342.

Stoddard, Olga, Christopher F. Karpowitz, and Jessica Preece. 2020. *Strength in Numbers: A Field Experiment in Gender, Influence, and Group Dynamics.* Working Paper.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. "The Increasing Dominance of Teams in Production of Knowledge." *Science* 316 (5827): 1036–1039.

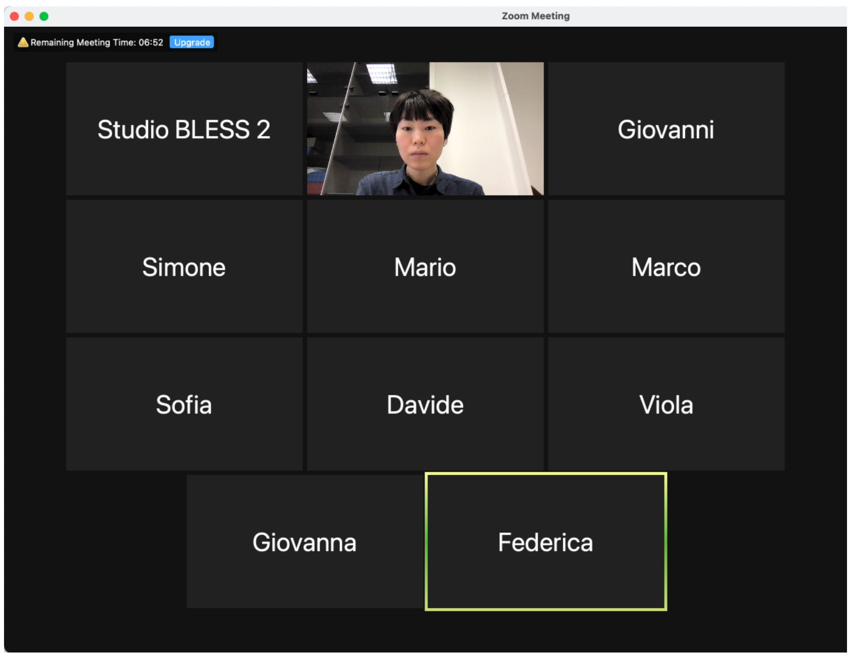
Figures

Figure 1: Puzzle screen



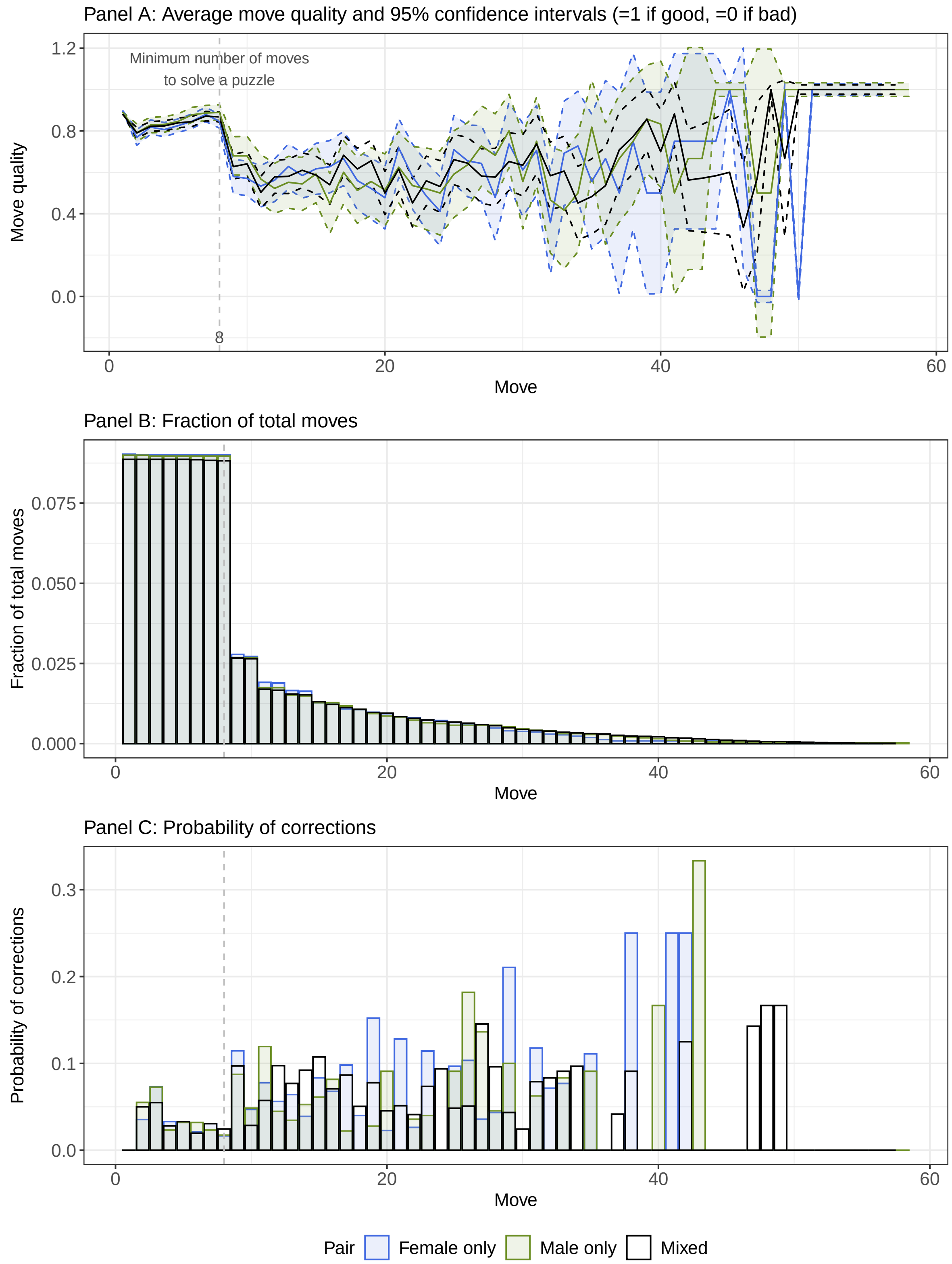
Notes: This shows a sample puzzle screen where a participant is matched with another participant called Valeria at the 4th round of the puzzle and making their move. All the texts are in Italian in the experiment; see Online Appendix section E for the original screen.

Figure 2: Zoom screen



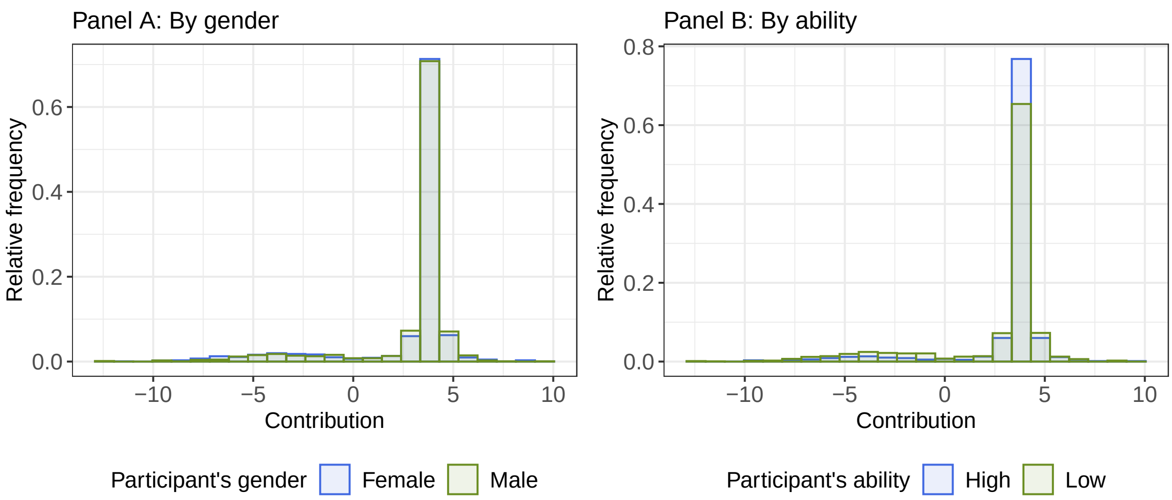
Notes: This figure shows a Zoom screen participants would see during the roll call. The experimenter's camera is on during the roll call. Participants would see this screen throughout the experiment but the experimenter's camera may be turned off.

Figure 3: Move quality, fraction of total moves, and probability of corrections



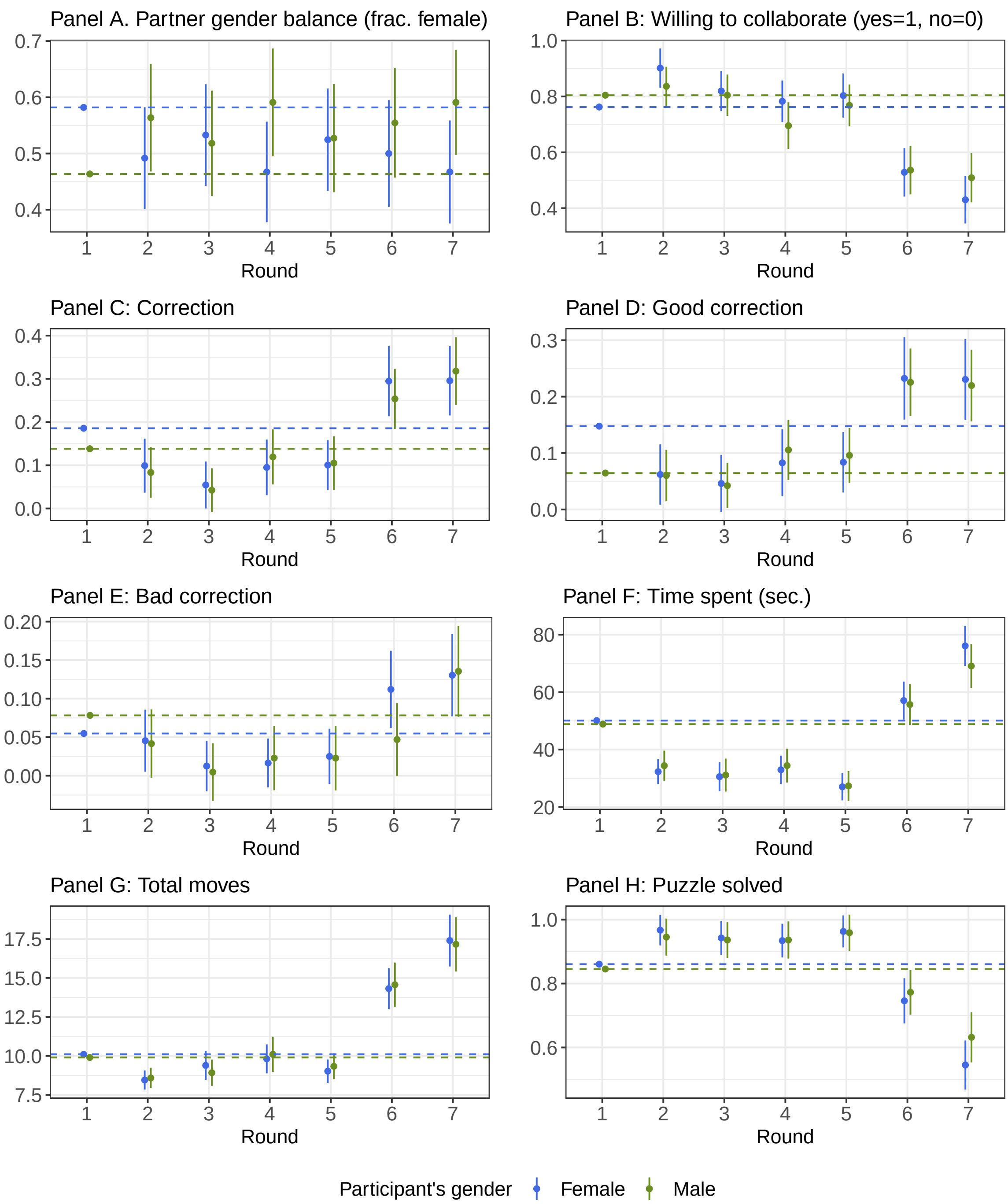
Notes: The average move quality along with 95% confidence intervals (panel A), the fraction of total moves in each move (panel B), and the probability of corrections in each move (panel C), separately for female only (gray), male only (white), and mixed gender pairs (blue). The confidence interval of panel A is 95% confidence intervals of βs from the following OLS regression: $MoveQuality_{ijt} = \beta_1 + \sum_{k=2}^{58} \beta_k \mathbb{1}[t_{ij} = k] + \epsilon_{ijt}$, where t_{ij} is the pair i-j's move round and $\mathbb{1}$ is an indicator variable. $MoveQuality_{ijt}$ takes a value of 1 if a move of a pair i-j in tth move is good and 0 if bad. I add an estimate of β_1 to estimates of β_2-β_{58} to make the figure easier to look at. Standard errors are clustered at the pair level.

Figure 4: Distribution of contribution



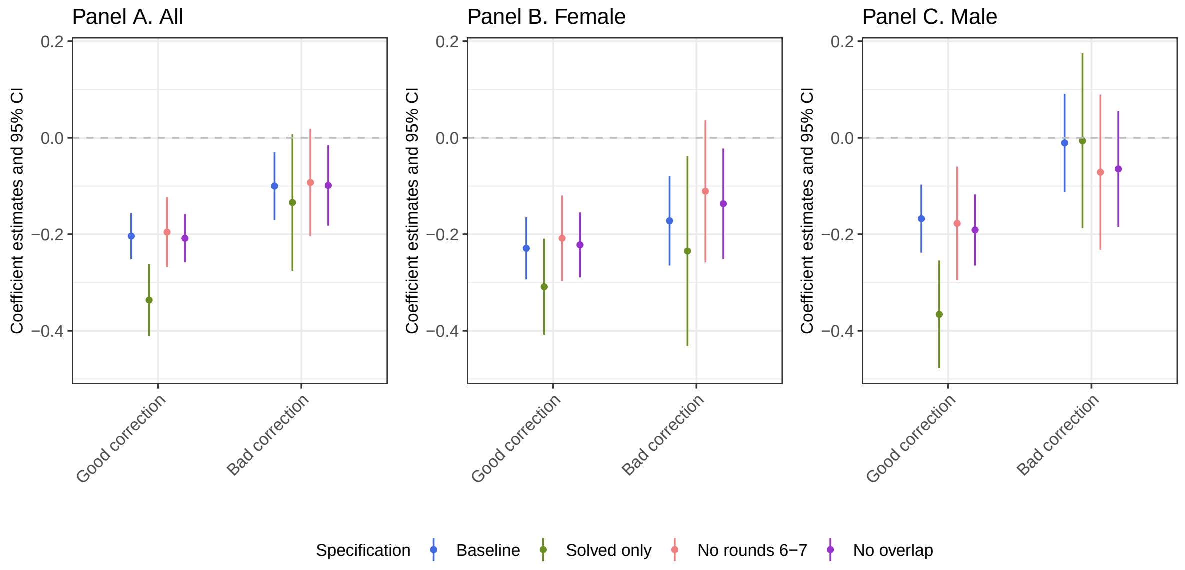
Notes: This figure shows the distribution of individual contribution by gender (panel A) and ability (panel B) and shows that most participants contributed to the same degree. Panel A further shows no gender difference in contribution, and panel B further shows that among high-ability people, higher fraction contributes to the puzzles to the same degree. Contribution is defined as one's net good moves in a given puzzle (the number of good moves minus the number of bad moves).

Figure 5: Balance across rounds



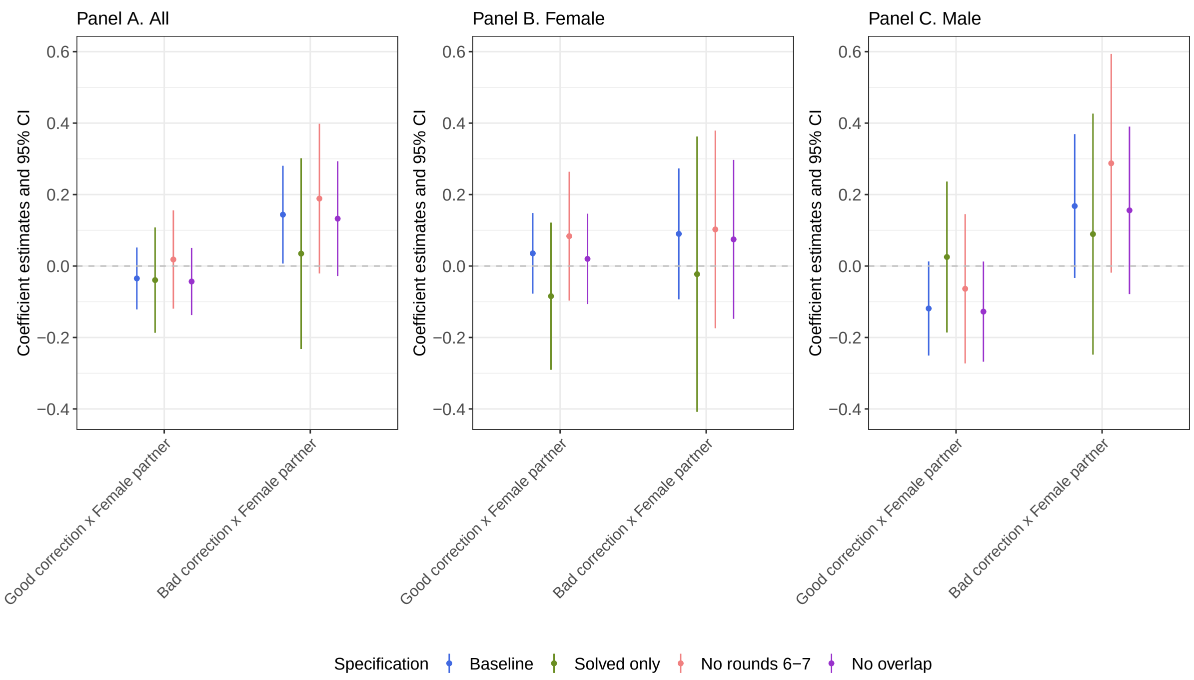
Notes: This figure shows point estimates and 95% confidence intervals of βs from the following OLS regression with gender balance (female dummy) and different puzzle outcomes separately for female (blue) and male participants (green): $y_{ij} = \beta_1 + \sum_{k=2}^{7} \beta_k \mathbb{1}[t_{ij} = k] + \epsilon_{ij}$, where $t_{ij} \in \{1, 2, 3, 4, 5, 6, 7\}$ is the puzzle round in which i and j are playing, $\mathbb{1}$ is an indicator variable, and y_{ij} is outcome variable indicated in each panel. I add the estimate of β_1 to estimates of β_2-β_7 to make the figure easier to look at. Standard errors are clustered at the individual level.

Figure 6: Response to corrections: Robustness



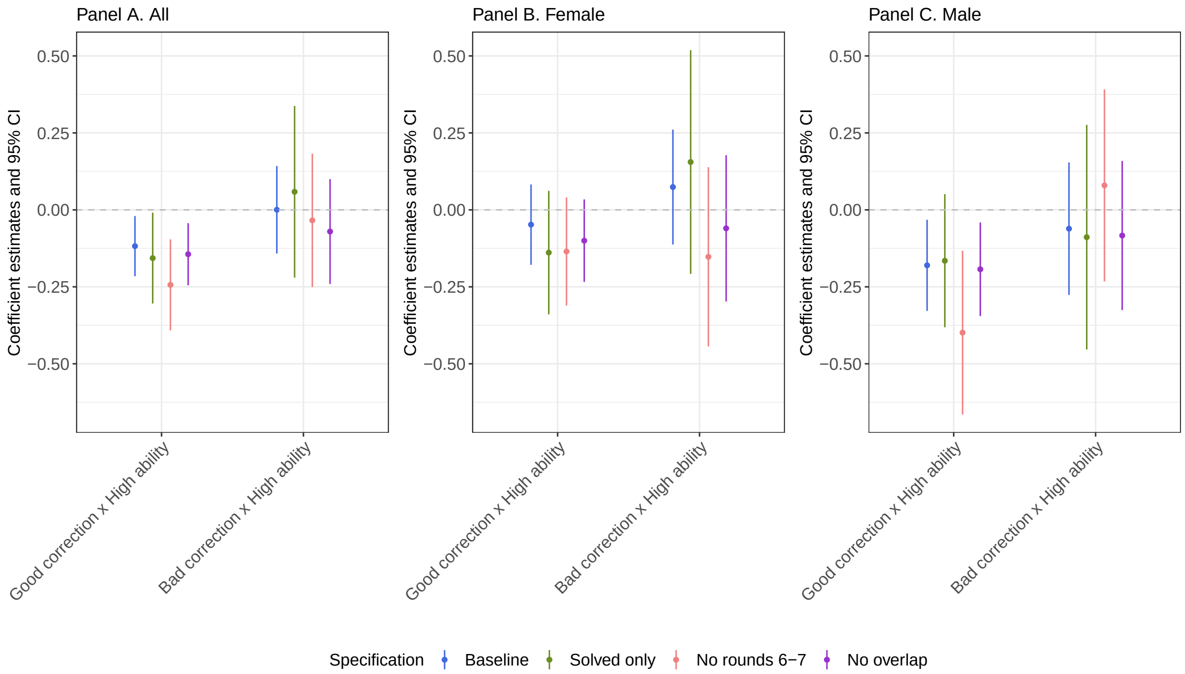
Notes: This figure plots the coefficient estimates and 95% confidence intervals of columns 3, 6, and 9 of Table 3 with solved puzzles only (the green dots and lines), with rounds 1-5 only (the red dots and lines), and with puzzles where only good or bad corrections occurred (the purple dots and lines). The blue dots and lines are the corresponding baseline estimates. They show that the findings in Table 3 are robust to limiting samples in these ways.

Figure 7: Response to corrections made by women vs. men: Robustness



Notes: This figure plots the coefficient estimates and 95% confidence intervals of columns 3, 6, and 9 of Table 4 with solved puzzles only (the green dots and lines), with rounds 1-5 only (the red dots and lines), and with puzzles where only good or bad corrections occurred (the purple dots and lines). The blue dots and lines are the corresponding baseline estimates. They show that the findings in Table 4 are robust to limiting samples in these ways.

Figure 8: Response to corrections made of high vs. low ability people: Robustness



Notes: This figure plots the coefficient estimates and 95% confidence intervals of columns 3, 6, and 9 of Table 5 with solved puzzles only (the green dots and lines), with rounds 1-5 only (the red dots and lines), and with puzzles where only good or bad corrections occurred (the purple dots and lines). The blue dots and lines are the corresponding baseline estimates. They show that the findings in Table 5 are robust to limiting samples in these ways.

Tables

Table 1: Participants' characteristics

	Female (N=244)			Male (N=220)			Difference (Female − Male)	
	Mean	SD	Median	Mean	SD	Median	Mean	P-value
Age	24.45	3.13	24	25.87	4.33	25	-1.41	0.00
Region of origin:								
North	0.32			0.36			-0.04	0.37
Center	0.23			0.24			-0.01	0.77
South	0.45			0.40			0.06	0.23
Abroad	0.00			0.00			0.00	0.32
Major:								
Humanities	0.45			0.22			0.23	0.00
Social sciences	0.24			0.27			-0.03	0.52
Natural sciences	0.12			0.20			-0.08	0.02
Engineering	0.05			0.23			-0.17	0.00
Medicine	0.13			0.08			0.05	0.08
Program:								
Bachelor	0.34			0.26			0.08	0.06
Master	0.63			0.68			-0.05	0.26
Doctor	0.03			0.06			-0.03	0.11

Notes: This table describes participants' characteristics. P-values of the difference between female and male participants are calculated with heteroskedasticity-robust standard errors.

Table 2: Own and partners' puzzle behaviors and puzzle outcomes

	Female (N=1708)		Male (N=1540)		Difference (Female – Male)		
	Mean	SD	Mean	SD	Mean	SE	P-value
Panel A: Own behaviors							
Contribution	2.98	2.93	3.14	2.64	-0.16	0.10	0.11
# puzzles solved in pt. 1	8.36	2.41	8.80	2.34	-0.44	0.22	0.05
Correction	0.15	0.36	0.16	0.36	0.00	0.01	0.85
Good correction	0.12	0.33	0.12	0.33	0.00	0.01	0.90
Bad correction	0.06	0.23	0.05	0.22	0.00	0.01	0.70
Panel B: Partner's behaviors							
Contribution	3.04	2.73	3.07	2.87	-0.03	0.10	0.77
# puzzles solved in pt. 1	8.58	2.35	8.57	2.43	0.01	0.16	0.93
Correction	0.16	0.37	0.15	0.36	0.01	0.01	0.51
Good correction	0.13	0.33	0.12	0.32	0.01	0.01	0.44
Bad correction	0.06	0.23	0.05	0.22	0.01	0.01	0.44
Panel C: Puzzle outcomes							
Willing to collaborate (yes=1, no=0)	0.72	0.45	0.71	0.45	0.01	0.02	0.49
Time spent (sec.)	43.74	36.15	42.99	35.76	0.74	1.28	0.56
Total moves	11.18	7.46	11.21	7.70	-0.03	0.28	0.92
Puzzle solved	0.85	0.36	0.86	0.35	-0.01	0.01	0.43
Consecutive correction	0.04	0.20	0.04	0.21	0.00	0.01	0.81

Notes: This table describes own (panel A) and partner's puzzle behaviors (panel B) and puzzle outcomes (panel C). P-values of the difference between female and male participants are calculated with standard errors clustered at the individual level. Contribution is defined as one's net good moves in a given puzzle (the number of good moves minus the number of bad moves).

Table 3: Response to corrections

Outcome:	Willing to collaborate (yes=1, no=0)								
Sample:		All			Female			Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Good correction	-0.208***	-0.238***	-0.204***	-0.240***	-0.269***	-0.229***	-0.169***	-0.197***	-0.168***
	(0.028)	(0.030)	(0.024)	(0.037)	(0.043)	(0.033)	(0.041)	(0.040)	(0.036)
Bad correction	-0.518***	-0.508***	-0.100***	-0.551***	-0.550***	-0.172***	-0.476***	-0.457***	-0.011
	(0.031)	(0.034)	(0.036)	(0.039)	(0.044)	(0.047)	(0.049)	(0.050)	(0.052)
Female partner	-0.003	-0.001	0.009	-0.018	-0.009	0.004	0.016	0.007	0.016
	(0.016)	(0.017)	(0.014)	(0.020)	(0.021)	(0.018)	(0.024)	(0.026)	(0.021)
Partner's contribution			0.084***			0.089***			0.080***
			(0.003)			(0.004)			(0.004)
Individual FE		✓	✓		✓	✓		✓	✓
Good correction	0.310***	0.271***	-0.104**	0.311***	0.281***	-0.057	0.307***	0.259***	-0.157**
−Bad correction	(0.048)	(0.052)	(0.045)	(0.067)	(0.075)	(0.061)	(0.071)	(0.071)	(0.065)
Baseline mean	0.780	0.780	0.780	0.780	0.780	0.780	0.778	0.778	0.778
Baseline SD	0.414	0.414	0.414	0.414	0.414	0.414	0.416	0.416	0.416
Adj. R-squared	0.104	0.100	0.335	0.134	0.111	0.369	0.074	0.090	0.306
Observations	3180	3180	3180	1670	1670	1670	1510	1510	1510
Clusters	464	464	464	244	244	244	220	220	220

Notes: This table presents the regression results of equation 6. Columns 1-3 include all participants' willingness to collaborate but column 1 excludes partner's contribution and individual fixed effects and column 2 partner's contribution. Columns 4-6 present the corresponding results for women and columns 7-9 for men. Columns 1, 4, and 7 show when we do not control for between-participants variation, the coefficient estimate on good correction is underestimated. Columns 2, 5, and 8 show when we do not control for partner's contribution, the coefficient estimate on bad correction is negative and very large, validating my experimental design that participants correctly understand the notion of good and bad moves and that participants are more willing to collaborate with partners who contributed more. Column 3, 6, and 9 show that participants are more willing to collaborate with partners who make good moves, but less willing to collaborate with those who made a correction, both good and bad. Baseline mean and standard deviation are that of male partners who do not make any corrections. Standard errors in parentheses are clustered at the individual level. Significance levels: * 10%, ** 5%, and *** 1%.

Table 4: Response to corrections made by women vs. men

Outcome:	Willing to collaborate (yes=1, no=0)								
Sample:		All			Female			Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Good correction	-0.205***	-0.230***	-0.187***	-0.283***	-0.312***	-0.248***	-0.109*	-0.127**	-0.104*
	(0.038)	(0.042)	(0.035)	(0.052)	(0.060)	(0.045)	(0.057)	(0.057)	(0.053)
Bad correction	-0.523***	-0.523***	-0.176***	-0.542***	-0.554***	-0.218***	-0.491***	-0.468***	-0.104
	(0.044)	(0.048)	(0.051)	(0.055)	(0.062)	(0.064)	(0.074)	(0.077)	(0.076)
Female partner	-0.002	-0.001	0.001	-0.028	-0.020	-0.002	0.028	0.021	0.003
	(0.016)	(0.017)	(0.022)	(0.021)	(0.023)	(0.032)	(0.025)	(0.026)	(0.030)
Partner's contribution			0.083***			0.089***			0.077***
			(0.004)			(0.006)			(0.006)
Good correction x Female partner	-0.007	-0.014	-0.035	0.083	0.085	0.035	-0.114	-0.133	-0.119*
	(0.052)	(0.056)	(0.044)	(0.065)	(0.076)	(0.057)	(0.082)	(0.081)	(0.067)
Bad correction x Female partner	0.010	0.027	0.144**	-0.022	0.005	0.090	0.028	0.023	0.168
	(0.062)	(0.067)	(0.070)	(0.074)	(0.081)	(0.093)	(0.103)	(0.110)	(0.102)
Partner's contribution x Female partner			0.002			-0.001			0.006
			(0.005)			(0.008)			(0.007)
Individual FE		✓	✓		✓	✓		✓	✓
Baseline mean	0.780	0.780	0.780	0.780	0.780	0.780	0.778	0.778	0.778
Baseline SD	0.414	0.414	0.414	0.414	0.414	0.414	0.416	0.416	0.416
Adj. R-squared	0.104	0.100	0.336	0.134	0.111	0.369	0.075	0.091	0.307
Observations	3180	3180	3180	1670	1670	1670	1510	1510	1510
Clusters	464	464	464	244	244	244	220	220	220

Notes: This table presents the regression results of equation 7. Columns 1-3 include all participants' willingness to collaborate but column 1 excludes partner's contribution and individual fixed effects and column 2 partner's contribution. Columns 4-6 present the corresponding results for women and columns 7-9 for men. The coefficient estimate on the interaction between partner's contribution and female partner in columns 3, 6, and 9 shows that people – both women and men – do not underestimate women's contribution when selecting a collaborator. The coefficient estimates on the interaction between good correction and female partner and between bad correction and female partner suggest that women's corrections do not receive stronger negative reactions from either women or men. Baseline mean and standard deviation are that of male partners who do not make any corrections. Standard errors in parentheses are clustered at the individual level. Significance levels: * 10%, ** 5%, and *** 1%.

Table 5: Response to corrections of high vs. low ability people

Outcome:	Willing to collaborate (yes=1, no=0)								
Sample:		All			Female			Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Good correction	-0.166***	-0.195***	-0.155***	-0.215***	-0.242***	-0.208***	-0.117**	-0.152***	-0.107***
	(0.034)	(0.038)	(0.030)	(0.048)	(0.058)	(0.042)	(0.048)	(0.050)	(0.041)
Bad correction	-0.524***	-0.520***	-0.100**	-0.567***	-0.579***	-0.201***	-0.482***	-0.462***	0.005
	(0.039)	(0.044)	(0.047)	(0.051)	(0.059)	(0.064)	(0.059)	(0.063)	(0.063)
Female partner	-0.003	-0.002	0.009	-0.019	-0.010	0.002	0.016	0.006	0.014
	(0.016)	(0.016)	(0.014)	(0.020)	(0.021)	(0.018)	(0.024)	(0.025)	(0.021)
Partner's contribution			0.084***			0.089***			0.082***
			(0.004)			(0.005)			(0.004)
High ability	0.006			-0.007			0.020		
	(0.015)			(0.020)			(0.023)		
Good correction x High ability	-0.103*	-0.102*	-0.118**	-0.056	-0.059	-0.048	-0.149*	-0.134*	-0.180**
	(0.056)	(0.060)	(0.050)	(0.075)	(0.086)	(0.066)	(0.086)	(0.080)	(0.075)
Bad correction x High ability	0.014	0.031	0.000	0.038	0.075	0.074	0.004	0.001	-0.061
	(0.063)	(0.068)	(0.072)	(0.081)	(0.090)	(0.095)	(0.102)	(0.103)	(0.109)
Partner's contribution x High ability			-0.001			-0.001			-0.003
			(0.005)			(0.007)			(0.008)
Individual FE		✓	✓		✓	✓		✓	✓
Baseline mean	0.780	0.780	0.780	0.780	0.780	0.780	0.778	0.778	0.778
Baseline SD	0.414	0.414	0.414	0.414	0.414	0.414	0.416	0.416	0.416
Adj. R-squared	0.105	0.101	0.336	0.134	0.110	0.368	0.075	0.090	0.308
Observations	3180	3180	3180	1670	1670	1670	1510	1510	1510
Clusters	464	464	464	244	244	244	220	220	220

Notes: This table presents the regression results of equation 8. Columns 1-3 include all participants' willingness to collaborate but column 1 excludes partner's contribution and individual fixed effects and column 2 partner's contribution. Columns 4-6 present the corresponding results for women and columns 7-9 for men. The coefficient estimate on the interaction between good correction and high ability in columns 3, 6, and 9 shows that even high ability people respond negatively to good corrections with men responding more negatively, which suggests i) correcting colleagues are really costly and that (ii) negative response to corrections is irrational. Baseline mean and standard deviation are that of male partners who do not make any corrections. Standard errors in parentheses are clustered at the individual level. Significance levels: * 10%, ** 5%, and *** 1%.

Online Appendix

Appendix A Pros and cons of the quasi-laboratory format

On top of logistical convenience and complying with the COVID pre-caution measures, the quasi-laboratory format has an additional benefit over physical laboratory experiments in that participants cannot see each other when they enter the laboratory which adds an additional layer of anonymity among participants. A drawback is that participants can potentially be distracted while participating.

However, unlike standard online experiments such as on MTurk and Prolific where participants' identity is fully anonymous by the platforms' rule, we have participants' personal information and participants know it as we recruit them from our standard laboratory subject pool. Also, they are connected to us via Zoom throughout the experiment. These mostly prevent participants' attrition that can be endogenous to their decisions or treatments and the main problem of online interactive experiments (Arechar, Gächter, and Molleman 2018) and experiments where treatments affect the probability of attrition, e.g., experiments with intertemporal decision making. In my experiment, we experienced no participant attrition. A drawback is that we could not collect a large number of observations.

Another benefit of quasi-laboratory experiments over standard online experiments is that we can screen participants based on their participation status in previous experiments. This allows us to collect cleaner data; in particular, this allows us to screen out participants who have participated in experiments with deception, which is another problem of online experiments (Arechar, Gächter, and Molleman 2018).

There are already a few other studies that use a quasi-laboratory format, for example, Goeschl, Oestreich, and Soldà (2021).

Appendix B Results with the original contribution measure

In the main text, I changed the definition of contribution from the one in the pre-analysis plan because there was truncation in the original contribution measure in more than 10% of the puzzle. Nonetheless, the same results hold when I use original contribution measure; see Online Appendix Tables B1, B2, and B3. Although the original measure is relative to one's pair while the measure I use in this paper is absolute, whether a measure is relative or absolute does not matter because I add individual fixed effects.

Table B1: Response to corrections: Original contribution measure

Outcome:	Willing to collaborate (yes=1, no=0)								
Sample:		All			Female			Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Good correction	-0.208***	-0.238***	-0.272***	-0.240***	-0.269***	-0.304***	-0.169***	-0.197***	-0.230***
	(0.028)	(0.030)	(0.026)	(0.037)	(0.043)	(0.035)	(0.041)	(0.040)	(0.038)
Bad correction	-0.518***	-0.508***	-0.160***	-0.551***	-0.550***	-0.234***	-0.476***	-0.457***	-0.065
	(0.031)	(0.034)	(0.037)	(0.039)	(0.044)	(0.048)	(0.049)	(0.050)	(0.054)
Female partner	-0.003	-0.001	0.008	-0.018	-0.009	0.004	0.016	0.007	0.012
	(0.016)	(0.017)	(0.014)	(0.020)	(0.021)	(0.019)	(0.024)	(0.026)	(0.022)
Partner's contribution			1.192***			1.164***			1.234***
			(0.058)			(0.078)			(0.084)
Individual FE		✓	✓		✓	✓		✓	✓
Good correction	0.310***	0.271***	-0.112**	0.311***	0.281***	-0.070	0.307***	0.259***	-0.165**
−Bad correction	(0.048)	(0.052)	(0.045)	(0.067)	(0.075)	(0.061)	(0.071)	(0.071)	(0.067)
Baseline mean	0.780	0.780	0.780	0.780	0.780	0.780	0.778	0.778	0.778
Baseline SD	0.414	0.414	0.414	0.414	0.414	0.414	0.416	0.416	0.416
Adj. R-squared	0.104	0.100	0.314	0.134	0.111	0.330	0.074	0.090	0.300
Observations	3180	3180	3180	1670	1670	1670	1510	1510	1510
Clusters	464	464	464	244	244	244	220	220	220

Notes: This table reports the same estimation results as Table 3 but with original contribution measure specified in the pre-analysis plan, and show that the results are robust to using the original measure. Baseline mean and standard deviation are that of male partners who do not make any corrections. Standard errors in parentheses are clustered at the individual level. Significance levels: * 10%, ** 5%, and *** 1%.

Table B2: Response to corrections made by women vs. men: Original contribution measure

Outcome:	Willing to collaborate (yes=1, no=0)								
Sample:		All			Female			Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Good correction	-0.205***	-0.230***	-0.268***	-0.283***	-0.312***	-0.338***	-0.109*	-0.127**	-0.180***
	(0.038)	(0.042)	(0.038)	(0.052)	(0.060)	(0.049)	(0.057)	(0.057)	(0.058)
Bad correction	-0.523***	-0.523***	-0.232***	-0.542***	-0.554***	-0.275***	-0.491***	-0.468***	-0.167**
	(0.044)	(0.048)	(0.054)	(0.055)	(0.062)	(0.069)	(0.074)	(0.077)	(0.084)
Female partner	-0.002	-0.001	-0.072	-0.028	-0.020	-0.090	0.028	0.021	-0.063
	(0.016)	(0.017)	(0.052)	(0.021)	(0.023)	(0.074)	(0.025)	(0.026)	(0.072)
Partner's contribution			1.109***			1.070***			1.147***
			(0.085)			(0.114)			(0.125)
Good correction x Female partner	-0.007	-0.014	-0.008	0.083	0.085	0.063	-0.114	-0.133	-0.090
	(0.052)	(0.056)	(0.046)	(0.065)	(0.076)	(0.061)	(0.082)	(0.081)	(0.071)
Bad correction x Female partner	0.010	0.027	0.143*	-0.022	0.005	0.085	0.028	0.023	0.188*
	(0.062)	(0.067)	(0.077)	(0.074)	(0.081)	(0.108)	(0.103)	(0.110)	(0.105)
Partner's contribution x Female partner			0.163			0.182			0.168
			(0.113)			(0.159)			(0.157)
Individual FE		✓	✓		✓	✓		✓	✓
Baseline mean	0.780	0.780	0.780	0.780	0.780	0.780	0.778	0.778	0.778
Baseline SD	0.414	0.414	0.414	0.414	0.414	0.414	0.416	0.416	0.416
Adj. R-squared	0.104	0.100	0.314	0.134	0.111	0.331	0.075	0.091	0.301
Observations	3180	3180	3180	1670	1670	1670	1510	1510	1510
Clusters	464	464	464	244	244	244	220	220	220

Notes: This table reports the same estimation results as Table 4 but with original contribution measure specified in the pre-analysis plan, and show that the results are robust to using the original measure. Baseline mean and standard deviation are that of male partners who do not make any corrections. Standard errors in parentheses are clustered at the individual level. Significance levels: * 10%, ** 5%, and *** 1%.

Table B3: Response to corrections of high vs. low ability people: Original contribution measure

Outcome:	Willing to collaborate (yes=1, no=0)								
Sample:		All			Female			Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Good correction	-0.166***	-0.195***	-0.216***	-0.215***	-0.242***	-0.269***	-0.117**	-0.152***	-0.168***
	(0.034)	(0.038)	(0.033)	(0.048)	(0.058)	(0.044)	(0.048)	(0.050)	(0.047)
Bad correction	-0.524***	-0.520***	-0.178***	-0.567***	-0.579***	-0.291***	-0.482***	-0.462***	-0.055
	(0.039)	(0.044)	(0.051)	(0.051)	(0.059)	(0.068)	(0.059)	(0.063)	(0.069)
Female partner	-0.003	-0.002	0.007	-0.019	-0.010	0.002	0.016	0.006	0.010
	(0.016)	(0.016)	(0.014)	(0.020)	(0.021)	(0.018)	(0.024)	(0.025)	(0.021)
Partner's contribution			1.196***			1.196***			1.222***
			(0.073)			(0.101)			(0.098)
High ability	0.006			-0.007			0.020		
	(0.015)			(0.020)			(0.023)		
Good correction x High ability	-0.103*	-0.102*	-0.137***	-0.056	-0.059	-0.079	-0.149*	-0.134*	-0.192**
	(0.056)	(0.060)	(0.052)	(0.075)	(0.086)	(0.070)	(0.086)	(0.080)	(0.076)
Bad correction x High ability	0.014	0.031	0.049	0.038	0.075	0.144	0.004	0.001	-0.051
	(0.063)	(0.068)	(0.072)	(0.081)	(0.090)	(0.094)	(0.102)	(0.103)	(0.108)
Partner's contribution x High ability			-0.004			-0.070			0.064
			(0.118)			(0.154)			(0.183)
Individual FE		✓	✓		✓	✓		✓	✓
Baseline mean	0.780	0.780	0.780	0.780	0.780	0.780	0.778	0.778	0.778
Baseline SD	0.414	0.414	0.414	0.414	0.414	0.414	0.416	0.416	0.416
Adj. R-squared	0.105	0.101	0.315	0.134	0.110	0.331	0.075	0.090	0.303
Observations	3180	3180	3180	1670	1670	1670	1510	1510	1510
Clusters	464	464	464	244	244	244	220	220	220

Notes: This table reports the same estimation results as Table 5 but with original contribution measure specified in the pre-analysis plan, and show that the results are robust to using the original measure. Baseline mean and standard deviation are that of male partners who do not make any corrections. Standard errors in parentheses are clustered at the individual level. Significance levels: * 10%, ** 5%, and *** 1%.

Appendix C Pre-analysis plan

Gender differences in the cost of contradiction
Pre-analysis plan

Yuki Takahashi

November 22, 2020

This document pre-specifies the main hypotheses, the experimental design, and the empirical specifications for a laboratory experiment that examines gender differences in the cost of contradiction. At the time this document is written, I ran 1 pilot session (with 16 participants) to make sure that the experimental design and procedure worked without any problems.

1 Main Hypotheses

H1: Men are less likely to work with a woman than with a man who contradicts them.
H2: The behavior conjectured in H1 leads to a suboptimal partner choice.
H3: A mechanism that underlies the behavior conjectured in H1 is gender bias.

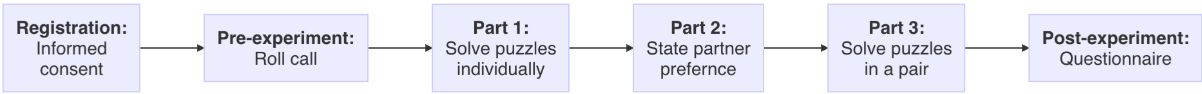
2 Design and procedure

The experiment will be computerized and conducted online with the University of Bologna's students in Italian. However, unlike standard online experiments, I will conduct the experiment as a "quasi-laboratory" where participants will be connected with the experimenter via Zoom throughout the experiment and listen to the instructions the experimenter will read out, ask questions to the experimenter via private chat, etc., just like the standard laboratory experiment. Their camera and microphone will be turned off throughout the experiment except when the experimenter calls their name at the beginning of the experiment (explained later).

Based on the power simulation in appendix A, I will recruit approximately 450 participants (225 female and 225 male). Each session will consist of a multiple of 8 participants and is expected to last for 1 hour. The average total payment per participant will be 10€, the maximum 25€, and the minimum 2€, all including the 2€ participation fee.

I use Isaksson (2018)'s 3x3 sliding puzzle as the real effort task for this experiment and define the difficulty (the number of moves away from the solution), good moves (a move that reduces the number of moves away from the solution), and bad moves (a move that increases the number of moves away from the solution) by the Breadth-First Search algorithm.

Figure 1: Flowchart of the experiment



The experiment will consist of 3 parts as summarized in figure 1. The details are below:

Registration

1. Upon receiving the invitation email to the experiment, participants will register for a session they want to participate in and upload their ID documents as well as a signed consent form. I will recruit a few more participants than I will need for a given session in case some participants would not show up to the session.

Pre-experiment

2. On the day and the time of the session they have registered, the participants will enter the Zoom waiting room. They receive a link to the oTree virtual room and enter their first name, last name, and their email they have used in the registration. They also draw a virtual coin that is numbered from 1 to 40.
3. Then I admit participants to the Zoom meeting room one by one and rename them by the first name they have entered on the oTree. If there is more than one participant with the same first name, I will add a number after their first name (e.g. Giovanni2).
4. After admitting all the participants, I will do roll call: I will call participants' first names and ask them to respond via microphone to ensure other participants that the called participants' first names correspond to their gender. If there are more participants than I would need to run the session, I will draw random numbers from 1 to 40 and ask those who drew the coins with the same number to leave. Those who will leave the session will receive the participation fee.

Part 1: Individual round

5. Participants will work on the puzzle individually with an incentive (0.2€ per puzzle solved). They can solve as many puzzles as possible with increasing difficulty (but maximum of 15 puzzles) in 4 minutes. This part will familiarize them with and measure their ability to solve the puzzle. The ability is measured by the number of puzzles they solve.

Part 2: Partner preference elicitation

6. Participants will be told the rules of part 3 and state their partner preference. This part will proceed as follows: participants will be grouped into 8 participants based on their ability similarity, then each participant will be randomly matched with another participant in the same group and solve 1 puzzle together by alternating their move. Which participant will make the first move will be randomized and this will be told to both participants. If they cannot solve the puzzle within 2 minutes, they will finish the puzzle without solving it. Reversing the matched participant's move will be used as the measure of contradiction. The matched participant's first name will be displayed on the computer screen throughout the puzzle to subtly inform that participant's gender. Each participant's contribution to a given puzzle is measured as defined in appendix C.
7. Once they finish the puzzle, participants will state whether they want to work with the matched participant (yes/no), which will be used as the measure of their partner preference.

Then they will be randomly re-matched with another participant with a perfect stranger algorithm and repeat point 6 with a different puzzle with the same difficulty and state their partner preference.

8. After all the participants solve the puzzle with all the other participants in the same group and state their partner preference, participants are matched according to the following algorithm:
   (a) 1 participant is randomly chosen
   (b) if they have a match (both them and the other person state "yes" when they are matched) they will work together in part 3
   (c) if they have more than 1 matches, 1 of the matches is randomly chosen
   (d) the match is excluded and (a)-(c) is repeated until there is no match
   (e) if some participants are still left unmatched, they are matched randomly

   Part 3: Group round

9. The matched participants will work together on the puzzles by alternating their move for 12 minutes and earn 1€ for each puzzle solved. Which participant will make the first move will be randomized at each puzzle and this will be told to both participants as in part 2. They can solve as many puzzles as possible with increasing difficulty (but maximum of 20 puzzles).

   Post-experiment

10. Participants will answer a short questionnaire which consists of (i) the 6 hostile and benevolent sexism questions in Stoddard, Karpowitz, and Preece (2020) which is originally from Glick and Fiske (1996) and measure gender bias,[1] and (ii) their basic demographic information and what they have thought about the experiment (see appendix B for the questions asked). I will ask them these questions in this order.
11. After participants answer all the questions, I will tell them their earnings and let them leave the virtual room and Zoom. They will receive their earnings via PayPal.

3 Specification

Test of H1 I test H1 by estimating the following OLS regression using male participants' partner preference observations elicited in part 2. I call participants who state their partner preference as decision-makers, participants who are evaluated by the decision-makers as participants:

$$\begin{aligned} Prefer_{ij} = \beta_1 Contradict_{ij} * Female_j + \beta_2 Contradict_{ij} + \beta_3 Female_j \\ + \delta Contribute_{ij} + IndividualFE_i + \epsilon_{ij} \end{aligned} \tag{1}$$

- $Prefer_{ij} \in \{0, 1\}$: a dummy variable indicating whether decision maker i preferred participant j as their partner.
- $Contradict_{ij} \in \{0, 1, ...\}$: the number of times j reverses i's move.

1. The Italian translation is from Manganelli Rattazzi, Volpato, and Canova (2008) and Rollero, Glick, and Tartaglia (2014). I score the participants' answer following Stoddard, Karpowitz, and Preece (2020) (assign 0 to strongly disagree and 4 to strongly agree, take the arithmetic average of all the 6 questions, and divide it by 24).

- $Female_j \in \{0, 1\}$: an indicator variable equals 1 if participant j is female, 0 otherwise.
- $IndividualFE_i$: fixed effects for decision-maker i. This is necessary for identification for 2 reasons. First, i's unobserved characteristics can affect both j's puzzle play (j's contradiction and contribution) and the probability that i prefers j as a partner. Second, the wealth effect is different across i because each i can earn a different amount in part 1.
- $Contribute_{ij} \in [0, 1]$: participant j's contribution to a puzzle played with decision-maker i as defined in appendix C. This is necessary for identification so that I can compare women and men who contradict i and make the same contribution. I add this variable as a linear term because the outcome must be increasing in j's contribution.

β_1 compares decision-makers' partner preference for female vs male participants who make the same number of contradictions and tests H1:

- $\beta_1 < 0$: men are less likely to work with a woman than with a man who contradicts them (so yes to H1).
- $\beta_1 > 0$: men are more likely to work with a woman than with a man who contradicts them (so no to H1).
- $\beta_1 = 0$: men are neither more nor less likely to work with a woman than with a man who contradicts them (so no to H1).

Test of H2 To test H2, I separate the effect of good contradictions in equation 1 by estimating the following OLS regression using the same sample as test of H1.

$$\begin{aligned} Prefer_{ij} = &\beta_1 Contradict_{ij} * Female_j + \beta_2 Contradict_{ij} + \beta_3 Female_j \\ &+ \beta_4 ContradictGood_{ij} * Female_j + \beta_5 ContradictGood_{ij} \\ &+ \delta Contribute_{ij} + IndividualFE_i + \epsilon_{ij} \end{aligned} \tag{2}$$

- $ContradictGood_{ij} \in \{0, 1, ...\}$: the number of times j reverses i's bad move.

other variables are as defined in equation 1.

β_4 picks up the part of β_1 in equation 1 that comes from j's good contradiction and tests H2:

- $\beta_4 < 0$: the behavior conjectured in H1 leads to a suboptimal partner choice (so yes to H2).
- $\beta_4 > 0$: the behavior conjectured in H1 leads to an optimal partner choice (so no to H2).
- $\beta_4 = 0$: the behavior conjectured in H1 leads to neither a suboptimal nor an optimal partner choice (so no to H2).

Test of H3 To test H3, I interact the contradictions, participants' gender, and their interaction with decision-makers' gender bias in 1 by estimating the following OLS regression using the same sample as test of H1.

$$\begin{aligned} Prefer_{ij} = &\beta_1 Contradict_{ij} * Female_j + \beta_2 Contradict_{ij} + \beta_3 Female_j \\ &+ \beta_4 Contradict_{ij} * Female_j * StrongerBias_i + \beta_5 Contradict_{ij} * StrongerBias_i \\ &+ \beta_6 Female_j * StrongerBias_i + \delta Contribute_{ij} + IndividualFE_i + \epsilon_{ij} \end{aligned} \tag{3}$$

- $StrongerBias_i \in \{0, 1\}$: an indicator variable equals 1 if decision-maker i's gender bias measured by the 6 hostile and benevolent sexism questions in the post-experimental questionnaire is above median of all the male decision-makers, 0 otherwise.

other variables are as defined in equation 1.

β_4 tests whether the behavior conjectured in H1 is stronger among decision-makers with stronger gender bias and tests H3:

- $\beta_4 < 0$: the behavior conjectured in H1 is stronger among decision-makers with stronger gender bias (so yes to H3).
- $\beta_4 > 0$: the behavior conjectured in H1 is weaker among decision-makers with stronger gender bias (so no to H3).
- $\beta_4 = 0$: the behavior conjectured in H1 is neither stronger nor weaker among decision-makers with stronger gender bias (so no to H3).

Standard error adjustment Because the treatment unit is i, I cluster standard error at i. Although the same individual appears twice (once as i and once as j), j is passive in preference elicitation.

Unsolved puzzles I include pairs who could not solve the puzzle.

Notes about the tests of H2 and H3 Interpreting the tests for H2 and H3 may require cautions. First, both tests are likely to be underpowered because they further split the effect of H1 for which the sample size is determined. Second, only for the test of H3, participants may not answer the gender bias questions honestly because gender is a socially sensitive issue, so the test may not be able to detect the effect even if H3 is true.

References

Glick, Peter, and Susan T. Fiske. 1996. "The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism." *Journal of Personality and Social Psychology* 70 (3): 491–512.

Isaksson, Siri. 2018. *It Takes Two: Gender Differences in Group Work.* Working Paper.

Manganelli Rattazzi, Anna Maria, Chiara Volpato, and Luigina Canova. 2008. "L'Atteggiamento ambivalente verso donne e uomini: Un contributo alla validazione delle scale ASI e AMI. [Ambivalent attitudes toward women and men: Contribution to the validation of ASI and AMI scales.]" *Giornale Italiano di Psicologia* 35 (1): 217–243.

Rollero, Chiara, Peter Glick, and Stefano Tartaglia. 2014. "Psychometric properties of short versions of the Ambivalent Sexism Inventory and Ambivalence Toward Men Inventory." *TPM-Testing, Psychometrics, Methodology in Applied Psychology* 21 (2): 149–159.

Stoddard, Olga, Christopher F. Karpowitz, and Jessica Preece. 2020. *Strength in Numbers: A Field Experiment in Gender, Influence, and Group Dynamics.* Working Paper.

Appendix A Power simulation

I estimate the number of participants I have to recruit to achieve 80% power for the test of H1 via Monte Carlo simulation.

I assume the following data generating process:

$$
\begin{aligned}
Prefer^*_{ij} = b_0 + b_1 Contradict_{ij} * Female_{ij} + b_2 Contradict_{ij} + b_3 Female_{ij} \\
+ \delta Contribute_{ij} + \sum_{k=1}^{3} \gamma^k \mathbb{1}(a_i = k) + \sum_{k=1}^{3} \theta^k \mathbb{1}(m_i = k) + e_{ij} \\
(i = 1, ..., N; j = 1, ..., 7)
\end{aligned}
\tag{A1}
$$

where each variable is drawn from the following distribution:

- $Contradict_{ij} \sim Pois(0.1\frac{L}{2} + 0.02(m_i - 1)\frac{L}{2})$ (10% of moves were reversed following Isaksson (2018); the meaner the decision-maker, the more likely they receive a contradiction)
- $Female_{ij} \sim^{iid} Bernoulli(0.5)$ (a matched participant is female by 50% chance)
- $Contribute_{ij} \sim TN(0.5 - 0.1(a_i - 1.5), 0.05, 0, 1)$ (a matched participant's contribution which negatively depends on the decision-maker's ability)
- $a_i \sim^{iid} Unif\{1, 3\}$ (the decision-maker's ability)
- $m_i \sim^{iid} Unif\{1, 3\}$ (the decision-maker's meanness)
- $e_{ij} \sim^{iid} N(0, \sigma^2)$ (large sample approximation)
- $Prefer_{ij} = \mathbb{1}(Prefer^*_{ij} > 0)$

Each parameter is defined as follows:

- $b_0 = 0$ (so that the unconditional probability that the decision-maker chooses a matched participant is 50%)
- $b_1 = MDE$
- $b_2 = MDE$ (being contradicted by a female participant reduces the probability of choosing that participant as a partner twice as much as being contradicted by a male participants)
- $b_3 = 0$ (the decision-maker has no underlying gender bias)
- $\delta = 0.2$ (this is the main determinant of partner preference: the higher a matched participant's contribution, the higher the probability that the decision-maker chooses them as a partner)
- $\gamma^k = -0.02 * (k - 1.5)$, k=1,2,3 (the higher the decision-maker's ability, the lower the probability that the decision-maker chooses a matched participant as a partner)
- $\theta^k = -0.02 * (k - 1.5)$, k=1,2,3 (the meaner the decision-maker, the lower the probability that the decision-maker chooses a matched participant as a partner)
- $\sigma = 0.1$

where L is total number of moves the decision-maker and a matched participant take to solve a puzzle, which I assume to be 15 (7.5 moves by the decision-maker). However, I also set it to 10 (5 moves by the decision-maker) for robustness check. $MDE = -0.02$ is my baseline assumption (being contradicted once reduces the probability of choosing a matched participant by the same degree as when the matched participant's contribution is 0.1 lower), but I also set it to -0.01 for robustness check, -0.03 to see what happens in a more optimistic scenario, and 0 to check that type I error rate is kept at 5% and that the estimated ATE is 0 when there is no underlying effect.

Thus, I estimate equation 1 with the sample drawn from equation A1 for $MDE \in \{0, -0.01, -0.02, -0.03\}$, $L \in \{15, 10\}$, and $N \in [50, 300]$. I draw 1000 independent sample.

Power is defined as the number of times the t-test rejects β_1 at 5% significance level (two-tailed) divided by the number of samples I draw:

$$Power(N, MDE, L) = \frac{\#Rejections(N, MDE, L)}{\#Draws} \tag{A2}$$

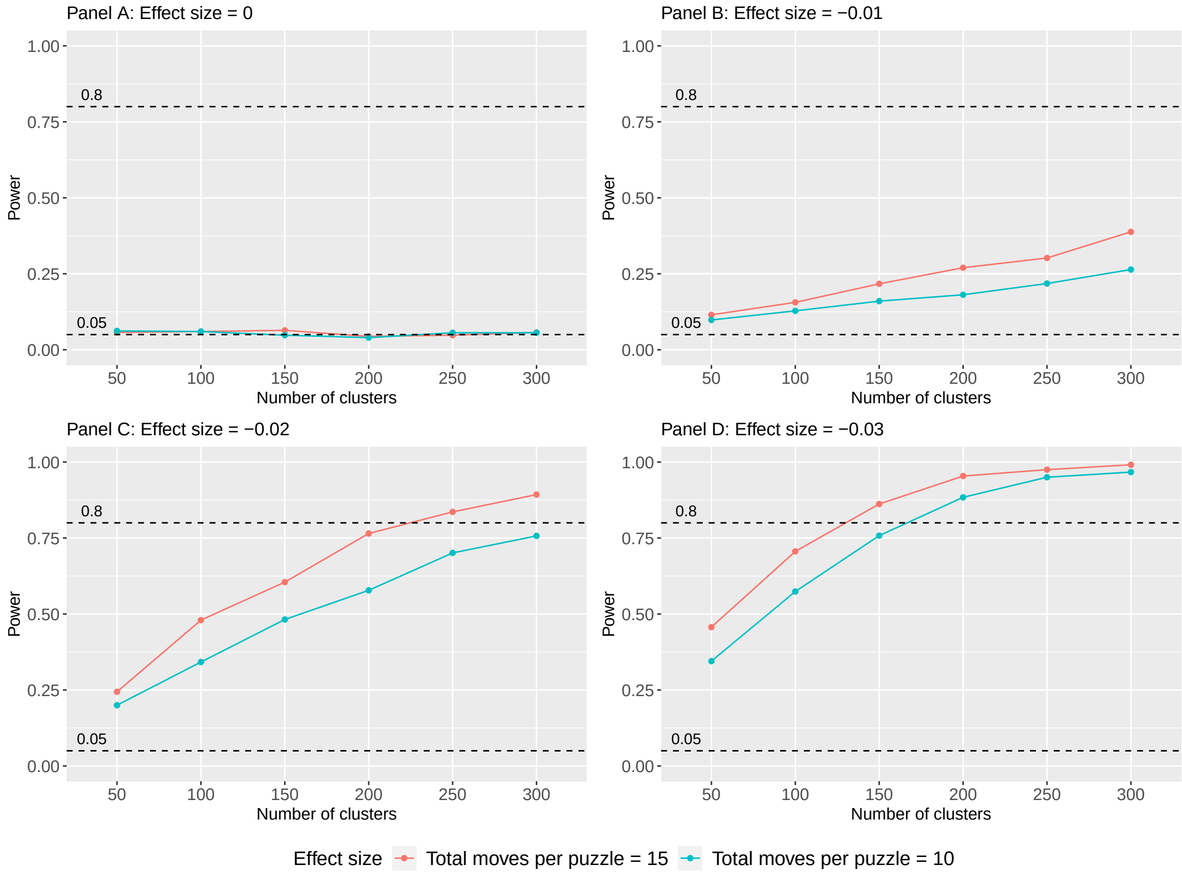
ATE is defined as the average of $\hat{\beta}_1$ across draws (its dependence on L is due to the non-linearity of the data generating process):

$$ATE(MDE, L) = \frac{\sum_{r=1}^{\#Draws} \hat{\beta}_1^r(MDE, L)}{\#Draws} \tag{A3}$$

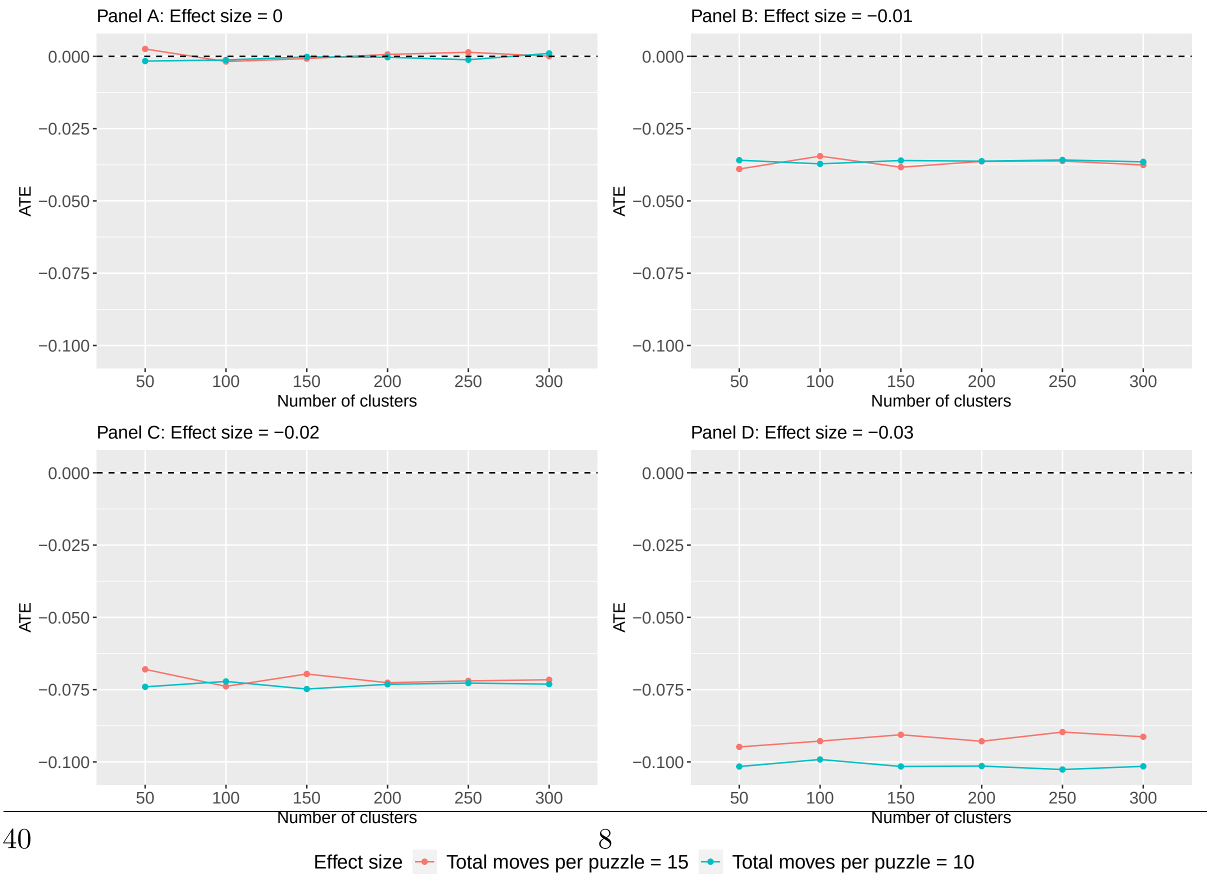
The results are presented in figure A1, which suggests that I need to recruit about 450 participants (so that I could have 225 clusters for testing H1). First, in the baseline scenario with $L = 15$, I can achieve about 80% power. Second, even under a tougher scenario where $L = 10$, I can still achieve about 60% power. The type I error rate is kept at 5%. ATE is larger than b_1 in magnitude because the data generating process is non-linear, but is 0 when the underlying effect size is 0. However, the power is very sensitive to the underlying effect size: if $MDE = -0.01$, I will likely not be able to detect the effect. If $MDE = -0.03$, on the other hand, my test is very high-powered: the power is close to 100% that I will almost always be able to detect the effect.

Figure A1: Estimated power and ATE (# draws=1000, $\alpha = 0.05$ two-tailed)

(a) Estimated power



(b) Estimated ATE



Appendix B Questions asked in the questionnaire

English version

- Your age: [Integer]
- Your gender: [Male, Female]
- Your region of origin: [Northwest, Northeast, Center, South, Islands, Abroad]
- Your major: [Humanities, Law, Social Sciences, Natural Sciences/Mathematics, Medicine, Engineering]
- Your degree program: [Bachelor, Master/Post-bachelor, Bachelor-master combined (1st, 2nd, or 3rd year), Bachelor-master combined (4th year or beyond), Doctor]
- What do you think this study was about? [Textbox]
- Was there anything unclear or confusing about this study? [Textbox]
- Were the puzzles difficult? [Difficult, Somewhat difficult, Just right, Somewhat easy, Easy]
- Do you have any other comments? (optional) [Textbox]

Italian translation

- Età: [Integer]
- Sesso: [Uomo, Donna]
- Regione di origine: [Nord-Ovest, Nord-Est, Centro, Sud, Isole, Estero]
- Campo di studi principale: [Studi umanistici, Giurisprudenza, Scienze sociali, Scienze naturali/Matematica, Medicina, Ingegneria]
- Tipo di corso: [Laurea, Laurea Magistrale/Post-Laurea, Ciclo Unico (1 °, 2 ° o 3 ° anno), Ciclo Unico (4 ° anno o oltre), Dottorato]
- Cosa pensi di questo studio? [Textbox]
- C'era qualcosa di poco chiaro o di confuso in questo studio? [Textbox]
- I puzzle erano difficili? [Difficili, Abbastanza difficili, Giusto, Abbastanza facili, Facili]
- Hai qualche altro commento? (opzionale) [Textbox]

Appendix C Calculation of contribution

Following Isaksson (2018), I define a participant's contribution to a given puzzle in part 2 as follows:

$$\text{Player i's contribution} = \frac{P_i}{P_i + P_j} \in [0, 1],\ i, j = 1, 2,\ i \neq j \tag{C1}$$

$$P_i = \max\{\text{i's \# good moves} - \text{i's \# bad moves}, 0\}\ i = 1, 2 \tag{C2}$$

If $P_i = 0$ and $P_j = 0$, I define both i's and j's contribution to 0.

Appendix D Experimental instructions: English translation

App: pt0

Page: Reg

Registration

Please fill out the following information in order for us to pay you after the session. Please make sure that they correspond to the information you registered on ORSEE.

N.B. Please capitalize only the first letter of your first name and last name.

Good examples: Marco Rossi; Maria Bianchi; Anna Maria Gallo

Bad examples: MARCO ROSSI; maria bianchi; Anna maria Gallo

- First name: [Textbox]
- Last name: [Textbox]
- Email address registered on ORSEE: [Textbox]

[Check if there are any same first names. If so, add an integer (starting from 2) at the end of the first name]

Page: Draw

Draw a coin

Please draw a virtual coin by clicking the button below.

[Draw]

[Assign random number ranging from 1 to 40]

Page: Wait

Your coin

You drew the following coin.

Please wait until the session starts.

Page: Excess

Please click an appropriate button

[I was chosen to participate] [I was chosen to leave]

Page: Intro

General instructions

<u>**Overview**</u>: This study will consist of **3 parts** and a follow-up survey and is expected to take **1 hour**. At the beginning of each part, you will receive specific instructions, followed by a set of understanding questions. You must answer these understanding questions correctly to proceed.

<u>**Your payment**</u>: For completing this study, you are guaranteed **2€** for your participation, but can earn up to **25€** depending on how good you are at the tasks. The tasks involve solving sliding puzzles, like the one shown below.

puzzle_2_0.png

<u>**Confidentiality**</u>: Other people participating in this study can see your first name. Aside from your first name, other participants will not see any information about you. **At the conclusion of the study, all identifying information will be removed and the data will be kept confidential**. If there is more than one participant with the same first name, we add a number at the end of your first name (e.g. Marco2).

<u>**General rules**</u>: During the study, please turn off your camera and microphone, and do not communicate with anyone other than us. Also, please do not reload the page or close your browser because it may make your puzzle unsolvable. If you have any questions or face any problems, please send us a private chat on Zoom.

<u>**App: pt1**</u>

Page: Intro

Instructions for part 1 out of 3

In this part, you will solve the puzzle alone to familiarize yourself with it. You can solve as many puzzles as possible (but a maximum of 15 puzzles) in **4 minutes**. You will earn **0.2€ for each puzzle** you solve.

Your goal is to move the tiles and order them as follows:

puzzle_goal.png

Before you start, please go through the three examples below to understand how to solve the puzzle.

Example 1:

First, consider the following puzzle.

puzzle_1.png

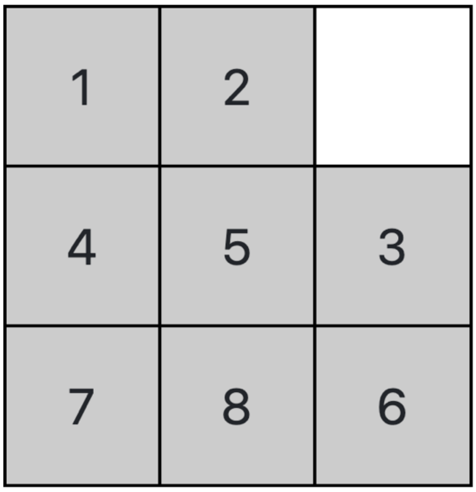
You can only move the tiles next to an empty cell and the tile you choose is moved to the empty cell. So, in this puzzle, there are 3 moves you can make: move 3 down, move 5 right, and move 6 up.

Among the 3 moves, moving 6 up is the only correct move: by moving 6 up, you can solve the puzzle. The other moves do not solve the puzzle.

When you click a tile next to an empty cell, the tile will be moved to the empty cell. So, in this case, you should click 6 to move it up.

Example 2:

Next, consider the following puzzle.



puzzle_2_0.png

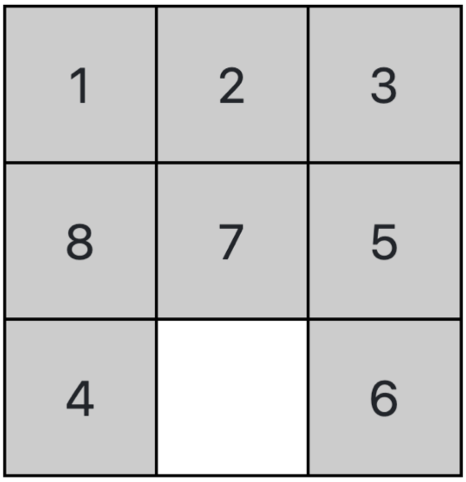
First, there are 2 moves you can make: move 2 right and move 3 up. Which moves should you make?

Observe that the only tiles that are not in the correct order are 3 and 6. So, you should move 3 up.

After moving 3 up, the puzzle will look like the one in example 1. Then you should move 6 up and the puzzle will be solved.

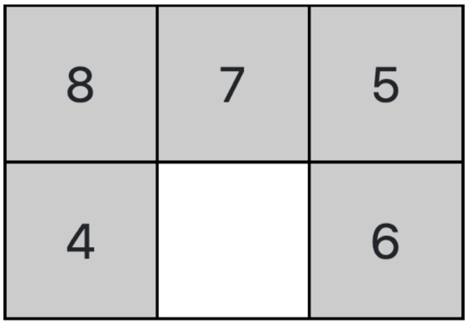
<u>Example 3</u>:

Finally, consider the following puzzle.



puzzle_3_0.png

This puzzle is a bit complicated but observe that the top row is already in the correct order. So, let's keep the top row as is, and think about the remaining part. **When the top row is in the correct order, you should always keep it as is**. So, think of this puzzle as the following simpler puzzle.



puzzle_3_0_2x3.png

You could solve the puzzle by trial and error. However, **after making the top row in the correct order, you should next make the left column in the correct order** to solve the puzzle faster. There are two moves you can make: move 4 right and move 7 down. Which is the faster way to make the left column in the correct order?

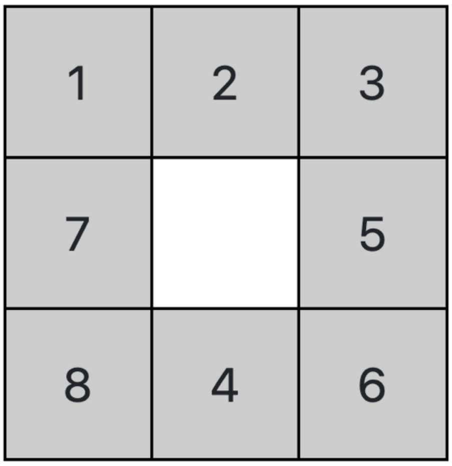
Let's try moving 4 right.

puzzle_3_1_bad_0.png

Now the only tile you can move is 8. So, let's move it down.

puzzle_3_1_bad_1.png

Now, if you ignore the top row which is already in the correct order, the only tile you can move is 7. So, let's move it to the left.

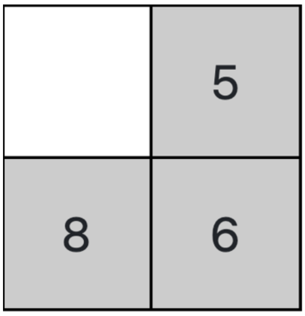


puzzle_3_1_bad_2.png

Then move 4 up, move 8 right, and move 7 down. Then you have made the left column in the correct order. You have moved tiles seven times until now.

puzzle_3_1_bad_3.png

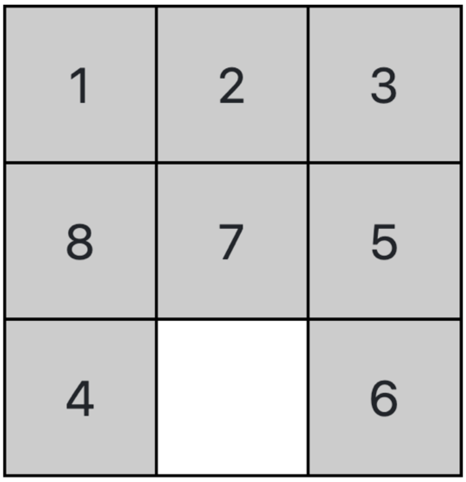
Now let’s also keep the left column as is.



puzzle_3_1_bad_3_2x2.png

Then you can solve the puzzle by moving 5 left and then 6 up. With this method, **you have moved tiles nine times in total**.

Let’s go back to the initial puzzle.

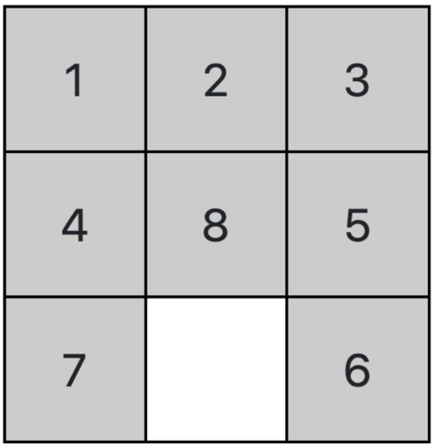


puzzle_3_0.png

This time, let's try moving 7 down.

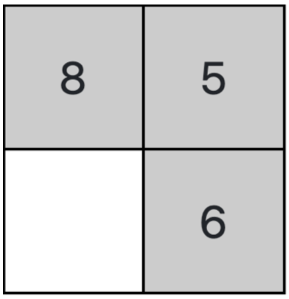
puzzle_3_1_good.png

Then move 8 right, 4 up, and 7 left. Now you have made the left column in the correct order only with four moves.



puzzle_3_4_good.png

Let's keep the left column as is (as well as the top row).



puzzle_3_4_good_2x2.png

Now it's easy to solve the puzzle: move 8 down, 5 left, and 6 up. With this method, **you have only moved tiles seven times in total**.

Because there is a time limit, it's better to solve the puzzle with the minimum number of moves. **We call a move a good move if it makes a puzzle closer to the solution, and a bad move if it makes a puzzle far from the solution. There are no neutral moves: all moves are either good or bad.**

In summary: when you solve the puzzle, first make the top row in the correct order, then make the left column in the correct order. Always try to make the number of moves as small as possible.

<u>**Understanding questions**</u>:

Before you proceed, please answer the following understanding questions. After you answer, please click Next.

1. Which of the following statements is true?

- ✓In this part, I will work on the puzzles individually for 4 minutes and earn 0.2€ for each puzzle I solve.
- In this part, I will work on the puzzles in pairs for 4 minutes and earn 0.2€ for each puzzle we solve.
- In this part, I will work on the puzzles individually for 4 minutes, but I will not earn anything.

2. Which of the following puzzles is in the correct order?

- A
- ✓B

A

puzzle_2_0.png

B

puzzle_goal.png

3. What is the strategy you should use to solve the puzzle as fast as possible?

- First, make the left column in the correct order, then the bottom row. Always minimize the number of moves I make.
- First, make the top row in the correct order, then the right column. Always minimize the number of moves I make.
- ✓First, make the top row in the correct order, then the left column. Always minimize the number of moves I make.

4. Look at the following puzzle. Which is the good move?

- Move 4 down.
- ✓Move 7 left.

puzzle_3_3_good.png

5. Consider the puzzle in question 4. What is the minimum number of moves to solve the puzzle?

- 2
- 3
- ✓4

6. Look at the following puzzle. Which is the good move?

- ✓Move 5 left.
- Move 8 up.

puzzle_3_5_good.png

7. Consider the puzzle in question 6. What is the minimum number of moves to solve the puzzle?

- ✓2
- 3
- 4

Page: Ready

Be ready

[5 seconds time count]

Please be ready for the individual round.

Page: Game

Individual round

[4 minutes time count]

[max. 15 puzzles with increasing difficulty]

Page: Proceed

The individual round is over

The individual round is over. You have solved **xx puzzles**.

Please click Next to proceed.

App: pt2

Page: Intro

Instructions for part 2 out of 3

In this part, you will **choose your partner for part 3**, the next part.

Although you will not earn anything in this part, it is important to choose the best partner possible: in part 3, you will work on the puzzles for 12 minutes in a pair by moving the tiles in turn, and both you and your partner will earn 1€ for each puzzle you two solve. There is a maximum of 20 puzzles you and your partner can solve (so the maximum earning is 20€).

You will **meet 7 other people** participating in this session one by one and solve 1 puzzle together by moving tiles in turn as you would do in part 3. One of you will be randomly chosen to make the first move at the beginning of each puzzle. You will have a **2-minute limit** for each puzzle.

After solving the puzzle, you will **choose whether you want to work with this person in part 3 too**. This person or other people in this session will not see your choice. **You can choose as many people as you want**.

After you meet all the 7 people and state your choices, we will check all the choices you and the 7 other people have made, and decide each person's partner for part 3 as follows:

1. We randomly choose 1 person out of you and the other 7 people. Call this person Giovanni.
2. We then check if Giovanni has a "match": among people Giovanni has chosen, we check whether these people also have chosen Giovanni. If there is such a person, we make Giovanni and this person as partners for part 3.
3. If Giovanni has more than one match, we randomly choose one of the matches and make them as partners for part 3.
4. If Giovanni has not chosen anyone, the people Giovanni has chosen have not chosen Giovanni, or those people already have their partner, we put Giovanni on a waiting list and repeat points 1-3 above.
5. After we choose all people, we randomly match people on the waiting list as partners for part 3.

So, **even if you choose a particular person, you may not be able to work with that person in part 3**. So, choose everyone whom you want to work with in part 3.

<u>**Understanding questions**</u>:

Before you proceed, please answer the following understanding questions. After you answer, please click Next.

1. Which of the following statements is true?

- ✓In this part, I will choose my partner for part 3.
- In this part, I will work on the puzzles for 12 minutes in a pair by moving the tiles in turn.

2. How many people can you choose whom you want to work with in part 3?

- 1 person.
- 2 people.
- ✓As many people as you want.

3. Why is it important to choose the best partner for part 3?

- ✓ because how many puzzles I can solve in part 3 depends on my partner's moves.
- because my partner will solve puzzles for me.

4. Suppose you have chosen Giovanni and Valeria. However, while Valeria has chosen you, Giovanni has not. If we have randomly chosen you first, who will be your partner for part 3?

- Giovanni
- ✓Valeria
- Someone on the waiting list
- Randomly chosen from Giovanni and Valeria

5. Suppose you have chosen Giovanni and Valeria. However, unlike question 4, while Giovanni has chosen you, Valeria has not. If we have randomly chosen you first, who will be your partner for part 3?

- ✓Giovanni
- Valeria
- Someone on the waiting list
- Randomly chosen from Giovanni and Valeria

6. Suppose you have chosen Giovanni and Valeria. Also, both Giovanni and Valeria have chosen you. If we have randomly chosen you first, who will be your partner for part 3?

- Giovanni
- Valeria
- Someone on the waiting list
- ✓Randomly chosen from Giovanni and Valeria

7. Suppose you have chosen Giovanni and Valeria. Also, both Giovanni and Valeria have chosen you. However, we already matched Valeria with Giovanni before we choose you. Who will be your partner for part 3?

- Giovanni
- Valeria
- ✓Someone on the waiting list
- Randomly chosen from Giovanni and Valeria

8. Suppose you have not chosen anyone. Also, both Giovanni and Valeria have chosen you. If we have randomly chosen you first, who will be your partner for part 3?

- Giovanni
- Valeria

- ✓Someone on the waiting list
- Randomly chosen from Giovanni and Valeria

9. Suppose you have chosen Giovanni and Valeria. However, neither Giovanni nor Valeria has chosen you. If we have randomly chosen you first, who will be your partner for part 3?

- Giovanni
- Valeria
- ✓Someone on the waiting list
- Randomly chosen from Giovanni and Valeria

Page: Puzzle

Puzzle 1/2/3/4/5/6/7 out of 7

You are playing the puzzle with **[this person's ID]**

[2 minutes time count]

Page: Pref

Puzzle 1/2/3/4/5/6/7 out of 7

You have played the puzzle with **[this person's ID]**. Do you want to work with [this person's ID] in part 3?

[Yes, No]

<u>App: pt3</u>

Page: Partner

Your partner for part 3

Based on your and the 7 other people's choices, **[the partner's ID]** became your partner for part 3.

Page: Intro

Instructions for part 3 out of 3

In this part, you will work on the puzzles with your partner for **12 minutes** by moving the tiles in turn, and both you and your partner will earn **1€ for each puzzle** you two solve. There is a maximum of 20 puzzles you and your partner can solve (so the maximum earning is 20€). As in part 2, one of you will be randomly chosen to make the first move at the beginning of each puzzle.

<u>Understanding questions</u>:

Before you proceed, please answer the following understanding questions. After you answer, please click Next.

1. Which of the following statements is true?

- ✓In this part, you and your partner will both earn 1€ for each puzzle you two solve, which means you will earn 1€ for each puzzle you two solve.
- In this part, you and your partner will earn 1€ for each puzzle you two solve, which means you will earn 0.5€ for each puzzle you two solve.

2. You and your partner…

- ✓will work on the puzzles for 12 minutes by moving the tiles in turn. Which of you will make the first move is randomly determined at the beginning of each puzzle.
- will work on the puzzles for 12 minutes. Which of you will make the first move is randomly determined at the beginning of this part and fixed afterward.

Page: Ready

Be ready

[5 seconds time count]

Please be ready for the group round.

Page: Game

Puzzle 1/2/3/…/20

Your partner: **[the partner's ID]**

[12 minutes time count]

[max. 20 puzzles with increasing difficulty]

Page: Proceed

The group round is over

The group round is over. You have solved **xx puzzles**.

Please click Next to proceed.

App: pt4

Page: Intro

A follow-up survey

As the last task, we will ask you a series of questions in which there are no right or wrong answers. We are only interested in your personal opinions. We are interested in what characteristics are associated with people's behaviors in this study. **The answers you provide will in no way affect your earnings in this study and are kept confidential.**

Please click Next to start the survey.

Page: SurveyASI

Survey page 1 out of 2

Below is a series of statements concerning men and women and their relationships in contemporary society. Please indicate the degree to which you agree or disagree with each statement.

- Women are too easily offended.
- Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for "equality."
- Men should be willing to sacrifice their own wellbeing in order to provide financially for the women in their lives.
- Many women have a quality of purity that few men possess.
- No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman.
- Women exaggerate problems they have at work.

[Choices: Strongly agree, Agree a little, Neither agree nor disagree, Disagree a little, Strongly disagree]

Page: SurveyDem

Survey page 2 out of 2

Please tell us about yourself and your opinion about this study.

- Your age: [Integer]
- Gender: [Male, Female]
- Region of origin: [Northwest, Northeast, Center, South, Islands, Abroad]
- Field of study: [Humanities, Law, Social Sciences, Natural Sciences/Mathematics, Medicine, Engineering]
- Degree program: [Bachelor, Master/Post-bachelor, Bachelor-master combined (1st, 2nd, or 3rd year), Bachelor-master combined (4th year or beyond), Doctor]
- What do you think this study was about? [Textbox]
- Was there anything unclear or confusing about this study? [Textbox]
- Were the puzzles difficult? [Difficult, Somewhat difficult, Just right, Somewhat easy, Easy]
- Do you have any other comments? (optional) [Textbox]

Page: ThankYou

Thank you for your participation

Thank you for your participation. You have completed the study.

Your earnings:

- **2€** for your participation.
- **xx.x€** for the puzzles you solved in part 1.
- **xx€** for the puzzles you and your partner solved in part 3.

Thus, you have earned **xx.x€** in this study. We will pay you your earnings via PayPal within 2 weeks. If you haven't received your earnings after 2 weeks, please contact us.

Optional: If you would like to know the results of this study, we are more than happy to send you the working paper via email once we finish this study.

[No, I do not want to receive the working paper] [Yes, I want to receive the working paper]

App: pt99

Page: ThankYou

Thank you for showing up

Thank you for showing up in this study. You will receive the show up fee of **2€** via PayPal within 2 weeks. If you haven't received your earnings after 2 weeks, please contact us.

Appendix E Screenshots

Pre-experiment

Pagina di registrazione 1 di 2

Compila i seguenti campi. Assicurati che corrispondano alle informazioni che hai registrato su ORSEE in modo da poterti inviare il pagamento dopo la sessione.

N.B. Inserisci in maiuscolo solo la prima lettera del tuo nome e cognome.

Buoni esempi: Marco Rossi; Maria Bianchi; Anna Maria Gallo

Cattivi esempi: MARCO ROSSI; maria bianchi; Anna maria Gallo

Nome:

Cognome:

Indirizzo e-mail che hai registrato su ORSEE:

Successivo

Lancia una moneta virtuale

Lancia una moneta virtuale facendo click sul pulsante in basso.

Lancia

La tua moneta virtuale

Hai ottenuto questo numero.

Si prega di attendere, la sessione inizierà a breve.

Fai click su un pulsante appropriato

Sono stato scelto per partecipare | Sono stato scelto per andarmene

Grazie per la tua presenza

Grazie per la tua presenza in questo studio. Ti pagheremo **2€** per la presenza tramite PayPal entro 2 settimane. Se non hai ricevuto i tuoi guadagni dopo 2 settimane, contattaci.

Istruzioni generali

Indicazioni generali: questo studio è composto da **3 parti** e da un questionario finale. Il tutto durerà circa **1 ora**. All'inizio di ogni parte riceverai istruzioni specifiche, seguite da una serie di domande di comprensione. Devi rispondere correttamente a queste domande di comprensione per procedere.

Il tuo pagamento: ti sono garantiti **2€** per la tua partecipazione, ma puoi guadagnare fino a **25€** a seconda di quanto sarai bravo a risolvere esercizi. Questo consistono nel risolvere dei puzzle scorrevoli, come quello mostrato qui di seguito.

1	2	
4	5	3
7	8	6

Privacy: le altre persone che partecipano a questo studio possono vedere il tuo nome. A parte il tuo nome, gli altri partecipanti non vedranno alcuna altra informazione su di te. **Alla conclusione dello studio, tutte le informazioni di identificazione verranno rimosse e i dati saranno trattati in maniera riservata**. Se c'è più di un partecipante con lo stesso nome, aggiungiamo un numero alla fine del tuo nome (ad esempio, Marco2).

Regole generali: durante lo studio, si prega di spegnere la videocamera e il microfono, e non comunicare con nessuno tranne noi. Inoltre, si prega di non aggiornare la pagina o chiudere il browser in quanto potrebbe rendere il puzzle irrisolvibile. In caso di domande o di problemi, inviaci una chat privata su Zoom.

Part 1

Istruzioni per la parte 1 di 3

In questa parte, risolverai il puzzle in maniera individuale per familiarizzare con esso. Hai **4 minuti** per risolvere il maggior numero di puzzle possibile (fino a un massimo di 15 puzzle). Guadagnerai **0,2€ per ogni puzzle** risolto.

Il tuo obiettivo è spostare le tessere e ordinarle come segue:

1	2	3
4	5	6
7	8	

Prima di iniziare, segui i tre esempi seguenti per capire come risolvere il puzzle.

Esempio 1:

Per prima cosa, considera il seguente puzzle.

1	2	3
4	5	
7	8	6

Puoi solo spostare le tessere vicine a una cella vuota e la tessera che scegli viene spostata automaticamente nella cella vuota. Ad esempio, in questo puzzle ci sono 3 mosse che puoi fare: muovere la tessera con il 3 in basso, muovere quella con il 5 a destra, oppure muovere la tessera con il 6 in alto.

Tra le 3 mosse, spostare la tessera con il numero 6 verso l'alto è l'unica mossa corretta: spostando il 6 in su, puoi risolvere il puzzle. Le altre mosse non risolvono il puzzle.

Quando fai click su una tessera vicina a una cella vuota, la tessera verrà spostata automaticamente nella cella vuota. Quindi, in questo caso, dovresti fai click sul 6 per spostarlo verso l'alto.

Esempio 2:

Ora, considera il seguente puzzle.

1	2	
4	5	3
7	8	6

Innanzitutto, ci sono 2 mosse che puoi fare: muovere il 2 a destra o muovere il 3 in alto. Quale mossa dovresti fare?

Osserva che le uniche tessere che non sono nell'ordine corretto sono il 3 e il 6. Quindi, prima di tutto devi spostare il 3 in su.

Dopo aver spostato il 3 in su, il puzzle apparirà come quello nell'esempio 1. Quindi dovresti spostare il 6 in su e il puzzle sarà risolto.

Esempio 3:

Infine, considera il seguente puzzle.

1	2	3
8	7	5
4		6

Questo puzzle è un po' complicato, ma osserva che la riga superiore è già nell'ordine corretto. Quindi, manteniamo la riga superiore così com'è e pensiamo alla parte rimanente. **Quando la riga superiore è nell'ordine corretto, dovresti sempre mantenerla così com'è**. Quindi, pensa a questo puzzle come al seguente puzzle più semplice.

8	7	5
4		6

Potresti risolvere il puzzle per tentativi ed errori. Tuttavia, c'è anche una regola generale per risolvere il puzzle più velocemente: **dopo aver sistemato la riga superiore nell'ordine corretto, dovresti procedere sistemando la colonna di sinistra nell'ordine corretto**. In questo esempio ci sono due mosse possibili: muovere il 4 a destra o muovere il 7 in basso. Qual è il modo più veloce per sistemare la colonna di sinistra nell'ordine corretto in base a quanto abbiamo appena detto?

Proviamo a spostare il 4 a destra.

1	2	3
8	7	5
	4	6

Ora l'unica tessera che puoi spostare è l'8. Quindi spostiamola verso il basso.

1	2	3
	7	5
8	4	6

Ora, se ignori la riga in alto che è già nell'ordine corretto, l'unica tessera che puoi spostare è la 7. Quindi, spostiamola a sinistra.

1	2	3
7		5
8	4	6

Poi, sposta il 4 in alto, l'8 a destra e infine il 7 in basso. Così facendo hai sistemato la colonna sinistra nell'ordine corretto. Hai spostato le tessere sette volte fino ad ora.

1	2	3
4		5
7	8	6

Ora lasciamo anche la colonna di sinistra così com'è.

	5
8	6

Quindi puoi risolvere il puzzle spostando il 5 a sinistra e poi il 6 in alto. Con questo metodo, **hai spostato le tessere nove volte in totale**.

Torniamo al puzzle iniziale.

1	2	3
8	7	5
4		6

Questa volta, proviamo a spostare il 7 verso il basso.

1	2	3
8		5
4	7	6

Poi sposta di l'8 a destra, il 4 in alta e, infine, il 7 a sinistra. Ora hai sistemato la colonna di sinistra nell'ordine corretto solo con quattro mosse.

1	2	3
4	8	5
7		6

Manteniamo la colonna di sinistra così com'è (così come la riga superiore).

8	5
	6

Ora è facile risolvere il puzzle: sposta l'8 in basso, il 5 a sinistra e, infine, il 6 in alto. Con questo metodo, **hai spostato le tessere solo sette volte in totale**.

Poiché esiste un limite di tempo, è meglio risolvere il puzzle con il numero minimo di mosse. **Definiremo una mossa come "giusta" se rende un puzzle più vicino alla soluzione, e una mossa come "sbagliata" se rende un puzzle più lontano dalla soluzione. Non ci sono mosse neutre: tutte le mosse sono giuste o sbagliate**.

In sintesi: quando risolvi il puzzle, per prima cosa, sistema le celle della prima la riga in alto nell'ordine corretto, poi procedi sistemando le celle della colonna di sinistra nell'ordine corretto. Cerca sempre di ridurre il più possibile il numero di mosse.

Domande di comprensione:
Prima di procedere, rispondi alle seguenti domande di comprensione. Dopo aver risposto, fai click su Successivo.

1. Quale delle seguenti affermazioni è vera?

○ In questa parte, lavorerò sui puzzle individualmente per 4 minuti e guadagnerò 0,2€ per ogni puzzle che risolvo.

○ In questa parte, lavorerò sugli enigmi in coppia per 4 minuti e guadagnerò 0,2€ per ogni puzzle che risolviamo.

○ In questa parte, lavorerò individualmente sui puzzle per 4 minuti, ma non guadagnerò nulla.

2. Quale dei seguenti puzzle è nell'ordine corretto?

○ A.

○ B.

A

1	2	
4	5	3
7	8	6

B

1	2	3
4	5	6
7	8	

3. Qual è la strategia da utilizzare per risolvere il puzzle il più velocemente possibile?

○ Per prima cosa, sistema le celle della colonna di sinistra nell'ordine corretto, quindi quelle della riga in basso. Riduci sempre al minimo il numero di mosse da fare.

○ Per prima cosa, sistema le celle della riga superiore nell'ordine corretto, quindi quelle della colonna di destra. Riduci sempre al minimo il numero di mosse da fare.

○ In primo luogo, sistema le celle della riga superiore nell'ordine corretto, quindi quelle della colonna di sinistra. Riduci sempre al minimo il numero di mosse da fare.

4. Guarda il seguente puzzle. Qual è la mossa giusta?

○ Sposta il 4 in basso.

○ Spostare il 7 a sinistra.

1	2	3
4	8	5
	7	6

5. Considera il puzzle in questione 4. Qual è il numero minimo di mosse per risolvere il puzzle?

○ 2

○ 3

○ 4

6. Guarda il seguente puzzle. Qual è la mossa giusta?

- ○ Spostare il 5 a sinistra.
- ○ Sposta l'8 in alto.

1	2	3
4		5
7	8	6

7. Considera il puzzle in questione 6. Qual è il numero minimo di mosse per risolvere il puzzle?

- ○ 2
- ○ 3
- ○ 4

Successivo

Preparati, il round individuale inizierà a breve

Tempo rimasto per completare questa pagina: **0:02**

Sii pronto per il round individuale.

Round individuale

Tempo rimasto per completare questa pagina: **3:36**

1	2	3
4	5	6
7		8

Il round individuale è finito

Il round individuale è finito. Hai risolto **12 puzzle**.

Fai click su Successivo per procedere.

Successivo

Part 2

Istruzioni per la parte 2 di 3

In questa parte, **sceglierai il tuo partner per la parte 3**, l'ultima parte dello studio.

Sebbene non guadagnerai nulla in questa parte, è importante scegliere il miglior partner possibile: nella parte 3, lavorerai su dei nuovi puzzle per 12 minuti in coppia. Tu e il tuo partner sposterete le celle del puzzle a turno fino a che non lo avrete risolto, e guadagnerete 1€ a testa per ogni puzzle che risolverete assieme. Nella parte 3 tu e il tuo partner potrete risolvere un massimo di 20 puzzle (quindi il guadagno massimo è di 20€ a testa per la parte 3).

Nella parte 2, incontrerai, una per volta, **altre 7 persone** che, come te, partecipano a questa sessione. Con ciascuna di queste 7 persone risolverai 1 puzzle spostando a turno le tessere, così come farai poi nella parte 3. Uno di voi sarà scelto a caso per fare la prima mossa all'inizio del puzzle. Avrete un **limite di 2 minuti** per ogni puzzle.

Per ognuna delle 7 persone che incontrerai, dopo aver risolto il puzzle, **ti verrà chiesto di indicare se vuoi lavorare con questa persona anche nella parte 3. Puoi scegliere quante persone vuoi tra le 7 che incontrerai. Nessuno degli altri partecipanti sarà informato delle tue scelte in questa parte**.

Dopo aver incontrato le 7 persone e aver indicato le tue preferenze, controlleremo tutte le scelte che tu e le altre 7 persone avete fatto e assegneremo un partner per la parte 3 a ciascuno di voi. Procederemo nel modo seguente:

1. Selezioneremo a caso una persona tra te e le altre 7 persone. Per esempio, chiamiamo questa persona Giovanni.
2. Verifichiamo se Giovanni ha una "corrispondenza": ossia se tra le persone che Giovanni ha indicato c'è qualcuno che - a sua volta - ha scelto Giovanni. Se c'è allora faremo in modo che Giovanni e questa persona diventino partner per la parte 3.
3. Se Giovanni ha più di una corrispondenza, selezioneremo in maniera casuale una delle corrispondenze e assegneremo un partner a Giovanni per la parte 3.
4. Se Giovanni non ha scelto nessuno, o le persone che Giovanni ha scelto non hanno a loro volta indicato Giovanni, o infine, se le persone che lo avevano indicato sono già state assegnate a un partner per la parte 3, metteremo Giovanni in lista d'attesa e ripeteremo i punti 1-3 sopra per gli altri partecipanti fino a che ognuno avrà un partner o sarà assegnato alla lista d'attesa.
5. Dopo aver assegnato tutte le persone, prenderemo coloro che sono stati posti in lista d'attesa e li abbineremo tra loro in maniera casuale, in modo che ciascuno abbia un partner per la parte 3.

Quindi, **se indichi una persona in particolare, potresti non essere in grado di lavorare con quella persona nella parte 3**. Pertanto, ti suggeriamo di indicare tutti coloro con cui sei disposto a lavorare nella parte 3.

<u>**Domande di comprensione**</u>:
Prima di procedere, rispondi alle seguenti domande di comprensione. Dopo aver risposto, fai click su Successivo.

1. Quale delle seguenti affermazioni è vera?

○ In questa parte, sceglierò il mio partner per la parte 3.

○ In questa parte, lavorerò sui puzzle per 12 minuti in coppia spostando le tessere a turno.

2. Quante persone puoi scegliere con cui vuoi lavorare nella parte 3?

○ 1 persona.

○ 2 persone.

○ Tutte le persone che vuoi.

3. Perché è importante scegliere il miglior partner per la parte 3?

○ perché quanti puzzle posso risolvere nella parte 3 dipende dalle mosse del mio partner.

○ perché il mio partner risolverà i puzzle per me.

4. Supponi di aver scelto Giovanni e Valeria. Tuttavia, mentre Valeria ha scelto te, Giovanni no. Se ti abbiamo scelto a caso per primo, chi sarà il tuo partner per la parte 3?

○ Giovanni

○ Valeria

○ Qualcuno in lista d'attesa

○ Scelto a caso da Giovanni e Valeria

5. Supponi di aver scelto Giovanni e Valeria. Tuttavia, a differenza della domanda 4, mentre Giovanni ha scelto te, Valeria no. Se ti abbiamo scelto a caso per primo, chi sarà il tuo partner per la parte 3?

- ○ Giovanni
- ○ Valeria
- ○ Qualcuno in lista d'attesa
- ○ Scelto a caso da Giovanni e Valeria

6. Supponi di aver scelto Giovanni e Valeria. Inoltre, sia Giovanni che Valeria hanno scelto te. Se ti abbiamo scelto a caso per primo, chi sarà il tuo partner per la parte 3?

- ○ Giovanni
- ○ Valeria
- ○ Qualcuno in lista d'attesa
- ○ Scelto a caso da Giovanni e Valeria

7. Supponi di aver scelto Giovanni e Valeria. Inoltre, sia Giovanni che Valeria hanno scelto te. Tuttavia, abbiamo già abbinato Valeria a Giovanni prima di scegliere te. Chi sarà il tuo partner per la parte 3?

- ○ Giovanni
- ○ Valeria
- ○ Qualcuno in lista d'attesa
- ○ Scelto a caso da Giovanni e Valeria

8. Supponi di non aver scelto nessuno. Inoltre, sia Giovanni che Valeria hanno scelto te. Se ti abbiamo scelto a caso per primo, chi sarà il tuo partner per la parte 3?

- ○ Giovanni
- ○ Valeria
- ○ Qualcuno in lista d'attesa
- ○ Scelto a caso da Giovanni e Valeria

9. Supponi di aver scelto Giovanni e Valeria. Tuttavia, né Giovanni né Valeria ti hanno scelto. Se ti abbiamo scelto a caso per primo, chi sarà il tuo partner per la parte 3?

- ○ Giovanni
- ○ Valeria
- ○ Qualcuno in lista d'attesa
- ○ Scelto a caso da Giovanni e Valeria

Successivo

Il puzzle 4 su 7

Tempo rimasto per completare questa pagina: **1:54**

Stai risolvendo il puzzle con **Giovanni**

1	2	3
8	7	5
	4	6

Aspetta il tuo partner!

Il puzzle 4 su 7

Hai risolto il puzzle con **Giovanni**. Sei disposto a lavorare con Giovanni nella parte 3?

- ○ Sì
- ○ No

Successivo

Part 3

Il tuo partner nella parte 3

In base alle tue scelte e a quelle di altre 7 persone, **Valeria** è diventato il tuo partner per la parte 3.

Successivo

Istruzioni per la parte 3 di 3

In questa parte, risolverai puzzle con il tuo partner per **12 minuti** spostando le tessere a turno, e sia tu che il tuo partner guadagnerete **1€ per ogni puzzle** che risolverete. C'è un massimo di 20 puzzle che tu e il tuo partner potete risolvere (quindi il guadagno massimo è di 20 € per ciascuno di voi). Come nella parte 2, uno di voi verrà scelto a caso per fare la prima mossa all'inizio di ogni puzzle.

Domande di comprensione:
Prima di procedere, rispondi alle seguenti domande di comprensione. Dopo aver risposto, fai click su Successivo.

1. Quale delle seguenti affermazioni è vera?

- ○ In questa parte, tu e il tuo partner guadagnerete 1€ per ogni puzzle risolto, il che significa che guadagnerete 1€ per ogni puzzle risolto.
- ○ In questa parte, tu e il tuo partner guadagnerete 1€ per ogni puzzle risolto da voi due, il che significa che guadagnerete 0,5€ per ogni puzzle risolto.

2. Tu e il tuo partner...

- ○ lavorerete sui puzzle per 12 minuti spostando le tessere a turno. Chi di voi farà la prima mossa viene determinato casualmente all'inizio di ogni puzzle.
- ○ lavorerete sui puzzle per 12 minuti. Chi di voi farà la prima mossa viene determinato casualmente all'inizio di questa parte e fissato in seguito.

Successivo

Preparati, il round con il tuo partner inizierà a breve

Tempo rimasto per completare questa pagina: **0:03**

Sii pronto per il round con il tuo partner.

Il puzzle 2

Tempo rimasto per completare questa parte: **11:32**

Il tuo partner è **Valeria**

1	2	3
4	5	
7	8	6

È il tuo turno!

Il round con il tuo partner è finito

Il round con il tuo partner è finito. Avete risolto **8 puzzle**.

Fai click su Successivo per procedere.

Successivo

Post-experiment

Questionario finale

Per concludere, ti faremo una serie di domande in cui non ci sono risposte giuste o sbagliate. Ci interessano solo le tue opinioni personali. Siamo interessati a quali caratteristiche sono associate ai comportamenti delle persone in questo studio. **Le risposte che fornisci non influiranno in alcun modo sui tuoi guadagni in questo studio e sono mantenuti riservati**.

Fai click su Successivo per avviare il questionario.

Successivo

Questionario pagina 1 di 2

Di seguito troverai una serie di affermazioni riguardanti uomini e donne e le loro relazioni nella società contemporanea. Indica per favore il grado di accordo o di disaccordo con ciascuna delle seguenti affermazioni.

	Del tutto d'accordo	**Abbastanza d'accordo**	**Nè d'accordo né in disaccordo**	**Abbastanza in disaccordo**	**Per niente d'accordo**
Le donne si offendono troppo facilmente.	○	○	○	○	○
Molte donne, sotto la veste dell'uguaglianza, cercano in realtà favoritismi, come per esempio politiche di assunzione che le favoriscano rispetto agli uomini.	○	○	○	○	○
Per mantenere economicamente le loro donne, gli uomini dovrebbero essere disposti a sacrificare il proprio benessere.	○	○	○	○	○
Molte donne hanno una qualità di purezza che pochi uomini posseggono.	○	○	○	○	○
Per quanto realizzato sia, un uomo non è mai veramente completo come persona se non ha l'amore di una donna.	○	○	○	○	○
Le donne tendono a ingigantire i problemi che hanno sul lavoro.	○	○	○	○	○

Successivo

Questionario pagina 2 di 2

Infine, ti chiediamo alcune informazioni personali ed opinioni su questo studio.

Età:

Sesso:

- Uomo
- Donna

Regione di origine:

Campo di studi principale:

Tipo di corso:

Cosa pensi di questo studio?

C'era qualcosa di poco chiaro o di confuso in questo studio?

I puzzle erano difficili?

Hai qualche altro commento? (opzionale):

Successivo

Grazie per la partecipazione

Grazie per la partecipazione. Hai completato lo studio.

I tuoi guadagni sono composti come segue:

- **2€** per la partecipazione.
- **2.4€** per i puzzle che hai risolto nella parte 1.
- **8€** per i puzzle che tu e il tuo partner avete risolto nella parte 3.

Quindi, hai guadagnato **12.4€** in totale. Ti invieremo il denaro tramite PayPal entro 2 settimane. Se non hai ricevuto i tuoi guadagni dopo 2 settimane, contattaci.

Opzionale: Se desideri conoscere i risultati di questo studio, saremo più che felici di inviarti l'articolo via e-mail una volta terminato questo studio.

No, non desidero ricevere l'articolo   Sì, desidero ricevere l'articolo