# Online Appendix for: What to blame? Self-serving attribution bias with multi-dimensional uncertainty

Alexander Coutts Leonie Gerhards Zahra Murad

January 2024

## 1 Belief Elicitation by Wave

### 1.1 Discussion

In both wave 1 and 2 we elicited beliefs about the probability that each teammate's performance was in the top 10. This can be seen in the experiment screenshots shown in Section 10.1 for wave 1. As noted in Section 2.2 of the main paper, calculation of the optimal weight requires information about the probabilities assigned to each of the four possible states: both top, both bottom, and teammate 1 top and teammate 2 bottom (and vice-versa). To derive these probabilities in wave 1, we assumed independence between beliefs about performance of the teammates.

In wave 2 we additionally elicited about the probabilities of all four states. The elicitation sequence was such that participants first had their beliefs about the two teammates' performances elicited, then the four states were pre-filled assuming independence, and lastly participants had full freedom to re-allocate the four state probabilities as they saw fit. This can be seen in the experiment screenshots shown in Section 10.2 for wave 2.

Reassuringly, 90% of the time participants chose not to alter beliefs in the four states, that is, they followed the independence assumption.[1] Strictly speaking, when faced with the $2 \times 2$ set which corresponds to each teammate being either in the top or bottom half, the elicitation procedure is only incentive compatible for the two payoff relevant states (in

[1] Independence fails to hold after feedback, which creates dependencies between beliefs about performance of the two teammates. For the 10% that reported beliefs that were inconsistent with the independence assumption, the average difference in the belief reported was 1.1 percentage points. Results are robust to excluding these observations. Piloting suggested it was not intuitive for participants to initially think about the probabilities of these four states. For this reason we first asked about the probability of teammate 1 and 2 being in the top half.

which only one of the two teammates ranked in the top half and the other in the bottom half). However, since the vast majority of participants do not alter beliefs in the four states, this suggests that participants were not strategically mis-representing beliefs in the other two states.

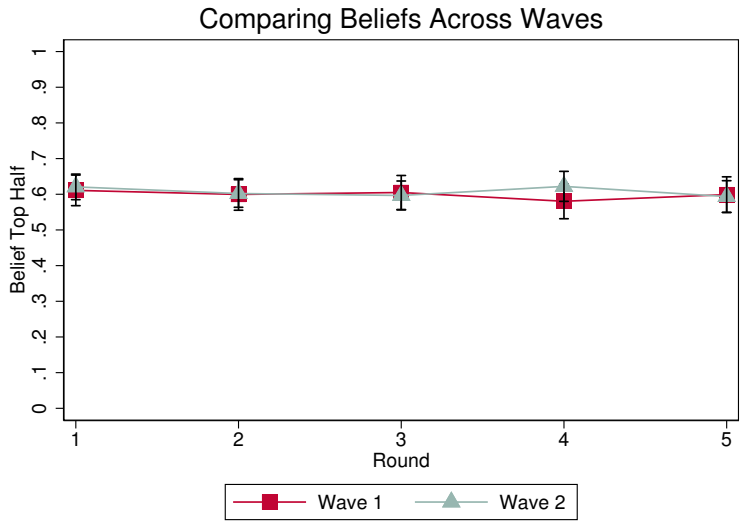
In the next subsection, we further investigate beliefs by wave. The fact that beliefs are nearly identical across the two waves suggests that participants did not alter their behaviour in response to these theoretical subtleties. This is sensible, as they are hard to perceive, but beyond this, they do not generate any additional strategic motivation to lie.

## 1.2 Beliefs by Wave

Figures OA.1 and OA.2 present average beliefs pooled across Main and Control about teammate 1 and 2 being in the top half of performers respectively. Due to slight differences in the belief elicitation we may have been concerned that these resulted in differences in reported beliefs. From these figures we can see that there are no differences in beliefs across waves, neither for teammate 1 nor 2.

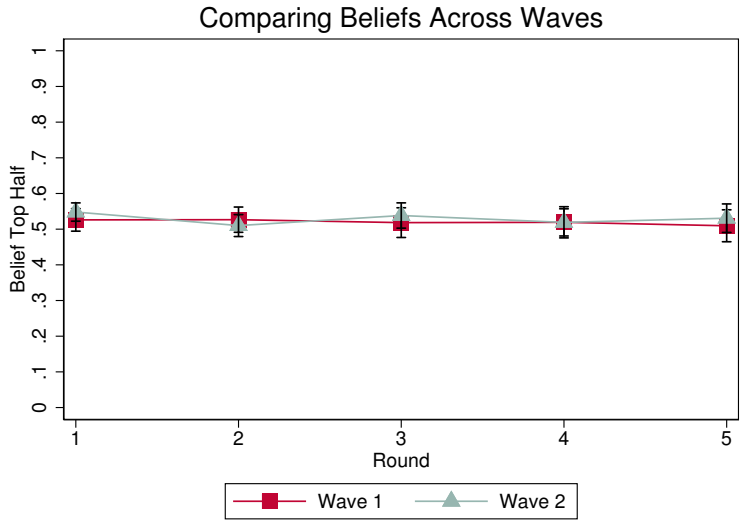
Note that the stability of beliefs in these figures belies an underlying positive asymmetry – since prior beliefs are larger than 50% and positive and negative signals are equally likely, stability is indicative of over-weighting positive relative to negative feedback. Later in this document, Section 5 examines related figures for Main versus Control.

Figure OA.1: Beliefs about Teammate 1 across Waves 1 and 2




Average beliefs about teammate 1 for rounds 1 to 5, split by wave. 95% confidence intervals shown.

Figure OA.2: Beliefs about Teammate 2 across Waves 1 and 2




Average beliefs about teammate 2 for rounds 1 to 5, split by wave. 95% confidence intervals shown.

# 2 Hard Easy Effect

In Part 1 of the experiment, participants took either an easy or a hard version of the test. The questions had similar scope, but differed in difficulty, which was assessed by the authors and confirmed by piloting of the test questions. The hard-easy effect (Larrick et al., 2007; Moore and Small, 2007) predicts that individuals will predict that their performance relative to others is greater on easy tasks relative to hard tasks.[2]

Figure OA.3 presents prior beliefs, before receiving feedback, about the probability that teammate 1 was in the top half. Recall that in the Main treatment, teammate 1 was the participant themselves, while in Control it was a third party (whose responses were visible to the participant). The hard-easy manipulation is successful, as individuals believe they are in the top half with 72% probability when the test was easy, compared with 62% when the test was hard (Wilcoxon ranksum p-value: 0.037). Interestingly, the hard-easy effect carries through to the Control treatment. There initial beliefs for the hard test are 60% and for the easy test are 52% (Wilcoxon ranksum p-value: 0.047).[3]

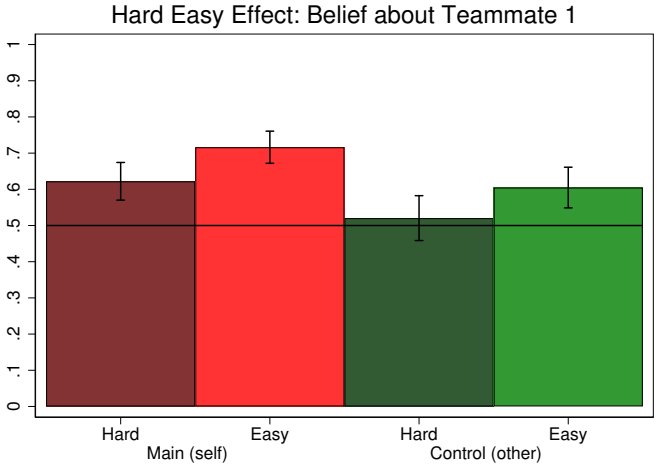
While we found hard-easy effects for teammate 1 in Main and Control, Figure OA.4 shows quite clearly that there are no effects for teammate 2. This is true in Main, when the participant is teammate 1, and also for Control, when the participant is not a member

[2]Overconfidence in relative performance is often referred to as over-placement. For estimation about absolute performance, this hard-easy effect is reversed, i.e. there is more over-estimation in hard tasks.

[3]Additionally, we continue to find evidence of overconfidence comparing Main and Control within either hard or easy sub-samples, Wilcoxon ranksum p-values are 0.014 and 0.006 respectively.

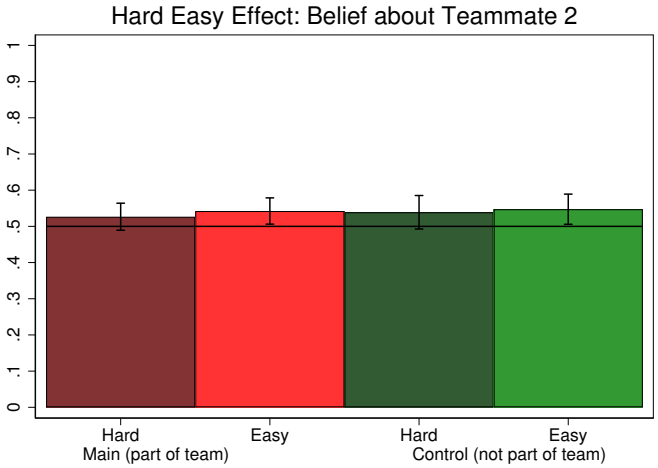
of the team. Section 6.4 studies belief updating for the hard and easy sub-samples.

Figure OA.3




Examining the hard-easy effect in initial beliefs about teammate 1. In Main treatment participant is teammate 1. In Control, participant is not on the team, but is aware of the test responses of teammate 1. 95% confidence intervals shown.

Figure OA.4




Examining the hard-easy effect in initial beliefs about teammate 2. Participant is aware of only the number of questions attempted by teammate 2. 95% confidence intervals shown.

# 3 Gender Differences

## 3.1 Prior Beliefs

For our Primary experiment sample, recruitment was balanced across gender.[4] Figure OA.5 (a) presents prior beliefs of men and women participants respectively for prior beliefs about the probability of teammate 1 being in the top half, for Main treatment (their own performance) and for the Control treatment (performance of other). It appears that men are substantially more overconfident than women about their own performance (Main), both relative to the Bayesian prediction of 0.5 and the to the Control treatment. However, there are also average performance differences on the test. On average, men scored statistically significantly better than women (25.5 versus 22.1 points; Wilcoxon ranksum p-value 0.006).[5]

To account for these potential differences we further split the sample into teammate 1's who scored in the top half and those who scored in the bottom half. Results should be interpreted cautiously, due to the smaller samples. Figures OA.5 (b) and (c) present prior beliefs, by gender, for these respective groups. Because we are conditioning on actual performance of teammate 1, the Bayesian prediction is in fact 1 for those in the top half, and 0 for those in the bottom half.[6] Thus these results show the distribution of underconfidence and overconfidence respectively. For those in the top, Figure OA.5 (b) shows significant, though more moderate underconfidence. Men are more confident than women both regarding own performance (Main, Wilcoxon ranksum test p-value: 0.004) and other performance (Control, Wilcoxon ranksum test p-value 0.016). Interestingly, for those in the top half, neither men nor women in Main are more confident about their own performance, compared to the same respective genders in Control (regarding other performance), thus this underconfidence likely reflects only the lack of full information. For those in the bottom Figure OA.5 (c) shows that there is significant overconfidence. Men *in the bottom half* believe on average there is a 66% probability they are in the

[4] We are missing gender data on 26 individuals, and 7 individuals chose not to identify as male or female. This section thus uses data from a total of 393 (out of 426) individuals, for which the percentage of women is 52%.

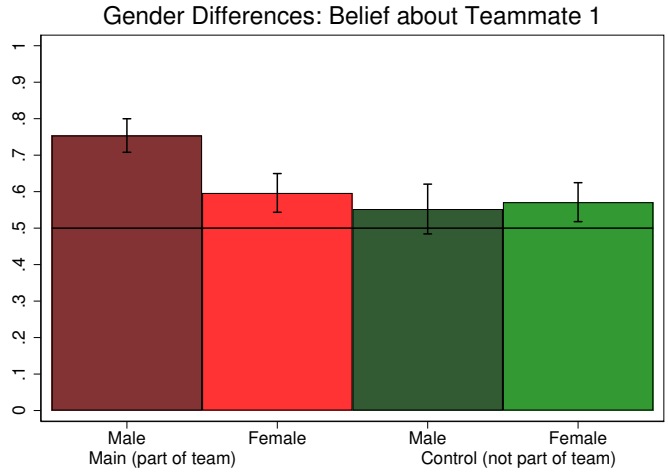
[5] As each question was worth 2.5 points and incorrect answers result in a penalty of 1.0 points, this difference is equivalent to one single correct versus wrong answer out of the 15 total questions. We note that these were not true IQ test but ad-hoc tests created with various logic and trivia questions. In our framing however, we emphasised that, "Questions similar to these are often used to measure a person's general intelligence (IQ)". It is possible that this generated pressure which may have disproportionately affected women rather than men, i.e. stereotype threat, see Spencer et al. (2016).

[6] Recall that in the Control treatment individuals observe the question responses of another different teammate 1. Thus they have additional information about performance of these individuals, explaining the large differences in reported beliefs, which are in the expected direction.

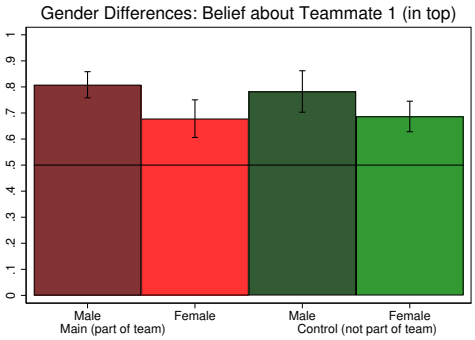
top half. For women this is also large, at 54%. The gender difference is significant at the 5% level, Wilcoxon ranksum test p-value 0.043. While this suggests overconfidence, we also note that the estimates of teammate 1 in Control are also significantly inflated, at 41% for both genders. To bolster the claim of overconfidence, we need to compare beliefs across Main and Control: indeed for men and women in the bottom half, beliefs about own performance are significantly higher than beliefs about another teammate 1 (Wilcoxon ranksum test p-values 0.000 and 0.012 respectively). Overall, overconfidence bias is strongly present for those who perform in the bottom half, and that it is significantly greater for men compared to women.

Finally Figure OA.6 presents prior beliefs, by gender, for the performance of teammate 2. Since, intuitively, there are no performance differences for teammate 2 by the participant's gender we do not resort to the sub-sample investigation in the previous figures. It suffices to report that there do not appear to be any gender differences in participant's beliefs about the performance of teammate 2 in Main or Control.

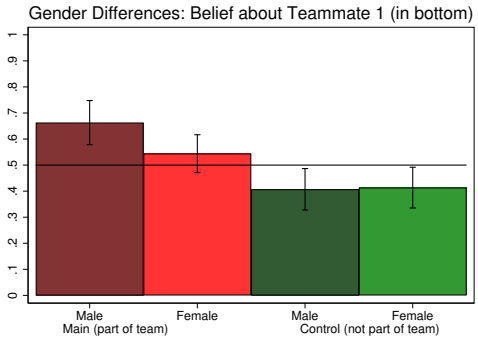
Figure OA.5: Gender Differences




(a) Full sample. Beliefs about teammate 1 being in the top half of performers. 95% confidence intervals shown.




(b) Sample restricted to teammate 1's that were in fact in the top half of performers (ranked 1-10 out of 20). 95% confidence intervals shown.

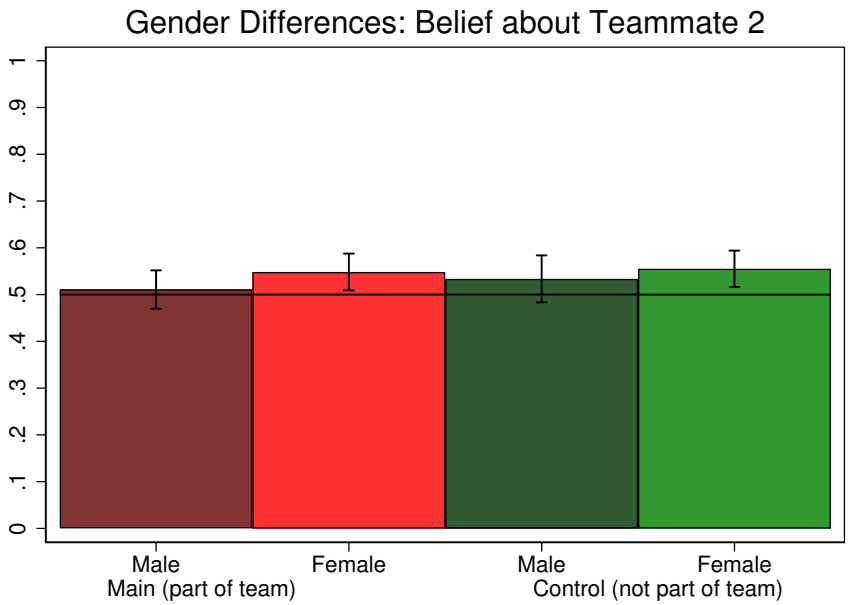



(c) Sample restricted to teammate 1's that were in fact in the bottom half of performers (ranked 11-20 out of 20). 95% confidence intervals shown.

## 3.2 Updating

After having noted significantly greater overconfidence for men relative to women, we now examine whether there are gender differences in updating these beliefs. Table OA.1 presents the main results (Table 2 in the main paper) for men and women separately. From these tables it is immediately apparent that the asymmetry in Main treatment appears to be driven by men. The positive asymmetry for men is large in magnitude, responsiveness to positive signals is 72% of a Bayesian, compared to negative signals which is 12% of a Bayesian, significant at the 1% level. In other words, men respond 6 times as much to positive feedback relative to negative feedback when it is about their own performance.

Figure OA.6




Gender differences in initial beliefs about teammate 2, before receiving feedback. In Main treatment participant is on the team, as teammate 1. In Control, participant is not on the team. In both cases participant is aware of only the number of questions attempted by teammate 2. 95% confidence intervals shown.

Examining the Control group, we find that in fact men respond slightly more to negative feedback when updating about others' performance, though the asymmetry is not significant. This observed asymmetry in Main is significantly different from that in Control at the 1% level (Chow-Test p-value 0.005). On the other hand, women exhibit very little asymmetry, and we are unable to reject that they respond equally to positive or negative signals. This is true both when they update about their own performance (Main) as well as when they update about another teammate 1's performance (Control). While men in Main respond slightly more to positive signals and significantly less to negative signals relative to men in Control, women in Main respond slightly less to positive signals and respond equivalently to negative signals relative to women in Control.

Table OA.1: Updating Beliefs about Teammate 1 (By Gender)

| | Male | | Female | |
|---|---|---|---|---|
| Regressor | Main (1) | Control (2) | Main (3) | Control (4) |
| $\delta$ | 0.779** | 0.758*** | 0.682*** | 0.740*** |
| | (0.086) | (0.069) | (0.073) | (0.058) |
| $\beta_1$ | 0.722** | 0.552*** | 0.409*** | 0.512*** |
| | (0.137) | (0.117) | (0.086) | (0.102) |
| $\beta_0$ | 0.119*** | 0.671*** | 0.390*** | 0.390*** |
| | (0.113) | (0.112) | (0.074) | (0.069) |
| P-Value ( $\delta = 1$ ) | 0.0123 | 0.0008 | 0.0000 | 0.0000 |
| P-Value ( $\beta_1 = 1$ ) | 0.0459 | 0.0003 | 0.0000 | 0.0000 |
| P-Value ( $\beta_0 = 1$ ) | 0.0000 | 0.0043 | 0.0000 | 0.0000 |
| P-Value ( $\beta_1 = \beta_0$ ) | 0.0072 | 0.4094 | 0.8796 | 0.3034 |
| $R^2$ | 0.70 | 0.62 | 0.50 | 0.59 |
| Observations | 354 | 321 | 444 | 493 |
| P-Value [Chow-test] for $\delta$ | | 0.8490 | | 0.5290 |
| P-Value [Chow-test] for $\beta_1$ | | 0.3450 | | 0.4371 |
| P-Value [Chow-test] for $\beta_0$ | | 0.0005 | | 0.9994 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | | 0.0053 | | 0.5489 |

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

Table OA.2 presents the analogue of Table 3 in the main paper, updating about teammate 2, for men and women separately. The only significant asymmetry arises for men evaluating the performance of their teammates when they themselves are a member of the team (Main). Men respond approximately 2 times as much to positive signals compared to negative signals, significant at the 5% level and echoing the direction of the overall results in the main paper. However, just as in the main paper, the asymmetry cannot be statistically distinguished from the asymmetry in the Control, in column 2. Regarding women, in columns 3 and 4 there does not appear asymmetry in response to feedback in Main or Control.[7]

[7]Given performance differences, one question is whether the asymmetric updating is driven by high performers and not men per-se. To the extent this is possible to examine, there is no clear evidence that this is the case, as studied in Section 6.3.

Table OA.2: Updating Beliefs about Teammate 2 (By Gender)

| | Male | | Female | |
|---|---|---|---|---|
| Regressor | Main | Control | Main | Control |
| | (1) | (2) | (3) | (4) |
| $\delta$ | 0.843** | 0.608*** | 0.685*** | 0.832*** |
| | (0.064) | (0.082) | (0.072) | (0.046) |
| $\beta_1$ | 0.509*** | 0.605*** | 0.290*** | 0.411*** |
| | (0.097) | (0.126) | (0.068) | (0.074) |
| $\beta_0$ | 0.255*** | 0.459*** | 0.268*** | 0.393*** |
| | (0.060) | (0.116) | (0.071) | (0.066) |
| P-Value ( $\delta = 1$ ) | 0.0154 | 0.0000 | 0.0000 | 0.0004 |
| P-Value ( $\beta_1 = 1$ ) | 0.0000 | 0.0023 | 0.0000 | 0.0000 |
| P-Value ( $\beta_0 = 1$ ) | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| P-Value ( $\beta_1 = \beta_0$ ) | 0.0421 | 0.2790 | 0.8115 | 0.8430 |
| $R^2$ | 0.64 | 0.40 | 0.46 | 0.62 |
| Observations | 444 | 370 | 482 | 531 |
| P-Value [Chow-test] for $\delta$ | | 0.0223 | | 0.0837 |
| P-Value [Chow-test] for $\beta_1$ | | 0.5428 | | 0.2258 |
| P-Value [Chow-test] for $\beta_0$ | | 0.1183 | | 0.1949 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | | 0.5545 | | 0.9734 |

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

# 4 Chosen Weights

While the primary focus of the empirical analysis is on determinants of beliefs and belief updating, it is informative to investigate how beliefs and updating affect participant's weighting decisions. Recall that individuals had to choose a weight from 0 to 1, with 0 representing all of the weight on teammate 2, and 1 representing all of the weight on teammate 1.

Table OA.3 shows regressions which examine impacts of participant characteristics and the Main treatment on weighting decisions. In both Main and Control treatments, positive feedback has a statistically significant effect on the weight chosen, in columns (1) and (2). Being a member of the team, i.e. our Main treatment, has a statistically significant effect on the chosen weight, though only in column (2).

Importantly, as columns (3) and (4) show, the positive effect of both a positive signal and the Main treatment are coming from the interaction between the two. In particular, this interaction increases the weight by 6.3 percentage points. This is about an 11% increase on the average weight chosen. Thus, when individuals are part of the team, when receiving a positive signal they increase the weight on their own performance by 6.3 percentage points.

The result that there is some evidence of a larger weight after positive signals is consistent with the results on positively biased updating. A moderate effect for positive signals is consistent with the slight over-weighting of positive signals for self relative to teammate 2. In the Control treatment the responsiveness to feedback was balanced across both teammate 1 and 2, and for both positive and negative feedback. This is consistent with the results in Table OA.3.

Table OA.3: Submitted Weight of Teammate 1 versus Teammate 2

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Initial Weight | $0.600^{***}$ | $0.507^{***}$ | $0.511^{***}$ | $0.461^{***}$ |
| | (0.033) | (0.042) | (0.042) | (0.045) |
| + Signal | $5.435^{***}$ | $5.315^{***}$ | 2.407 | 0.443 |
| | (1.458) | (1.428) | (2.139) | (2.086) |
| Main Treatment | 3.540 | $5.267^{**}$ | 2.293 | 1.969 |
| | (2.320) | (2.289) | (2.852) | (2.730) |
| + Signal × Main Treatment | | | $5.807^{**}$ | $6.298^{**}$ |
| | | | (2.819) | (2.718) |
| Female | 2.767 | 2.143 | 2.393 | 2.085 |
| | (2.271) | (2.188) | (2.174) | (2.133) |
| Age | −0.387 | $-0.414^{*}$ | $-0.415^{*}$ | −0.386 |
| | (0.237) | (0.242) | (0.241) | (0.236) |
| # Attempted by Teammate 1 | | $3.252^{***}$ | $3.235^{***}$ | $1.944^{***}$ |
| | | (0.637) | (0.637) | (0.684) |
| # Attempted by Teammate 2 | | $-1.401^{**}$ | $-1.403^{**}$ | $-1.847^{***}$ |
| | | (0.561) | (0.558) | (0.530) |
| Score of Teammate 1 on IQ Test | | | | $0.645^{***}$ |
| | | | | (0.162) |
| Round Fixed Effects | ✓ | ✓ | ✓ | ✓ |
| $R^2$ | 0.38 | 0.40 | 0.41 | 0.42 |
| Observations | 2595 | 2595 | 2595 | 2595 |

Analysis uses OLS regression. Difference is significant from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level.

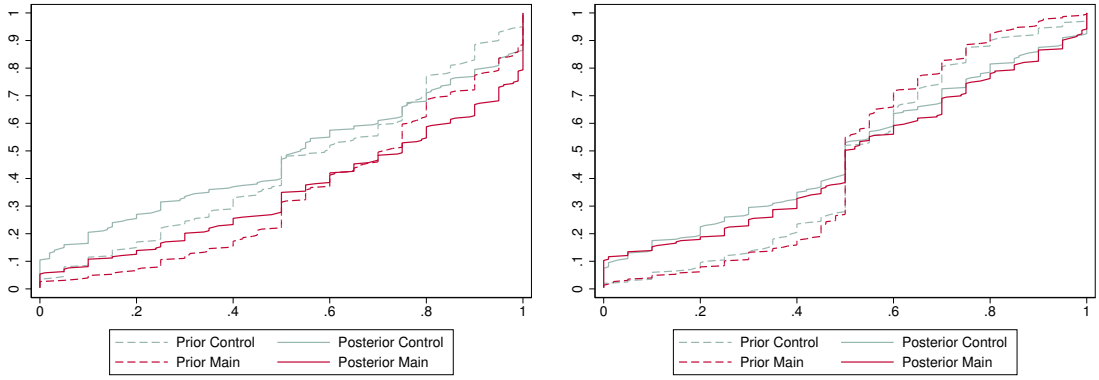
# 5 Examining Prior and Posterior Beliefs

Figure OA.7 presents the CDFs of beliefs in the Primary experiment, for teammate 1 and teammate 2 respectively, separated by prior and posterior, and by treatment. We reject that beliefs are normally distributed for all sub-samples at the 1% level, with the exception of prior beliefs in Control about teammate 2 which we reject at the 10% level (Shapiro-Wilk test).

Regarding teammate 1, relative to the Control treatment, participants in the Main treatment are initially more overconfident (significantly different at 1% level; Kolmogorov-Smirnov test). After four rounds of feedback, the belief distributions diverge even further, and the distributions continue to be significantly different at the 1% level. Regarding teammate 2, there are no prior differences in confidence, and the distributions are not significantly different at conventional levels (Kolmogorov-Smirnov test). After four rounds of feedback, although there is a larger divergence in mean beliefs, with more relative confidence in Main compared to Control, the distributions are not significantly different (Kolmogorov-Smirnov test).

Figure OA.8 presents the analogous CDFs for the Follow-up experiment. We also reject that beliefs are normally distributed for all sub-samples at the 1% level, with the exception of posterior beliefs about the random fundamental (Shapiro-Wilk test).

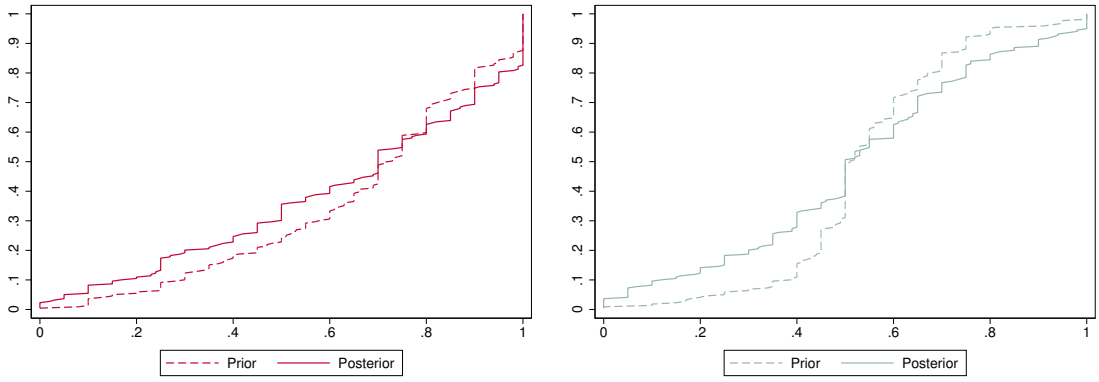
Figure OA.7: Prior and Posterior Belief Distributions




(a) Teammate 1 (b) Teammate 2

Cumulative distribution functions. Note that the posterior is calculated after receiving four signals.

Figure OA.8: Prior and Posterior Belief Distributions (Follow-up)




(a) Teammate 1 (b) Random Fundamental

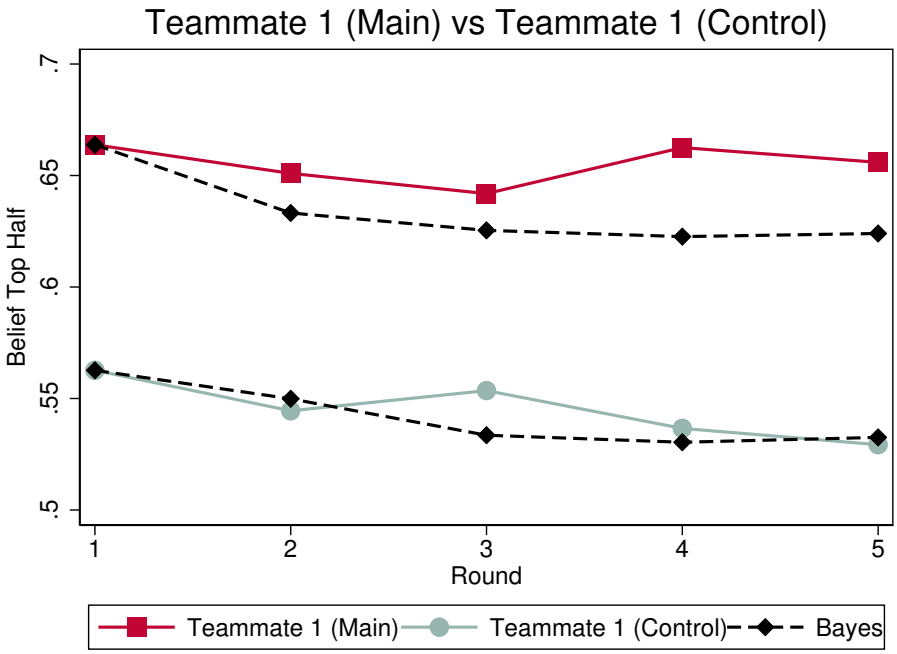
Cumulative distribution functions. Note that the posterior is calculated after receiving four signals.

Figures OA.9 and OA.10 examine the evolution of beliefs in response to feedback for teammate 1 and 2 respectively, starting from the first prior, before receiving any feedback. While posterior beliefs about one's self (Main, teammate 1) are significantly greater than beliefs about teammate 1 in the Control, this is in large part driven by differences in prior beliefs due to overconfidence. In both figures one can see a pattern that posterior beliefs in the final round deviate further from the Bayesian prediction in Main compared to Control, both for teammate 1 and 2.[8]

Figure OA.11 examines this more closely, presenting the difference between reported posteriors and the Bayesian prediction given participants' initial priors, after four rounds of feedback. This corresponds to round 5 in the two figures above. While this does present evidence that positive deviations are more pronounced in the Main treatment, we also note that the difference between the deviations in Main and Control are not significantly different at conventional levels.

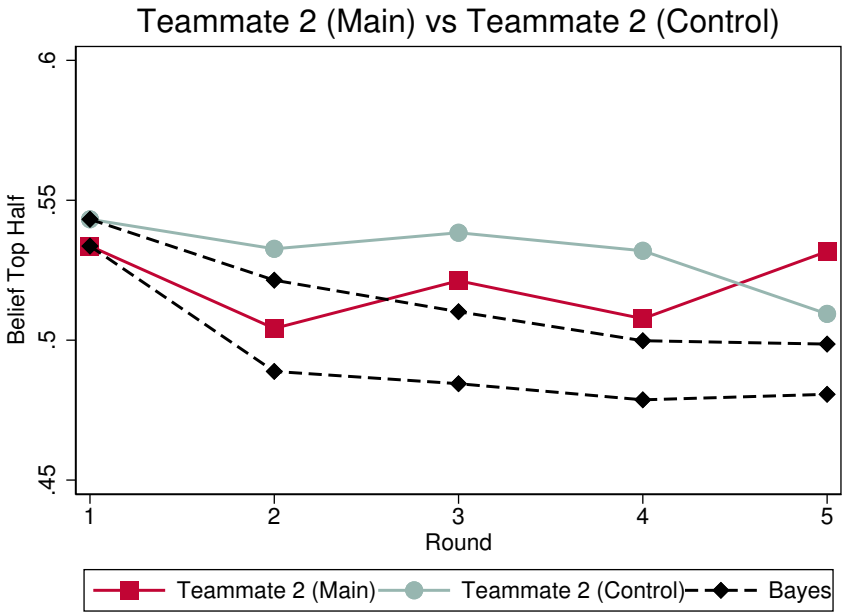
[8]One pattern that can be seen is that Bayesian beliefs appear to flatten out above 50%. In fact, this corresponds to Bayesian updating from overconfident priors, as a result of the signal structure in our experiment. Specifically, the probability of a positive or negative signal depends on which of the four states is true for a given participant: $TT$, $TB$, $BT$, $BB$. For states $TT$ or $BB$ the probability of a positive or negative signal is 90%, respectively. For these states, a Bayesian will eventually learn the true state in the long run. However, for states $TB$ or $BT$, the probability of a positive signal is equal at 50%. Therefore, participants who are members of teams in these states will not learn whether the true state is $TB$ or $BT$ and their beliefs in the long run will converge to an intermediate probability that is a function of their prior. For this reason, the average Bayesian belief will converge at a level which is greater than 50%.

Figure OA.9: Evolution of Beliefs: Teammate 1



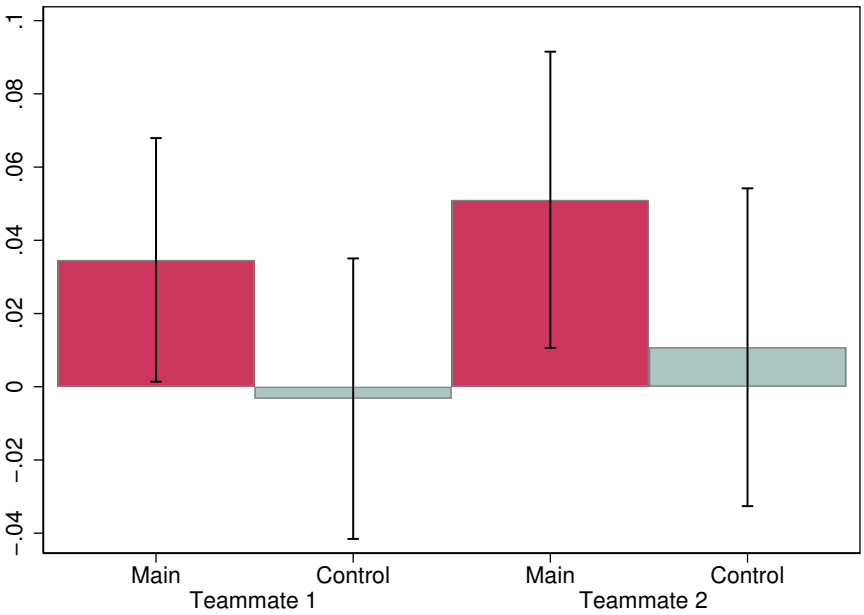

Evolution of average beliefs about teammate 1 starting from prior beliefs with 4 round of feedback. Bayesian benchmark is calculated from participant's first prior, then evolves given actual signals observed. Confidence intervals omitted for clarity (intervals are always overlapping with Bayesian predictions).

Figure OA.10: Evolution of Beliefs: Teammate 2




Evolution of average beliefs about teammate 2 starting from prior beliefs with 4 round of feedback. Bayesian benchmark is calculated from participant's first prior, then evolves given actual signals observed. Confidence intervals omitted for clarity (intervals are always overlapping with Bayesian predictions).

Figure OA.11: Raw Deviation of Posterior Beliefs from Bayesian Benchmark




Plot of the difference between actual posterior beliefs and Bayesian beliefs after four rounds of feedback. Bayesian beliefs are calculated using participant priors before any feedback. 95% confidence intervals shown.

# 6 Belief Updating Alternative Sub-samples

## 6.1 Excluding Part 3 from analysis

At the end of Part 2, participants had the opportunity to re-match with new teammate 2's. In Part 3, those who are not re-matched, continue to update about the same teammate 2, while those who are re-matched state a prior belief about the new teammate 2, and then update about this new teammate. In the statistical analysis of the main paper these situations are pooled together. Here we replicate the main analysis, the analogous tables for Table 2 and 3 in the main paper, excluding Part 3 of the experiment. Table OA.4 presents this analysis for teammate 1, while Table OA.5 presents the analysis for teammate 2. From these tables it is apparent that the results are not significantly affected by the removal of Part 3.

Table OA.4: Updating Beliefs about Teammate 1

| Regressor | (1)<br>Main Treatment | (2)<br>Control Treatment |
|---|---|---|
| $\delta$ | $0.702^{***}$ | $0.709^{***}$ |
| | $(0.060)$ | $(0.053)$ |
| $\beta_1$ | $0.671^{***}$ | $0.581^{***}$ |
| | $(0.074)$ | $(0.095)$ |
| $\beta_0$ | $0.325^{***}$ | $0.609^{***}$ |
| | $(0.066)$ | $(0.082)$ |
| P-Value ( $\delta = 1$ ) | 0.0000 | 0.0000 |
| P-Value ( $\beta_1 = 1$ ) | 0.0000 | 0.0000 |
| P-Value ( $\beta_0 = 1$ ) | 0.0000 | 0.0000 |
| P-Value ( $\beta_1 = \beta_0$ ) | 0.0018 | 0.8032 |
| $R^2$ | 0.57 | 0.55 |
| Observations | 569 | 525 |
| P-Value [Chow-test] for $\delta$ | (1) vs (2): | 0.9325 |
| P-Value [Chow-test] for $\beta_1$ | (1) vs (2): | 0.4524 |
| P-Value [Chow-test] for $\beta_0$ | (1) vs (2): | 0.0067 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | (1) vs (2): | 0.0163 |

Analysis uses OLS regression, for only Part 2 of the experiment. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

Table OA.5: Updating Beliefs about Teammate 2

| Regressor | (1) Main Treatment | (2) Control Treatment |
|---|---|---|
| $\delta$ | $0.690^{***}$ | $0.664^{***}$ |
| | $(0.062)$ | $(0.062)$ |
| $\beta_1$ | $0.405^{***}$ | $0.523^{***}$ |
| | $(0.064)$ | $(0.080)$ |
| $\beta_0$ | $0.238^{***}$ | $0.419^{***}$ |
| | $(0.058)$ | $(0.076)$ |
| P-Value ( $\delta = 1$ ) | 0.0000 | 0.0000 |
| P-Value ( $\beta_1 = 1$ ) | 0.0000 | 0.0000 |
| P-Value ( $\beta_0 = 1$ ) | 0.0000 | 0.0000 |
| P-Value ( $\beta_1 = \beta_0$ ) | 0.0597 | 0.2914 |
| $R^2$ | 0.41 | 0.43 |
| Observations | 672 | 582 |
| P-Value [Chow-test] for $\delta$ | (1) vs (2): | 0.7704 |
| P-Value [Chow-test] for $\beta_1$ | (1) vs (2): | 0.2508 |
| P-Value [Chow-test] for $\beta_0$ | (1) vs (2): | 0.0566 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | (1) vs (2): | 0.6250 |

Analysis uses OLS regression, for only Part 2 of the experiment. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

## 6.2 Sub-sample robustness checks

Here we will change some of the inclusion criteria used in Tables 2 and 3 in the main paper. In theses tables we had excluded wrong updates as well as boundary observations (beliefs of 0 or 1). An update is defined as being in the wrong direction when individuals update at least one of the two beliefs per feedback round, for either teammate 1 or 2, in the opposite direction than feedback suggests, *without* adjusting the other belief in the direction feedback suggests. Another restriction we use in the main paper is to exclude boundary observations of beliefs equal to either 0 or 1, since these drop out of the logit function.

Table OA.6 presents the implications of these sampling restrictions for analysing reactions to feedback of teammate 1. In Columns 1 and 2 of Table OA.6 we include observations which involve updating in the wrong direction following our earlier definition.

One can see that while the magnitudes are affected, the patterns are virtually unchanged. If anything there is slightly more evidence that positive signals receive more weight in Main (beliefs about own performance) relative to Control (beliefs about another's performance). In Columns 3 and 4 we include boundary observations, noting that this as well does not significantly alter the main pattern of results, though it does create some asymmetry in Control.[9] Finally in Columns 5 and 6 we remove all restrictions to the data. The patterns are similar, though again with some asymmetry in Control. We note that the overall results appear to be even stronger than our main analysis suggests, pointing to a stronger reaction to positive signals ($\beta_1$) in Main, a weaker reaction to negative signals ($\beta_0$), and a rejection that the size of the asymmetry is equal between Main and Control.

Similarly, in Table OA.7 we investigate the implications of removing these sampling restrictions on updating about teammate 2. We see similar patterns, with some stronger evidence that positive signals are incorporated less in Main (participant is part of the team) rather than Control, compared to Table 3 in the paper. Overall these tables suggest that our patterns of results are consistent for other sampling strategies.

[9]We do not have an explanation for this asymmetry, but we note that including boundary observations essentially amounts to adding noise to the model, since once beliefs arrive at the extremes, they typically stay there. Regardless, the asymmetry in Main is greater in magnitude than Control, though the p-value is 0.115.

Table OA.6: Updating Beliefs about Teammate 1

| | Inc. Wrong Dir. | | Inc. Boundary | | Inc. All | |
|---|---|---|---|---|---|---|
| Regressor | Main (1) | Control (2) | Main (3) | Control (4) | Main (5) | Control (6) |
| $\delta$ | $0.703^{***}$ | $0.710^{***}$ | $0.826^{***}$ | $0.842^{***}$ | $0.790^{***}$ | $0.805^{***}$ |
| | $(0.056)$ | $(0.048)$ | $(0.048)$ | $(0.044)$ | $(0.052)$ | $(0.046)$ |
| $\beta_1$ | $0.513^{***}$ | $0.365^{***}$ | $0.743^{***}$ | $0.646^{***}$ | $0.658^{***}$ | $0.467^{***}$ |
| | $(0.072)$ | $(0.070)$ | $(0.084)$ | $(0.084)$ | $(0.083)$ | $(0.080)$ |
| $\beta_0$ | $0.071^{***}$ | $0.339^{***}$ | $0.281^{***}$ | $0.445^{***}$ | $0.054^{***}$ | $0.258^{***}$ |
| | $(0.058)$ | $(0.053)$ | $(0.075)$ | $(0.078)$ | $(0.069)$ | $(0.067)$ |
| P-Value ( $\delta = 1$ ) | $0.0000$ | $0.0000$ | $0.0004$ | $0.0005$ | $0.0001$ | $0.0000$ |
| P-Value ( $\beta_1 = 1$ ) | $0.0000$ | $0.0000$ | $0.0024$ | $0.0000$ | $0.0001$ | $0.0000$ |
| P-Value ( $\beta_0 = 1$ ) | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_1 = \beta_0$ ) | $0.0001$ | $0.7577$ | $0.0003$ | $0.0653$ | $0.0000$ | $0.0416$ |
| $R^2$ | 0.55 | 0.54 | 0.59 | 0.57 | 0.54 | 0.52 |
| Observations | 996 | 985 | 938 | 899 | 1077 | 1062 |
| P-Value [Chow-test] for $\delta$ | (1) vs (2): | 0.9210 | (3) vs (4): | 0.8037 | (5) vs (6): | 0.8263 |
| P-Value [Chow-test] for $\beta_1$ | (1) vs (2): | 0.1408 | (3) vs (4): | 0.4127 | (5) vs (6): | 0.0959 |
| P-Value [Chow-test] for $\beta_0$ | (1) vs (2): | 0.0006 | (3) vs (4): | 0.1297 | (5) vs (6): | 0.0342 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | (1) vs (2): | 0.0026 | (3) vs (4): | 0.1153 | (5) vs (6): | 0.0141 |

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

Table OA.7: Updating Beliefs about Teammate 2

| | Inc. Wrong Dir. | | Inc. Boundary | | Inc. All | |
|---|---|---|---|---|---|---|
| Regressor | Main (1) | Control (2) | Main (3) | Control (4) | Main (5) | Control (6) |
| $\delta$ | $0.741^{***}$ | $0.687^{***}$ | $0.868^{***}$ | $0.761^{***}$ | $0.845^{***}$ | $0.733^{***}$ |
| | $(0.044)$ | $(0.047)$ | $(0.044)$ | $(0.054)$ | $(0.040)$ | $(0.051)$ |
| $\beta_1$ | $0.330^{***}$ | $0.342^{***}$ | $0.435^{***}$ | $0.666^{***}$ | $0.330^{***}$ | $0.497^{***}$ |
| | $(0.054)$ | $(0.066)$ | $(0.073)$ | $(0.080)$ | $(0.069)$ | $(0.075)$ |
| $\beta_0$ | $0.133^{***}$ | $0.275^{***}$ | $0.389^{***}$ | $0.549^{***}$ | $0.239^{***}$ | $0.379^{***}$ |
| | $(0.038)$ | $(0.056)$ | $(0.066)$ | $(0.071)$ | $(0.059)$ | $(0.065)$ |
| P-Value ( $\delta = 1$ ) | $0.0000$ | $0.0000$ | $0.0030$ | $0.0000$ | $0.0002$ | $0.0000$ |
| P-Value ( $\beta_1 = 1$ ) | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_0 = 1$ ) | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_1 = \beta_0$ ) | $0.0042$ | $0.4056$ | $0.6331$ | $0.1979$ | $0.3120$ | $0.1903$ |
| $R^2$ | 0.52 | 0.46 | 0.48 | 0.46 | 0.46 | 0.42 |
| Observations | 1166 | 1088 | 1085 | 974 | 1244 | 1150 |
| P-Value [Chow-test] for $\delta$ | (1) vs (2): | 0.4052 | (3) vs (4): | 0.1231 | (5) vs (6): | 0.0838 |
| P-Value [Chow-test] for $\beta_1$ | (1) vs (2): | 0.8942 | (3) vs (4): | 0.0316 | (5) vs (6): | 0.0995 |
| P-Value [Chow-test] for $\beta_0$ | (1) vs (2): | 0.0367 | (3) vs (4): | 0.0993 | (5) vs (6): | 0.1103 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | (1) vs (2): | 0.2163 | (3) vs (4): | 0.5880 | (5) vs (6): | 0.8273 |

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

## 6.3 Updating By Teammate 1 Performance

In this subsection we examine updating by whether teammate 1's performance was in the top half or bottom half, in Tables OA.8 and OA.9. These results can be compared with those in Table 2 of the main paper. However, in addition to the standard disclaimer that this analysis should be interpreted cautiously due to the smaller sample size, another issue arises in this analysis. In particular, when comparing Main versus Control within the top and bottom half sub-samples, the coefficient estimate of the $\delta$ parameter is significantly different. This parameter is the weight the updating model places on the log prior odds ratio. Since the parameter estimates impact one-another, this may hinder the ability to make direct comparisons between $\beta_s$ across Main and Control.[10] To account for this, in addition to the standard unrestricted estimation in Columns (1) and (2), in this subsection we also include another estimation which fixes $\delta$ at the average value across Columns (1) and (2). This estimation in Columns (3) and (4) allows us to more confidently compare estimates of $\beta_0$ and $\beta_1$ across Main and Control.

Examining Columns (3) and (4) in Table OA.8 shows positive asymmetry in Main, significant at the 1% level (p-value 0.000), and positive but not significant asymmetry in Control (p-value 0.185). Comparing Main and Control directly, the asymmetry in Main is significantly greater at the 10% level (Chow-test p-value 0.074). $\beta_1$ and $\beta_0$ are not significantly different between Main and Control, although the Chow-test p-value for $\beta_0$ is 0.139.[11]

Next, examining Columns (3) and (4) in Table OA.9 interestingly shows no asymmetry in Main, and negative asymmetry in Control, significant at the 5% level (p-value 0.016). Comparing Main and Control directly, the difference is significant at the 5% level (Chow-test p-value 0.037). In these specifications, $\beta_0$ is significantly different between Main and Control (Chow-test p-value 0.000), but not $\beta_1$.[12]

Overall, the difference in asymmetry across Main and Control is significant for both the top half and bottom half sub-samples. In the former case it appears to be driven by positive asymmetry in Main, while in the latter case it appears to be driven by negative asymmetry in Control.

[10] Differences in $\beta_s$ across Main and Control may arise due to interactions with $\delta$, in addition to direct differences in weighting positive or negative signals.

[11] In the unrestricted estimation, both Columns (1) and (2) show significant positive asymmetry. However, given the significant differences in $\delta$ across these columns, in line with our earlier discussion, we do not see a valid way to interpret these comparisons of asymmetry.

[12] In the unrestricted estimation, both Columns (1) and (2) show similar patterns. Once again, given the differences in $\delta$ across these columns, we do not see a valid way to interpret these comparisons of asymmetry.

Table OA.8: Updating Beliefs about Teammate 1 – Top Half Only

| Regressor | (1) Main | (2) Control | (3) Main | (4) Control |
|---|---|---|---|---|
| $\delta$ | $0.933^{**}$ | $0.694^{***}$ | 0.813 | 0.813 |
| | $(0.029)$ | $(0.081)$ | — | — |
| $\beta_1$ | $0.479^{***}$ | $0.605^{***}$ | $0.615^{***}$ | $0.503^{***}$ |
| | $(0.074)$ | $(0.106)$ | $(0.069)$ | $(0.078)$ |
| $\beta_0$ | $0.250^{***}$ | $0.214^{***}$ | $0.144^{***}$ | $0.343^{***}$ |
| | $(0.087)$ | $(0.114)$ | $(0.096)$ | $(0.095)$ |
| P-Value ( $\delta = 1$ ) | $0.0214$ | $0.0003$ | — | — |
| P-Value ( $\beta_1 = 1$ ) | $0.0000$ | $0.0004$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_0 = 1$ ) | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_1 = \beta_0$ ) | $0.0606$ | $0.0338$ | $0.0004$ | $0.1851$ |
| $R^2$ | 0.86 | 0.60 | 0.25 | 0.13 |
| Observations | 375 | 416 | 375 | 416 |
| P-Value [Chow-test] for $\delta$ | | 0.0054 | | — |
| P-Value [Chow-test] for $\beta_1$ | | 0.3269 | | 0.2845 |
| P-Value [Chow-test] for $\beta_0$ | | 0.8040 | | 0.1389 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | | 0.4549 | | 0.0739 |

Analysis uses OLS regression restricting sample to top half performers. Columns (1) and (2) estimate $\delta, \beta_1$, and $\beta_0$, while columns (3) and (4) fix $\delta = 0.813$ and estimate $\beta_1$ and $\beta_0$. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

Table OA.9: Updating Beliefs about Teammate 1 – Bottom Half Only

| Regressor | (1) Main | (2) Control | (3) Main | (4) Control |
|---|---|---|---|---|
| $\delta$ | $0.490^{***}$ | $0.736^{***}$ | 0.613 | 0.613 |
| | $(0.103)$ | $(0.058)$ | — | — |
| $\beta_1$ | $0.347^{***}$ | $0.349^{***}$ | $0.340^{***}$ | $0.252^{***}$ |
| | $(0.128)$ | $(0.132)$ | $(0.111)$ | $(0.158)$ |
| $\beta_0$ | $0.283^{***}$ | $0.604^{***}$ | $0.300^{***}$ | $0.662^{***}$ |
| | $(0.064)$ | $(0.075)$ | $(0.062)$ | $(0.079)$ |
| P-Value ( $\delta = 1$ ) | 0.0000 | 0.0000 | — | — |
| P-Value ( $\beta_1 = 1$ ) | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| P-Value ( $\beta_0 = 1$ ) | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| P-Value ( $\beta_1 = \beta_0$ ) | 0.6819 | 0.0382 | 0.7724 | 0.0161 |
| $R^2$ | 0.27 | 0.61 | 0.06 | 0.19 |
| Observations | 488 | 413 | 488 | 413 |
| P-Value [Chow-test] for $\delta$ | (1) vs (2): | 0.0361 | (3) vs (4): | — |
| P-Value [Chow-test] for $\beta_1$ | (1) vs (2): | 0.9950 | (3) vs (4): | 0.6468 |
| P-Value [Chow-test] for $\beta_0$ | (1) vs (2): | 0.0011 | (3) vs (4): | 0.0003 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | (1) vs (2): | 0.1056 | (3) vs (4): | 0.0366 |

Analysis uses OLS regression restricting sample to bottom half performers. Columns (1) and (2) estimate $\delta, \beta_1$, and $\beta_0$, while columns (3) and (4) fix $\delta = 0.613$ and estimate $\beta_1$ and $\beta_0$. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

## 6.4 Updating by Hard/Easy Test

We conclude this section by examining updating by the hard-easy test sub-samples. To our knowledge we believe we are the first to study whether the belief updating process differs depending on whether the task was hard or easy. That is, we answer the question of whether we can observe the hard-easy effect in updating. Table OA.10 presents the analogue to Table 2 in the paper, but split into hard and easy sub-samples for Teammate 1. As expected the estimates are noisier, due to the sample size being cut in half. However, one can see that the asymmetry observed for individuals regarding their own performance is significant for both hard and easy test versions, at the 5% and 10% level respectively. It is also worth noting that asymmetry in the hard version is greater, and significantly different from the Control. For the easy test, some asymmetry is also observed in the Control, and while not significant, means we are unable to rule out that the asymmetry in Main is statistically different from Control for the easy test.[13] Overall this presents only suggestive evidence that positive asymmetry may be stronger for more difficult tasks, an interesting question for future research.

Table OA.11 presents these results for Teammate 2, i.e. the analogue to Table 3 in the paper. The patterns for the hard version of the test mimic those found in the overall analysis, whereas for the easy test, there appear to be contrary patterns, which we cannot explain, other to note again the smaller sample sizes.

[13]The lower sample sizes suggest some caution in interpreting patterns seen here. Moreover, the difference in the asymmetry across the Main treatments (own performance) for easy versus hard test versions is not significant.

Table OA.10: Updating Beliefs about Teammate 1 (By Test Difficulty)

| | Easy | | Hard | |
|---|---|---|---|---|
| Regressor | Main (1) | Control (2) | Main (3) | Control (4) |
| $\delta$ | $0.790^{***}$ | $0.736^{***}$ | $0.673^{***}$ | $0.758^{***}$ |
| | $(0.055)$ | $(0.062)$ | $(0.091)$ | $(0.060)$ |
| $\beta_1$ | $0.538^{***}$ | $0.559^{***}$ | $0.608^{***}$ | $0.466^{***}$ |
| | $(0.086)$ | $(0.110)$ | $(0.115)$ | $(0.101)$ |
| $\beta_0$ | $0.303^{***}$ | $0.430^{***}$ | $0.244^{***}$ | $0.586^{***}$ |
| | $(0.088)$ | $(0.087)$ | $(0.075)$ | $(0.086)$ |
| P-Value ( $\delta = 1$ ) | $0.0003$ | $0.0001$ | $0.0005$ | $0.0001$ |
| P-Value ( $\beta_1 = 1$ ) | $0.0000$ | $0.0001$ | $0.0010$ | $0.0000$ |
| P-Value ( $\beta_0 = 1$ ) | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_1 = \beta_0$ ) | $0.0757$ | $0.2990$ | $0.0207$ | $0.3439$ |
| $R^2$ | 0.66 | 0.59 | 0.52 | 0.61 |
| Observations | 412 | 392 | 451 | 437 |
| P-Value [Chow-test] for $\delta$ | (1) vs (2): | 0.5121 | (3) vs (4): | 0.4341 |
| P-Value [Chow-test] for $\beta_1$ | (1) vs (2): | 0.8802 | (3) vs (4): | 0.3532 |
| P-Value [Chow-test] for $\beta_0$ | (1) vs (2): | 0.2989 | (3) vs (4): | 0.0028 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | (1) vs (2): | 0.5502 | (3) vs (4): | 0.0148 |

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

Table OA.11: Updating Beliefs about Teammate 2 (By Test Difficulty)

| | Easy | | Hard | |
|---|---|---|---|---|
| Regressor | Main (1) | Control (2) | Main (3) | Control (4) |
| $\delta$ | $0.740^{***}$ | $0.705^{***}$ | $0.776^{***}$ | $0.732^{***}$ |
| | $(0.073)$ | $(0.068)$ | $(0.061)$ | $(0.070)$ |
| $\beta_1$ | $0.286^{***}$ | $0.617^{***}$ | $0.492^{***}$ | $0.394^{***}$ |
| | $(0.074)$ | $(0.104)$ | $(0.087)$ | $(0.085)$ |
| $\beta_0$ | $0.371^{***}$ | $0.398^{***}$ | $0.110^{***}$ | $0.435^{***}$ |
| | $(0.059)$ | $(0.088)$ | $(0.061)$ | $(0.086)$ |
| P-Value ( $\delta = 1$ ) | $0.0006$ | $0.0000$ | $0.0004$ | $0.0002$ |
| P-Value ( $\beta_1 = 1$ ) | $0.0000$ | $0.0004$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_0 = 1$ ) | $0.0000$ | $0.0000$ | $0.0000$ | $0.0000$ |
| P-Value ( $\beta_1 = \beta_0$ ) | $0.3282$ | $0.0940$ | $0.0009$ | $0.6856$ |
| $R^2$ | 0.48 | 0.48 | 0.59 | 0.53 |
| Observations | 476 | 440 | 540 | 476 |
| P-Value [Chow-test] for $\delta$ | (1) vs (2): | 0.7259 | (3) vs (4): | 0.6337 |
| P-Value [Chow-test] for $\beta_1$ | (1) vs (2): | 0.0094 | (3) vs (4): | 0.4229 |
| P-Value [Chow-test] for $\beta_0$ | (1) vs (2): | 0.8034 | (3) vs (4): | 0.0019 |
| P-Value [Chow-test] for $(\beta_1 - \beta_0)$ | (1) vs (2): | 0.0498 | (3) vs (4): | 0.0047 |

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. $\delta$ is the coefficient on the log prior odds ratio. $\beta_1$ and $\beta_0$ are coefficients on the log likelihood of observing positive and negative signals respectively. Constant omitted because of collinearity. Bayesian updating corresponds to $\delta = \beta_1 = \beta_0 = 1$. $\beta_1, \beta_0 < 1$ indicates conservative updating. $\beta_1 - \beta_0 > 0$ indicates positive asymmetric updating.

# 7 Discussion of Alternative Explanations

## 7.1 Anchoring

One potential alternative explanation for our results is that there is some form of anchoring bias which causes individuals to update in a self-serving way about themselves, and then simply update in a similar way about their teammate; in a purely mechanical (and naive) sense.[14] With the results of our Follow-up experiment we are able to rule out such an explanation. Specifically, if the reason for observing a positive bias for teammate 2 was simply due to anchoring about ones' self, then the same patterns should have been observed in the Follow-up. Since this was not the case, and we observed no bias, we can conclude that the presence of the human teammate themselves was responsible for the bias observed (and they were not just a passive factor that absorbed positive bias due to anchoring). Note that if anchoring were to occur *only* in cases when the teammate is human, we would categorise this as being an in-group updating bias. More details can be found both in the explanation below and our concluding discussion in the paper.

## 7.2 Selective Discounting or Ignorance

Here we consider a possible alternative explanation which involves selectively discounting or ignoring particular signals. First we briefly consider a hypothetical individual who discounts the true signal strength of negative feedback overall, independently of whether they are updating about themselves (teammate 1) or their teammate (teammate 2). We note that such a model would predict *equivalent* asymmetry across self and teammate. However, examining the first columns in Tables 2 and 3 in the main paper, we note that we rejected this prediction, at the 10% level.

Next we turn to selective ignorance of negative signals. To set up this explanation, we consider a different variation of our theoretical model where the sub-conscious process can selectively choose whether to ignore new information. That is, the sub-conscious process chooses the optimal action of whether or not to update, conditional on observing the signal. In this model the action is binary, update or not. We assume otherwise updating would follows Bayes' rule.

We note that it would always be optimal to update after a positive signal, since the

[14] We note that beliefs about oneself and teammate 2 are not strongly correlated ($\rho = 0.058$). For the Control treatment this correlation is ($\rho = -0.003$). It is true that when we compare the aggregate difference between posteriors (in the final round) and priors (in the first round), there is a significant positive correlation, though this is not significantly different between the Main ($\rho = 0.298$) and Control ($\rho = 0.253$) treatments.

benefits from overconfidence and the benefits from accuracy go in the same direction. It may or may not be optimal to update after a negative signal, since the costs to holding lower beliefs are in opposition to the benefits of accuracy. Finally, we note that adding a cognitive fixed cost of updating, would potentially lower the number of updates, but not in a biased way between positive versus negative signals.

To test the predictions of such a model we investigate whether the propensity to update differed across our Main and Control treatments. In fact in the Control treatment the probability of updating after any signal is 67%, while in Main it is 63%. The difference is not significant when accounting for potential correlation of standard errors at the individual level. Moreover, this slight difference is driven entirely by positive signals, not negative signals. Thus there is no evidence that participants ignore negative feedback more in the Main treatment. Beyond this, in fact participants are more likely to update after a negative signal than a positive signal (in both Main and Control), which further goes against this potential theoretical explanation.

### 7.3 In-Group Bias

We also consider the possibility that individuals may become overconfident or optimistic about team performance, rather than individual performance. First we note that we don't find any bias when individuals update about a team in our Control treatment, so such a theory would nonetheless require that individuals themselves be a part of the team.

Next we don't find any initial overestimation of teammate 2's performance in Main compared to teammate 2's performance in Control – the belief that teammate 2 scores in the top 50% are 53.4% and 54.3% (Wilcoxon ranksum test p-value: 0.572) in the Main and Control treatment, respectively. Thus, the in-group bias would need to selectively apply to belief updating from signals, but not appear in initial belief formation. For these reasons we find that a classical type of in-group bias is unlikely to explain our results. However, if in-group bias specifically affects updating only (and not prior belief formation) we would no longer be able to exclude it. We discuss this at the end of Section 4.2.1 in the main paper, as well as in the concluding discussion.

## 8 Alternative Models

Here we present two alternative models of self-serving attribution bias which focus on two relatively stark predictions involving self-serving beliefs which are formed due to mis-attribution towards either (i) noise, or (ii) the external fundamental (teammate 2).

These models formed an earlier basis for our initial hypotheses, but lacked the micro-foundations that we later embedded into the model of the current paper.

We denote the first alternative model by: noisy attribution bias (NAB). With NAB the individual processes information about other factors (teammate 2) accurately, but is positively biased about own performance (teammate 1) at the expense of noise. In the second model, fundamental attribution bias (FAB), the individual respects the amount of noise contained in the signal, but is biased about their own performance (teammate 1) at the expense of teammate 2.

## 8.1 Updating with Noisy Attribution Bias

With NAB, individuals over-attribute positive feedback to their own performance, and under-attribute negative feedback to bad luck. We additionally specify that NAB predicts that individuals update using Bayes' rule regarding the performance of this teammate. Someone who exhibits NAB will update in a way that is consistent with mis-interpreting the strength of the binary signal. That is, when they receive a positive signal, they believe it is more informative about their performance than it really is. When they receive a negative signal, they believe it is less informative about their performance than it really is. They over-interpret the strength of the signal by a factor of $\gamma_p$, where $\gamma_p \geq 1$ in the case of a positive signal, and $\gamma_n \geq 1$ in the case of a negative signal. However, updating about teammate 2's performance occurs as if the strength of the signal were correctly interpreted in all states.

Thus, regarding own performance, and following the notation in the main paper, biased updating in response to positive and negative feedback through NAB results in upward biased beliefs:

$$\left[b_{t+1}^{1,NAB}|s_t = p\right] = \frac{\gamma_p \left[\Phi_{TT} b_t^{TT} + \Phi_{TB} b_t^{TB}\right]}{\gamma_p \left[\Phi_{TT} b_t^{TT} + \Phi_{TB} b_t^{TB}\right] + \Phi_{BT} b_t^{BT} + \Phi_{BB} b_t^{BB}} \geq \left[b_{t+1}^{1,BAYES}|s_t = p\right], \tag{1}$$

$$\begin{aligned}\left[b_{t+1}^{1,NAB}|s_t = n\right] &= \frac{\gamma_n \left[(1-\Phi_{TT}) b_t^{TT} + (1-\Phi_{TB}) b_t^{TB}\right]}{\gamma_n \left[(1-\Phi_{TT}) b_t^{TT} + (1-\Phi_{TB}) b_t^{TB}\right] + (1-\Phi_{BT}) b_t^{BT} + (1-\Phi_{BB}) b_t^{BB}} \qquad (2)\\ &\geq \left[b_{t+1}^{1,BAYES}|s_t = n\right].\end{aligned}$$

With NAB, updating about their own performance exhibits positive asymmetry - in-

dividuals over-weight positive signals and under-weight negative signals relative to a Bayesian. As noted, they are Bayesian with regards to their teammate's performance.

## 8.2 Updating with Fundamental Attribution Bias

With FAB, individuals over-attribute positive feedback to their own performance, at the expense of the other source of uncertainty, i.e. their teammate. Similarly, they under-attribute negative feedback to themselves, and over-attribute it to their teammate.

FAB takes the same functional form as NAB with regards to own performance.

$$\left[b_{t+1}^{1,FAB}|s_t=p\right]=\left[b_{t+1}^{1,NAB}|s_t=p\right]= \frac{\gamma_p\left[\Phi_{TT}b_t^{TT}+\Phi_{TB}b_t^{TB}\right]}{\gamma_p\left[\Phi_{TT}b_t^{TT}+\Phi_{TB}b_t^{TB}\right]+\Phi_{BT}b_t^{BT}+\Phi_{BB}b_t^{BB}}\geq\left[b_{t+1}^{1,BAYES}|s_t=p\right], \tag{3}$$

$$\left[b_{t+1}^{1,FAB}|s_t=n\right]=\left[b_{t+1}^{1,NAB}|s_t=n\right]= \frac{\gamma_n\left[(1-\Phi_{TT})b_t^{TT}+(1-\Phi_{TB})b_t^{TB}\right]}{\gamma_n\left[(1-\Phi_{TT})b_t^{TT}+(1-\Phi_{TB})b_t^{TB}\right]+(1-\Phi_{BT})b_t^{BT}+(1-\Phi_{BB})b_t^{BB}} \geq\left[b_{t+1}^{1,BAYES}|s_t=n\right]. \tag{4}$$

With FAB the individual updates in a biased but consistent manner across themselves and their teammate, unlike with NAB, where they update as a standard Bayesian with respect to their teammate. In response to positive and negative signals respectively FAB implies for teammate 2:

$$\left[\hat{b}_{t+1}^{2}|s_t=p\right]=\frac{\gamma_p\Phi_{TT}b_t^{TT}+\Phi_{BT}b_t^{BT}}{\gamma_p\left[\Phi_{TT}b_t^{TT}+\Phi_{TB}b_t^{TB}\right]+\Phi_{BT}b_t^{BT}+\Phi_{BB}b_t^{BB}} \tag{5}$$

$$\left[\hat{b}_{t+1}^{2}|s_t=n\right]=\frac{\gamma_n(1-\Phi_{TT})b_t^{TT}+(1-\Phi_{BT})b_t^{BT}}{\gamma_n\left[(1-\Phi_{TT})b_t^{TT}+(1-\Phi_{TB})b_t^{TB}\right]+(1-\Phi_{BT})b_t^{BT}+(1-\Phi_{BB})b_t^{BB}} \tag{6}$$

Because $\Theta=\Phi_{TT}\Phi_{BB}-\Phi_{TB}\Phi_{BT}\leq 0$, this guarantees that $\left[b_{t+1}^{2,FAB}|s_t=s\right]\leq\left[b_{t+1}^{2,BAYES}|s_t=s\right]$. Thus, FAB implies that when considering their teammate's performance they update asymmetrically in the negative direction, i.e. they under (over)-weight positive (negative) signals.

We note that both NAB and FAB predict positive asymmetry with respect to own performance beliefs, but either (i) no asymmetry or (ii) negative asymmetry, respectively, regarding updating about the performance of teammate 2. As our main results in the paper show, neither of these two theoretical predictions are borne out in the data.

# 9 Validation Experiment

## 9.1 Experiment Design

To explore the hypothesis that individuals find it less cognitively costly to distort their beliefs about human sources of uncertainty, rather than non-human, we conducted an online Validation experiment using Prolific.co with 600 participants.[15] We selected participants who were students based in either the UK or Germany, we balanced the sample on gender, and participants had to pass comprehension checks to be included in the study. Full instructions are shown in Section 10.3.

The key between-subject treatment was whether a participant was matched to a person who previously participated in our Primary experiment (treatment Human; $N = 301$) or with a random fundamental (treatment RF; $N = 299$). Participants were introduced to a "matched person" or "matched random factor". In the former case, they were shown a screenshot the IQ-style test from the Primary experiment, and told that their matched person was ranked either in the Top 10 or the Bottom 10, out of 20 performers. In the latter case, they were shown an image of a spin-wheel with numbers from 1 to 20, and told that their random factor would be determined by a randomly generated number that would be either HIGH or LOW, based on whether the generated number was in the Top 10 or Bottom 10.

After being introduced to the source of uncertainty, participants were asked to report their initial beliefs about the probability that the Human/RF was Top 10/High. We incentivised this initial belief elicitation using the binarised scoring rule (BSR) with the potential to earn a fixed payment of £0.50. Next, we asked participants to report a revised belief. We did not provide any new information about the Human/RF. However, we modified the BSR in order to incentivise belief revision.

This unique incentive scheme worked as follows. Participants were truthfully told that in addition to the standard £0.50 BSR payment, they would receive fixed payments of

[15] We pre-registered this experiment on AsPredicted.org, with $N = 600$ participants, where we stated the hypothesis that: "We hypothesise that individuals find it easier to distort beliefs concerning a human than a neutrally termed 'random factor'."

£0.01 for every 5 percentage points they revised their belief, away from the initial belief. These fixed payments accrued independent of whether beliefs were revised upwards or downwards. Participants were explicitly told that they faced a trade-off: their chances of receiving the £0.50 were increasing in their accuracy, but that the more they revised away from their initial belief the more guaranteed fixed payments they would earn.[16]

## 9.2 Theory

We assume that individuals are risk neutral and hold (true) initial belief about the probability that the Human/RF was Top 10/High, $b \in [0,1]$.[17] We call this initial belief true because given the incentives, participants will truthfully report their belief in the initial elicitation. We then denote the revised belief by $\hat{b}$. Following our theoretical framework, cognitive costs which are source ($s$) dependent are denoted by $J_s(\hat{b}, b)$, noting that individuals face mental costs of distorting beliefs $\hat{b}$ away from the truth, $b$. We assume that cognitive costs are symmetric such that $J_s(y) = J_s(-y)$ for any $y$. An individual's objective is then to choose $\hat{b} \in [0,1]$ to maximise the following equation, where $x > 0$ is the BSR payment.[18]

$$b \cdot \left(x \cdot \left(1 - (1-\hat{b})^2\right)\right) + (1-b) \cdot \left(x \cdot \left(1 - \hat{b}^2\right)\right) + \left(\frac{x}{50}\left|\frac{\hat{b}-b}{0.05}\right|\right) - J_s(\hat{b}, b)$$

[16] For example, if a participant reported an initial belief of 50%, they would receive additional fixed payments of £0.03 if they reported 65%, or £0.05 if they reported 25%. The BSR would then work as normal, using the reported belief of either 65% or 25% respectively for these examples, to determine whether they also earned the £0.50 payment.

[17] We assume risk neutrality so that the optimal distortion can be explicitly calculated. Adding risk preferences would result in different levels of optimal distortion, but does not alter the predictions with respect to cognitive costs.

[18] Noting that $b$ is the true subjective belief of the participant, with probability $b$ the participant believes the outcome occurs, and hence the payoff of $x$ under the BSR occurs with probability $(1-(1-\hat{b})^2)$, noting that $\hat{b}$ is the report. Similarly, with probability $(1-b)$ the participant believes the outcome does not occur, and hence the payoff of $x$ under the BSR occurs with probability $(1-\hat{b}^2)$. Noting that participants are paid $\frac{x}{50}$ for each 5 percentage point change, since in the experiment $x = 0.50$, the third term is the fixed payment for revising beliefs. For simplicity, here we assume these additional fixed payments are continuous not discrete.

We next take the first order condition with respect to $\hat{b}$, and setting it equal to zero provides us with two equivalent maxima:

$$\hat{b} = b + \frac{1}{5} - \frac{J_s'(\hat{b}, b)}{2x},$$
$$\hat{b} = b - \frac{1}{5} - \frac{J_s'(\hat{b}, b)}{2x}.$$

We note that given the restriction on $\hat{b} \in [0, 1]$, one of these solutions may not exist. In these cases there will be one unique maximum.

When there are no cognitive costs, it is straightforward to see that the optimal distortion is to revise the initial belief by 20 percentage points either upwards or downwards. With cognitive costs optimal distortions are smaller. Critically, when cognitive costs are state dependent, optimal distortion will differ across states. In particular, if we assume that there are two states $s \in \{H, N\}$ such that it is more cognitively costly to distort beliefs in state $N$, (i.e. $J_N'(x) > J_H'(x)$ for all $x$), then the optimal $\hat{b}$ in state $H$ will lead to greater distortion, i.e. a larger value of $|\hat{b} - b|$, than will the optimal $\hat{b}$ in state $N$.

## 9.3 Results

As noted in the main paper, initial beliefs about being Top/High are similar between the Human and RF treatments, at 55.1% and 55.0% respectively (Wilcoxon ranksum p-value 0.9464). Initial beliefs show an upward bias, and are significantly different from 50% at the 1% level (Wilcoxon signed-rank test p-value 0.0000), mimicking the upward bias we observed for teammate 2 in the Primary and Follow-up experiments. Average absolute belief revisions are 15.3 percentage points, which is lower than the risk-neutral, zero cognitive-cost benchmark of 20. In the Human treatment the value is 16.1, and in the RF treatment it is 14.6, a difference which is significant at the 10% level (Wilcoxon ranksum p-value 0.0653). Through the lens of our theory, this result suggests that individuals face greater cognitive costs when distorting beliefs about a random fundamental rather than a human. We further note that the proportion of individuals who revised their beliefs at all was significantly different between the Human and RF treatments, at 76.1% versus 67.6% ($\chi 2$ test p-value 0.0200).[19]

To try and understand some of the mechanisms underlying this result, we also examined whether the effects were moderated by other variables. To this end, we elicited

[19] We are not aware of any studies which have compared belief distortion of human versus non-human sources of uncertainty. Though in a very distinct context, Awad et al. (2020) find that individuals assign more blame to human than machine sources when both sources are operating an automated car.

three other measures for both the Human and RF. First, we elicited *cognitive uncertainty* regarding how certain individuals were about their initial belief estimate. We follow precisely the methodology of Enke and Graeber (2022). Second, we elicited *warmth*, which directly asked participants how warm or cold they felt towards their matched human or RF (Bayes, 1972). We also asked an exploratory question about how difficult it was to envision uncertain outcomes on a scale from 1 to 10, which we refer to as *imagination*. The exact wording of all of these variables can be seen in the instructions in Section 10.3.[20]

First we examined whether these three measures differed for the human versus the RF. We found statistically significant differences for *cognitive uncertainty* (Wilcoxon ranksum p-value 0.0634) and *warmth* (Wilcoxon ranksum p-value 0.0502), but not for *imagination* (Wilcoxon ranksum p-value 0.8756). Individuals expressed greater cognitive uncertainty with respect to the human, and interestingly, expressed more warmth towards the random fundamental. Regarding the latter finding, this result suggests that feelings of warmth may be unlikely as an alternative explanation for positively biased updating towards the teammate found in our Primary experiments.

Focusing now on *cognitive uncertainty* and *warmth*, Table OA.12 presents Probit regressions (showing marginal effects) for these measures, and their interactions with the treatment. We can see that on its own, more *cognitive uncertainty* is associated with greater belief distortion (Column 4). However, the interaction with the indicator variable for human is not significant (Column 5). There are no significant interactions for *warmth* (nor for *imagination* – not shown).

Overall the results suggest that cognitive uncertainty may be positively related to belief distortion, though it cannot account for all of the direct effects we observe of the source of uncertainty being human, given that it remains significant (Column 4).[21] Given that Enke and Graeber (2022) show how cognitive uncertainty is associated with responses trending towards a cognitive default (here the intuitive default is initial beliefs), this is slightly surprising. However, cognitive uncertainty also may proxy for the richness or ambiguity of a given unknown. In this sense, participants who have more uncertainty about the outcome may also have more flexibility in distorting their beliefs about this outcome. While all of these patterns are exploratory and move beyond the narrow exercise we wished to capture with the Validation experiment, we believe they present on their own a promising avenue for future research into the manipulability of beliefs.

[20] We also included a question about "oneness" but only regarding the human. It is uncorrelated with belief distortions.

[21] We note that it does reduce the coefficient estimate by 22% from 0.085 to 0.066, indicating that cognitive uncertainty does appear to partly capture the result of belief distortion.

Table OA.12: Human versus Random Fundamental

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Human | 0.085** | 0.087** | 0.083 | 0.066* | 0.099** |
| | (0.036) | (0.036) | (0.126) | (0.035) | (0.051) |
| Warmth | | 0.007 | 0.007 | | |
| | | (0.011) | (0.015) | | |
| Human × Warmth | | | 0.001 | | |
| | | | (0.021) | | |
| Cognitive Uncertainty | | | | 0.004*** | 0.005*** |
| | | | | (0.001) | (0.001) |
| Human × Cognitive Uncertainty | | | | | -0.001 |
| | | | | | (0.001) |
| Pseudo $R^2$ | 0.008 | 0.008 | 0.008 | 0.072 | 0.073 |
| Observations | 600 | 600 | 600 | 600 | 600 |

Probit marginal effects with robust standard errors in parentheses. Dependent variable is an indicator variable for whether the participant altered their belief estimate. *** p<0.01, ** p<0.05, * p<0.1

# 10 Instructions

In total we used six sets of instructions.

- Four versions for the Primary experiment which consisted of two waves × two versions (Main and Control). [*German*]
- One version for the Follow-up experiment. [*German*]
- One version for the Validation experiment. [*English*]

We present here three sets of instructions for brevity: **Primary experiment: wave 1 (Main treatment)**, **Primary experiment: wave 2 (Control treatment)**, and the **Validation experiment (Human treatment)**. All instructions have been translated to English (when applicable).

The differences between wave 1 and 2 are that wave 2 included an additional Part 3 where participants had their willingness to pay (WTP) elicited to be matched to a new teammate 2. Part 3 then repeated Part 2 (updating with feedback) with either the old or new teammate, depending on the outcome of the WTP procedure. An additional difference between waves 1 and 2 is that wave 2 elicited beliefs about the full distribution (four states), while wave 1 elicited only beliefs about each teammate being in the top half.

The difference between Main and Control was that in Main the participants themselves were playing the role of teammate 1, and hence their test performance was directly

relevant for payoffs. In Control the participants were making identical decisions for randomly selected teammates 1 and 2. Hence in Control the participants own performance was not relevant.

The key difference between Main (wave 2) and Follow-up was that the human teammate 2 was replaced by a random fundamental (referred to as a "random factor") that could either be HIGH or LOW. Rather than being shown the number of questions attempted by teammate 1, participants were shown the prior probability of the random factor being HIGH (which was given as a range of $\pm$ 15 percentage points – see Section 5.1 in the main paper). Additionally the Follow-up did not include Part 3.

The Validation experiment is distinct and hence presented separately. The differences in the RF treatment (not shown) referred to a matched random factor instead of a matched human. Instead of a picture of university students, it included a graphic of a computer with HIGH and LOW on the screen, along with a spin wheel with numbers from 1 to 20. There was no test shown in the RF treatment, and the measure of oneness was not elicited.

## 10.1 Primary Experiment Wave 1 (Main Treatment Shown)

**Welcome to this experiment!**
In this experiment you will be asked to make several decisions that determine your earnings. For showing up on time you receive 5 EUR. During the experiment you can earn additional money. How much? That depends on the decisions that you make during the experiment. Please read and follow the instructions carefully. They contain everything you need to know to participate. All decisions are taken anonymously, that is, the identity of the participants is not revealed to the other participants at any point. Also the payout is going to take place anonymously at the end of the experiment.

Please note that from now on and during the entire experiment, no communication is permitted. Throughout the experiment the use of mobile phones, smartphones, tablets or the like is prohibited. An infringement of these rules will result in exclusion from the experiment and all payments.

If you have questions at any point, please stretch your hand out of the booth and the experimenter will come to you to answer them in private.

The experiment consists of 2 parts. In the following we will explain to you Part 1 of the experiment. We will present you with the instructions for Part 2 after you have finished Part 1.

**Part 1 – The knowledge and IQ test**
The computer screen will present you with a set of **15 trivia and logic questions**. Questions similar to these are often used to measure a person's general intelligence (IQ). Your task is to answer as many of these questions correctly as possible.

Your score is determined as follows. For every correct answer you receive 2.5 points, for every wrong answer you lose 1 point (you cannot end up with less than 0 points). Questions that you do not answer will not affect your score. Your score will be transformed into Euros at the end of the experiment. The exchange rate is 1 point = 0.10 EUR.

| Payoff example:<br>Suppose you attempted 12 out of the 15 trivia and logic questions and answered 9 of them correctly and the remaining 3 incorrectly.<br>That means, your score is 9*2.5 – 3*1 = 19.5 points<br>Your payments from this part of the experiment then amount to 1.95 EUR. |
|---|

You will be given **10 minutes** to solve all questions. A timer in the upper right corner of your screen will inform you how much time (in seconds) is left to finish the test. You may use the backside of theses instructions to take notes, if you wish to do so.

You will not learn your score until the end of the experiment.

***Attention! The test will begin shortly on the screen!***

**Part 2 – The team task**

*There are no more tests for the remainder of this experiment.*

In this part of the experiment you will be matched randomly with a teammate. You will have the chance to earn additional 10 EUR. How likely this is will depend on your score and your teammate's score in the previous test and the decisions that you are going to take in this part of the experiment.

Remember, the score depends on the number of correctly answered questions (+2.5 points each) and the number of incorrectly answered questions (-1 point each).

The identity of the teammate will never be revealed to you. The only information that the computer is going to present you with is the number of questions that you and the teammate *attempted* to answer in the test.

Before the start of Part 2 the computer has separately compared your and your teammate's scores with the *same* randomly selected group of 19 other participants in today's experiment:

- **You** have been compared with the 19 participants and are either **ranked in the Top 10 or Bottom 10** of the 20 performances.

- Similarly, **your teammate** has been compared with the same 19 participants and is similarly **ranked in either the Top 10 or Bottom 10**.

**What determines whether you earn the additional 10 EUR in this part of the experiment?**

- If both your and your teammate's scores on the test were in the **Top 10**, you will earn the 10 EUR for sure (100% probability of winning).

- If both your and your teammate's scores on the test were in the **Bottom 10**, you will NOT earn the 10 EUR for sure (0% probability of winning).

- If ONE of either your score or your teammate's score was in the Top 10, and the other was in the Bottom 10**, your probability of earning the 10 EUR depends on you choosing a "weight" from 0 to 100** that you assign to your own performance in the team.

*Being in the Top 10 indicates that one is in the top 50% (half) of performers. Being in the Bottom 10 indicates that one is in the bottom 50% (half) of performers.*

We will now explain this "weight" in more detail.

**Part 2 – Choosing a weight**

If ONE of either your score or your teammate's score was in the Top 10, and the other was in the Bottom 10, the probability of earning the 10 EUR is calculated as follows:

a) If you scored in the Top 10 (and your teammate did not): $\sqrt{\text{Weight}/100}$

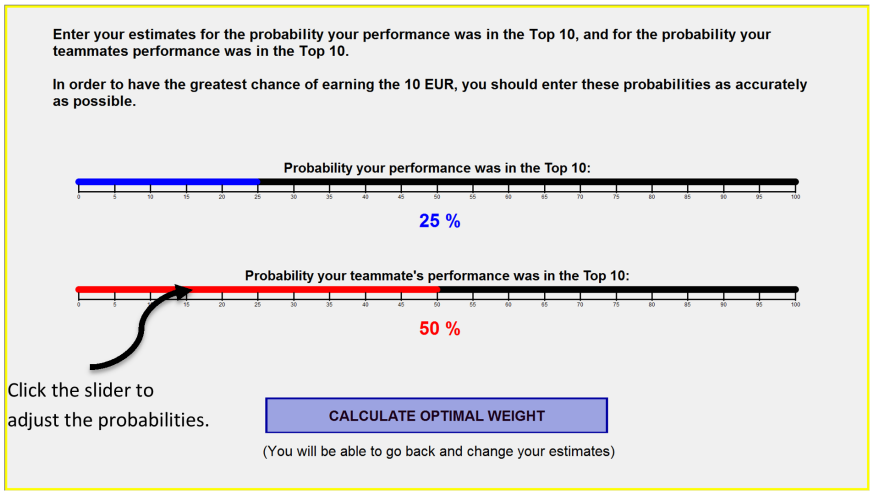
b) If your teammate scored in the Top 10 (and you did not): $\sqrt{(100\text{- Weight})/100}$.

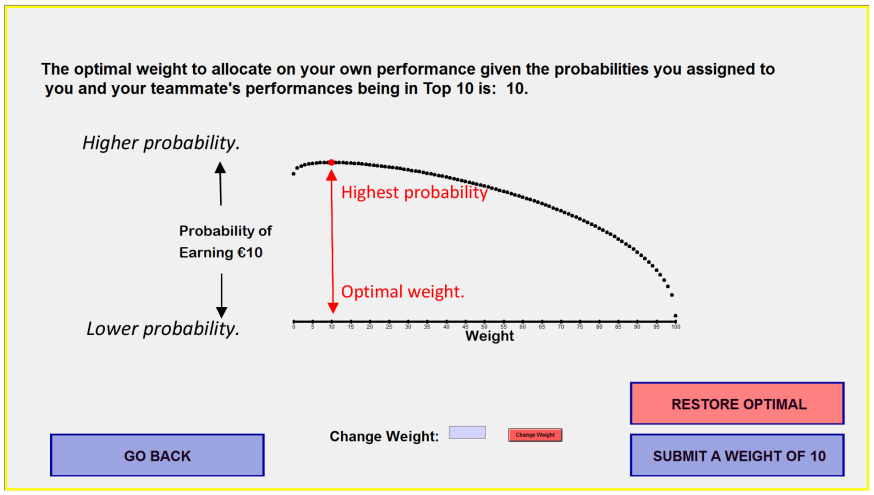
Don't worry if this looks confusing. **You will be asked to enter the probabilities you believe you and your teammate scored in the Top 10.** Then the computer will automatically calculate the weight on your own performance that gives you the highest probability of earning the 10 EUR.
To guarantee the highest probability of earning the 10 EUR, you should choose a weight **closer to 100 if you believe you were more likely than your teammate to score in the Top 10**, and **closer to 0 if you believe your teammate**

**was more likely to score in the Top 10**. If you think that both you and your teammate have the same likelihood of scoring in the Top 10, then the optimal weight would be 50.

Here's an example of how it works. Suppose you believe that there is a 25% chance your performance was in the Top 10, but you think there's a 50% chance your teammate's performance was in the Top 10. The computer will calculate the probability of earning the 10 EUR for *every* possible weight from 0 to 100 and inform you which weight you should *optimally* choose given the probabilities that you enter.

Given the above probabilities, this is how the information would be presented to you on the screen:



**Summary**

In order to have the best chance of earning the 10 EUR, it pays to estimate your performance and your teammate's performance as accurately as possible. The computer then calculates the optimal weight for you.

If you are satisfied with this weight, you can "confirm" this weight on the following screen and submit it for evaluation.

**OR**

If you want to put a different weight, you may either do this directly,

**OR**

You may go back to the previous screen and enter different probabilities with which you expect that your and the teammate's performances were in the Top 10. Based on this, a new weight will be calculated and displayed to you.

**Feedback**

After you will have submitted the weight, on the next screen, a Team Evaluator will give you either a "Green Check" or a "Red X", which is related to your performance and your teammate's performance.





However note that Team Evaluators can sometimes make mistakes:

- If you both scored in the Top 10, a Team Evaluator will give you a Green Check in 9 out of 10 cases. In 1 out of 10 cases he will give you a Red X.
- If one of you scored Top 10 and the other one scored in the Bottom 10, a Team Evaluator will give you a Green Check in 5 out of 10 cases. In other 5 out of 10 cases, he will give you a Red X.
- If both of you scored in the Bottom 10, a Team Evaluator will give you a Green Check in 1 out of 10 cases. In 9 out of 10 cases he will give you a Red X.

A **Green Check** indicates that it is **unlikely** that **both of you** scored in the **Bottom 10**, while a **Red X** indicates that it is **unlikely** that **both of you** scored in the **Top 10**.

That means, while a team evaluator gives you some information about you and your teammate's performances, you do not learn your or your teammate's performances for sure.

**What happens next?**

After receiving the Team Evaluator's feedback the computer will ask you once more to **enter the probabilities you believe you and your teammate scored in the Top 10.** Then the computer will again automatically calculate the weight that gives you the highest probability of earning the 10 EUR.

As before, you may plug in and try out different numbers. Once you are satisfied with the weight that the computer calculates for you – or, alternatively, once you have entered a weight yourself – you can "confirm" this weight and submit it for evaluation.

**This procedure (evaluator feedback – entering performance expectations and assigning weights) will be repeated 4 times**. So that, in total, you receive feedback from 4 different evaluators, after which the computer asks you to enter your beliefs and weights again. In each round you have the chance to earn the 10 EUR in the way described above.

Please note: Each time you receive feedback from a Team Evaluator, it is *completely independent* of the feedback you previously received. Also note that the weight(s) you have chosen have no effect on feedback. Feedback is just about you and your teammate's performances.

**What determines your earnings today?**

Besides your earnings in the trivia and logic test in Part 1 of the experiment, ***one out of the five*** weighting decisions that you enter in Part 2 (your initial one and the four that you enter after having received evaluator feedback) will be randomly selected for payment. **Thus, every decision you make is important as it may affect your earnings if it is chosen for payment.**

At the end of the experiment the computer will inform you about its selection on your screen.

On the next page, we will ask you to answer some control questions. This is to make sure that all participants understand the instructions of this experiment. Once all participants have answered all questions correctly, Part 2 of the experiment will start.

If you have questions about any of the instructions (now or later), please stretch your hand out of the booth and the experimenter will come to you to answer them in private. Please don't hesitate to ask us questions about any doubts you might have!

**Control questions**

**Control Question #1**

If both your performance and your teammate's performance were in the Bottom 10, what is the probability you will earn 10 EUR in this experiment?

_____ Percent

**Control Question #2**

In the screenshot that is presented on page 3, given the probabilities that you assigned to your and your teammate's performances being in the Top 10, what is the weight that you should choose to maximize your chances of earning 10 EUR?

_____

**Control Question #3**

a) Assume you and your teammate are both in the Top 10, in how many cases will you receive a Green Check from an evaluator?

_____ out of 10 cases

b) Assume you and your teammate are both in the Top 10, in how many cases will you receive a Red X from an evaluator?

_____ out of 10 cases

c) Assume exactly ONE of you or your teammate is in the Top 10, the other one is in the Bottom 10, in how many cases will you receive a Red X from an evaluator?

_____ out of 10 cases

**Control Question #4**

True or false? The weight you choose can also influence the type of feedback the Team Evaluators give you.

☐ True ☐ False

**Control Question #5**

In this experiment how are your earnings determined in Part 2? Please tick the correct answer below:

☐ One of your five weighting decisions will be selected at random for payment.

☐ All five of your weighting decisions will be selected for payment.

## 10.2 Primary Experiment Wave 2: (Control Treatment Shown)

**Welcome to this experiment!**
In this experiment you will be asked to make several decisions that determine your earnings. For showing up on time you receive 5 EUR. During the experiment you can earn additional money. How much? That depends on the decisions that you make during the experiment. Please read and follow the instructions carefully. They contain everything you need to know to participate. All decisions are taken anonymously, that is, the identity of the participants is not revealed to the other participants at any point. Also the payout is going to take place anonymously at the end of the experiment.

Please note that from now on and during the entire experiment, no communication is permitted. Throughout the experiment the use of mobile phones, smartphones, tablets or the like is prohibited. An infringement of these rules will result in exclusion from the experiment and all payments.

If you have questions at any point, please stretch your hand out of the booth and the experimenter will come to you to answer them in private.

The experiment consists of 3 parts. In the following we will explain to you Part 1 of the experiment. We will present you with the instructions for Part 2 and Part 3 after you have finished the previous parts.

**Part 1 – The trivia and logic test**
The computer screen will present you with a set of **15 trivia and logic questions**. Questions similar to these are often used to measure a person's general intelligence (IQ). Your task is to answer as many of these questions correctly as possible.

Your score is determined as follows. For every correct answer you receive 2.5 points, for every wrong answer you lose 1 point (you cannot end up with less than 0 points). Questions that you do not answer will not affect your score. Your score will be transformed into Euros at the end of the experiment. The exchange rate is 1 point = 0.10 EUR.

| Payoff example:<br>Suppose you attempted 12 out of the 15 trivia and logic questions and answered 9 of them correctly and the remaining 3 incorrectly.<br>That means, your score is 9*2.5 – 3*1 = 19.5 points<br>Your payments from this part of the experiment then amount to 1.95 EUR. |
|---|

You will be given **10 minutes** to solve all questions. A timer in the upper right corner of your screen will inform you how much time (in seconds) is left to finish the test. You may use the backside of theses instructions to take notes, if you wish to do so.

You will not learn your score until the end of the experiment.

***Attention! The test will begin shortly on the screen!***

**Part 2 – The team task**

*There are no more tests for the remainder of this experiment.*

In this part of the experiment you will manage a team. Two persons from this experiment will be matched randomly as teammates. They form your team. You will have the chance to earn additional 10 EUR. How likely this is will depend on the teammates' scores in the previous test and the decisions that you are going to take in this part of the experiment.

Remember, the score depends on the number of correctly answered questions (+2.5 points each) and the number of incorrectly answered questions (-1 point each).

The identity of the teammates will never be revealed to you. The only information that the computer is going to present you with is the number of questions that the teammates *attempted* to answer in the test.

Before the start of Part 2 the computer has separately compared the two teammates' scores with the *same* randomly selected group of 19 other participants in today's experiment:

- **Teammate 1** has been compared with the 19 participants and is either **ranked in the Top 10 or Bottom 10** of the 20 performances.

- Similarly, **Teammate** 2 has been compared with the same 19 participants and is similarly **ranked in either the Top 10 or Bottom 10**.
-

**What determines whether you earn the additional 10 EUR in this part of the experiment?**

- If both Teammate 1's and Teammate 2's scores on the test were in the **Top 10**, you will earn the 10 EUR for sure (100% probability of winning).

- If both Teammate 1's and Teammate 2's scores on the test were in the **Bottom 10**, you will NOT earn the 10 EUR for sure (0% probability of winning).

- If ONE of either Teammate 1's score or Teammate 2's score was in the Top 10, and the other was in the Bottom 10**, your probability of earning the 10 EUR depends on you choosing a "weight" from 0 to 100** that you assign to Teammate 1's performance in the team.

*Being in the Top 10 indicates that one is in the top 50% (top half) of performers. Being in the Bottom 10 indicates that one is in the bottom 50% (bottom half) of performers.*

We will now explain this "weight" in more detail.

**Part 2 – Choosing a weight**

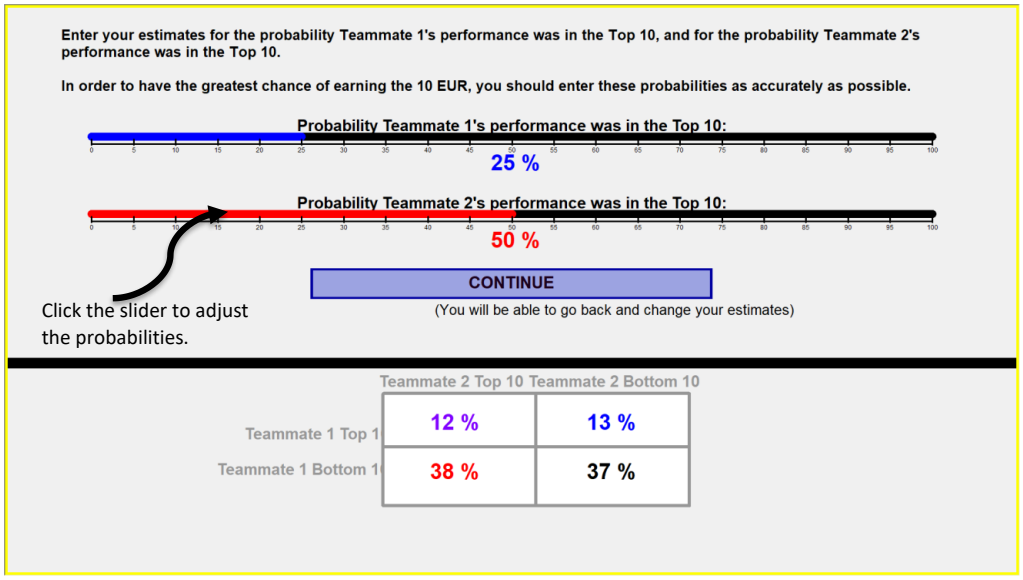
If ONE of either Teammate 1's score or Teammate 2's score was in the Top 10, and the other was in the Bottom 10, the probability of earning the 10 EUR is calculated as follows:

a) If Teammate 1 scored in the Top 10 (and Teammate 2 did not): $\sqrt{\text{Weight}/100}$
b) If Teammate 2 scored in the Top 10 (and Teammate 1 did not): $\sqrt{(100\text{- Weight})/100}$.

Don't worry if this looks confusing. **You will be asked to enter the probabilities you believe Teammate 1 and Teammate 2 scored in the Top 10.** Then the computer will automatically calculate the weight on Teammate 1's performance that gives you the highest probability of earning the 10 EUR.

To guarantee the highest probability of earning the 10 EUR, you should choose a weight **closer to 100 if you believe Teammate 1 was more likely than Teammate 2 to score in the Top 10**, and **closer to 0 if you believe Teammate 2 was more likely to score in the Top 10**. If you think that both Teammate 1 and Teammate 2 have the same likelihood of scoring in the Top 10, then the optimal weight would be 50.

Here's an example of how it works. Suppose you believe that there is a 25% chance Teammate 1's performance was in the Top 10, but you think there's a 50% chance Teammate 2's performance was in the Top 10.



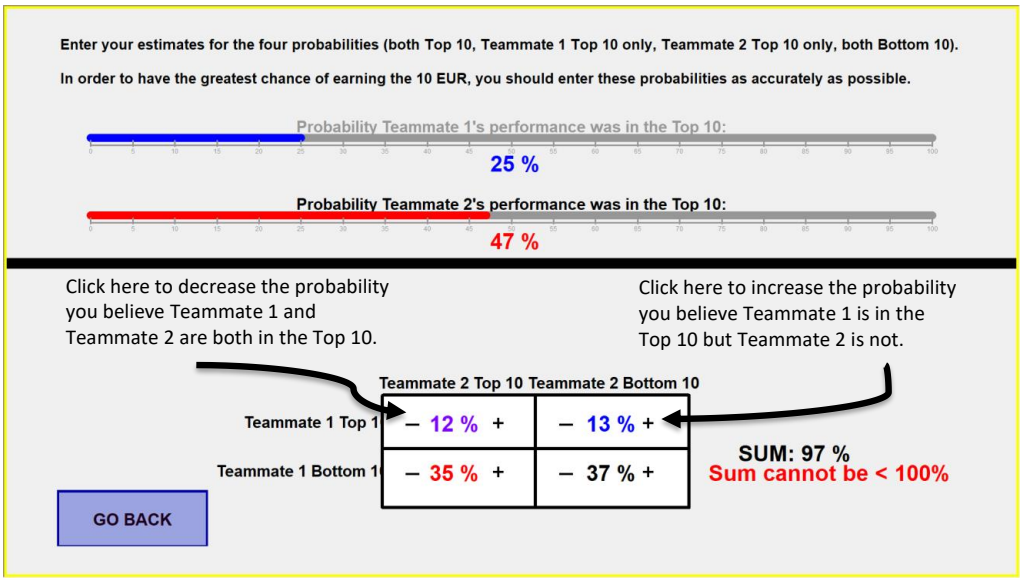

Given what you have entered in the slider, the computer will estimate the probability of each of the following four scenarios in the table at the bottom of the screen, by assuming that Teammate 1's and Teammate 2's scores are independent. The four scenarios are:

1) Teammate 1 and Teammate 2 are both in the Top 10 [top left]
2) Teammate 1 is in the Top 10 and Teammate 2 is in the Bottom 10 [top right]
3) Teammate 1 is in the Bottom 10 and Teammate 2 is in the Top 10 [bottom left]
4) Teammate 1 and Teammate 2 are both in the Bottom 10 [bottom right]

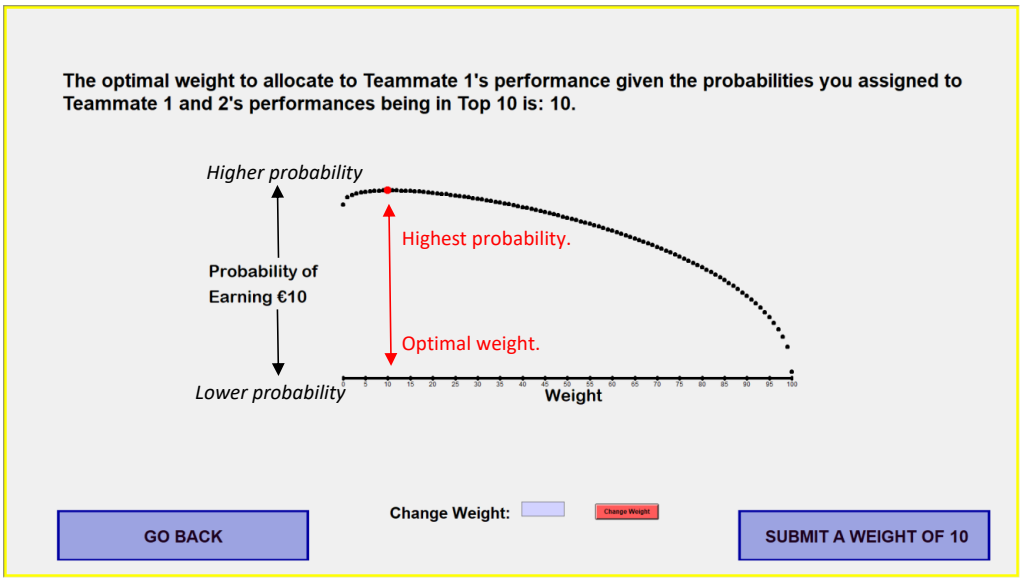
Note that these four scenarios cover all of the possibilities. Therefore they must sum up to 100%.

If you would like, you can adjust any of the probabilities in the table, if you think one scenario is more likely than another. For this, you simply have to click on the "+" and "-" signs that appear next to the figures in the table. You can change the probabilities in the table as many times as you can until you decide that they reflect your true belief. To be able to proceed to the next screen you will only have to make sure that the 4 probabilities sum up to 100 % before you click "CONTINUE".

Given the above probabilities, this is how the information would be presented to you on the screen:




Once you click "CONTINUE", the computer will calculate the probability of earning the 10 EUR for *every* possible weight from 0 to 100 and inform you which weight you should *optimally* choose given the probabilities that you enter.



**Summary**

In order to have the best chance of earning the 10 EUR, it pays to estimate Teammate 1's and Teammate 2's performance as accurately as possible. The computer then calculates the optimal weight for you.

If you are satisfied with this weight, you can "confirm" this weight on the following screen and submit it for evaluation.

**OR**

If you want to put a different weight, you may either do this directly,

**OR**

You may go back to the previous screen and enter different probabilities with which you expect that Teammate 1's and Teammate 2's performances were in the Top 10. Based on this, a new weight will be calculated and displayed to you.

**Feedback**

After you will have submitted the weight, on the next screen, a Team Evaluator will give your team either a "Green Check" or a "Red X", which is related to Teammate 1's performance and Teammate 2's performance.





However note that Team Evaluators can sometimes make mistakes:

- If both teammates scored in the Top 10, a Team Evaluator will give your team a Green Check in 9 out of 10 cases. In 1 out of 10 cases he will give your team a Red X.
- If one of the teammates scored Top 10 and the other one scored in the Bottom 10, a Team Evaluator will give your team a Green Check in 5 out of 10 cases. In other 5 out of 10 cases, he will give your team a Red X.
- If both teammates scored in the Bottom 10, a Team Evaluator will give your team a Green Check in 1 out of 10 cases. In 9 out of 10 cases he will give your team a Red X.

A **Green Check** indicates that it is **unlikely** that **both teammates** scored in the **Bottom 10**, while a **Red X** indicates that it is **unlikely** that **both teammates** scored in the **Top 10**.

That means, while a team evaluator gives you some information about Teammate 1's and Teammate 2's performances, you do not learn their performances for sure.

**What happens next?**

After receiving the Team Evaluator's feedback the computer will ask you once more to **enter the probabilities you believe Teammate 1 and Teammate 2 scored in the Top 10.** Then the computer will again automatically calculate the weight that gives you the highest probability of earning the 10 EUR.

As before, you may plug in and try out different numbers. Once you are satisfied with the weight that the computer calculates for you – or, alternatively, once you have entered a weight yourself – you can "confirm" this weight and submit it for evaluation.

**This procedure (evaluator feedback – entering performance expectations and assigning weights) will be repeated 4 times**. So that, in total, your team receives feedback from 4 different evaluators, after which the computer asks you to enter your beliefs and weights again. In each round you have the chance to earn the 10 EUR in the way described above.

Please note: Each time you receive feedback from a Team Evaluator, it is *completely independent* of the feedback your team previously received. Also note that the weight(s) you have chosen have no effect on feedback. Feedback is just about Teammate 1's and Teammate 2's performances.

**What determines your earnings today?**

Besides your earnings in the trivia and logic test in Part 1 of the experiment, you will be paid for Part 2 and for Part 3. ***One out of the five*** weighting decisions that you enter in Part 2 (your initial one and the four that you enter after your team has received evaluator feedback) will be randomly selected for payment. **Thus, every decision you make is important as it may affect your earnings if it is chosen for payment.** Part 3 will be explained after Part 2 is over, but don't worry, your decisions in Part 2 will not affect your earnings in Part 3.

At the end of the experiment the computer will inform you about its selection on your screen.

On the next page, we will ask you to answer some control questions. This is to make sure that all participants understand the instructions of this experiment. Once all participants have answered all questions correctly, Part 2 of the experiment will start.

If you have questions about any of the instructions (now or later), please stretch your hand out of the booth and the experimenter will come to you to answer them in private. Please don't hesitate to ask us questions about any doubts you might have!

**Control questions**

**Control Question #1**

If both teammates' performances were in the Bottom 10, what is the probability you will earn 10 EUR in Part 2?

_____ Percent

**Control Question #2**

In the lower of the two screenshots that are presented on page 4, given the probabilities that you assigned to Teammate 1's and Teammate 2's performances being in the Top 10 (in the given example: 25% and 50%, respectively), what is the weight that you should choose to maximize your chances of earning 10 EUR?

_____

**Control Question #3**

a) Assume both teammates are in the Top 10, in how many cases will your team receive a Green Check from an evaluator?

_____ out of 10 cases

b) Assume both teammates are in the Top 10, in how many cases will your team receive a Red X from an evaluator?

_____ out of 10 cases

c) Assume exactly ONE of the teammates is in the Top 10, the other one is in the Bottom 10, in how many cases will your team receive a Red X from an evaluator?

_____ out of 10 cases

**Control Question #4**

True or false? The weight you choose can also influence the type of feedback the Team Evaluators give your team.

☐ True ☐ False

**Control Question #5**

In this experiment how are your earnings determined in Part 2? Please tick the correct answer below:

☐ One of your five weighting decisions will be selected at random for payment.

☐ All five of your weighting decisions will be selected for payment.

**Part 3 – The new team task**

In this final part of the experiment, everything is the same as in Part 2. Just like before, you will be asked to enter the probabilities that you believe Teammate 1 and Teammate 2 scored in the Top 10. Then the computer will automatically calculate the weight that gives you the highest probability of earning 10EUR.

As before, you will next receive feedback, and then once more you can enter the probabilities you believe that Teammate 1 and Teammate 2 scored in the Top 10. This procedure – entering performance expectations and assigning weights – will be repeated 5 times.

**The only difference to Part 2 is that you will have an opportunity to change Teammate 2 that Teammate 1 is matched with:**

**Changing the teammate**

If you change to a new Teammate 2, that person will be randomly selected from this session, and his or her score will be compared with the same 19 other participants as Teammate 1's score was compared to, just as before.

If you change to a new Teammate 2, the computer will present you with the number of questions that the new Teammate 2 attempted to answer in the test in Part 1.

If you do not change to a new Teammate 2 but stay with the old Teammate 2, you will be reminded of the number of questions the old Teammate 2 attempted to answer in the test in Part 1.

*As before, your decisions do not affect the payoffs of any of the old or new teammates (and they will never learn about your decisions). Also, the new matching depends ONLY on your decisions, the old or new teammate's decisions are completely irrelevant.*

We will now explain to you in detail how you can change Teammate 2:

Changing Teammate 2 comes at a price. On the first screen you will be asked to enter the MAXIMUM amount (from 0.00 EUR to 5.00 EUR) you are willing to pay to change Teammate 2.

The computer has already randomly generated the true price for changing teammates. It is a random amount between 0.00 EUR and 5.00 EUR, too. After entering your maximum amount, the computer will reveal the true price to you.

- If the true price is lower than your stated maximum amount, then Teammate 2 will change and you will pay this true price.
- If the true price is higher than your stated maximum amount then Teammate 2 will not change. You will not pay anything.

As you can see, the best you can do is to enter the highest amount you would be willing to pay to change Teammate 2. If you enter too high an amount, you might pay more than you wanted to change Teammate 2. If you enter too low an amount, you might not change Teammate 2 for a price you'd have been willing to pay.

There are no right or wrong answers, it just depends how much you would be willing to pay.

→If you *wouldn't be willing to pay anything* (even 0,01 EUR) to change Teammate 2, then enter 0.
→If you would be willing to pay up to 2 EUR, then enter 2 EUR.
→If you would be willing to pay *any* price, then enter 5 EUR.

**What determines your earnings today?**

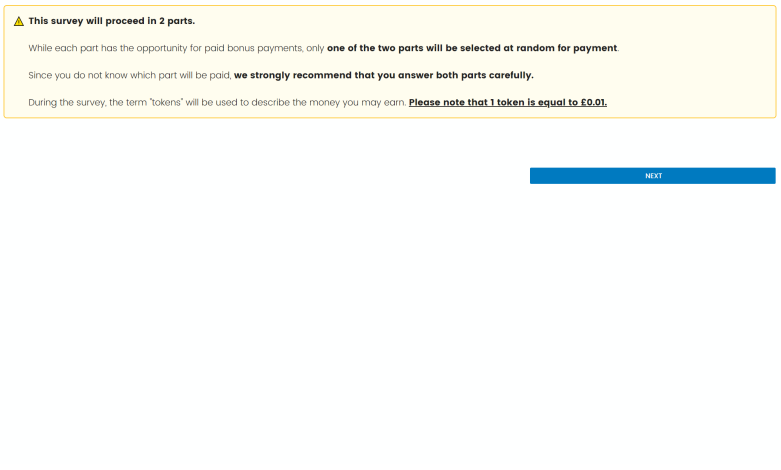
In addition to your earnings in the test in Part 1 and one randomly selected decision in Part 2, you will be paid for one randomly chosen decision in Part 3. **Every decision you make is important as it may affect your earnings if it is chosen for payment.**

At the end of the experiment the computer will inform you about its selection on your screen.

Just like Part 2, one of your weighting decisions will be selected for payment in Part 3. If this decision involved a new Teammate 2, the price for changing to that teammate will be subtracted from your Part 3 earnings

If you have questions about any of the instructions (now or later), please stretch your hand out of the booth and the experimenter will come to you to answer them in private. Please don't hesitate to ask us questions about any doubts you might have!

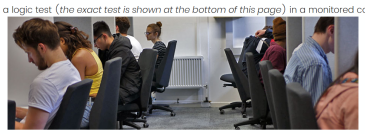
## 10.3 Validation Experiment: (Human Treatment Shown)



**PART 1 OF 2**

You have been **matched with a real person** who took part in a previous study.

This person was a university student, who took a logic test (*the exact test is shown at the bottom of this page*) in a monitored computer room.



**Their performance was then compared to 19 other randomly selected university students who took the exact same test in the same conditions as them.**
Everyone was ranked between 1 and 20 according to their score, where rank 1 was awarded to the best performer, rank 2 to the second best performer, ... and so on.

Your matched **person** was either in the **TOP 10** or in the **BOTTOM 10**.

- If your matched **person** was ranked between 1 and 10, then they were in the TOP 10.
- If your matched **person** was ranked between 11 and 20, then they were in the BOTTOM 10.

You will soon have the chance to earn an additional **50 tokens** *if Part 1 is selected for payment.*

On the next page you will enter your estimate of the probability your matched **person** was ranked in the TOP 10.

**The more accurate your estimate is, the greater your chances are of earning the 50 tokens.**

***Screenshot of the logic test that your matched person completed:***

| | |
|---|---|
| Q1. Which number should come next in this series? 25, 24,22,19,15 | A. 4 B. 5 C. 10 D. 14 |
| Q2. At the end of a banquet 10 people shake hands with each other. How many handshakes will there be in total? | A. 100 B. 20 C. 45 D. 50 |
| Q3. 165135 is to peace as 1215225 is to | A. lead B. love C. loop D. candle |
| Q4. The day before the day before yesterday is three days after Saturday. What day is it today? | A. Tuesday B. Wednesday C. Thursday D. Friday |
| Q5. Library is to book as book is to | A. page B. copy C. binding D. cover |
| Q6. A table and 4 chairs cost $300 in total. The 4 chairs jointly cost $100 more than the table. How much does the table cost? | A. $200 B. $180 C. $160 D. $100 |
| Q7. If it takes 30 shoemakers 30 minutes to repair 30 shoes, how long would it take 150 shoemakers to repair 150 shoes? | A. 150 minutes B. 50 minutes C.30 minutes D.5 minutes |
| Q8. On a tree, there is a patch of moss. Every day, the patch doubles in size. If it takes 96 days for the patch to cover the entire tree, how long would it take for the patch to cover half of the tree? | A. 95 days B. 64 days C. 48 days D. 32 days |
| Q9. Divide 30 by half and add ten. What do you get? | A. 15 B. 20 C. 25 D. 70 |
| Q10. If doctor gives you 3 pills and tells you take one pill every half hour, how long would it be before all pills had been taken? | A. 1 hour B. 1 hour 20 minutes C. 1 hour 30 minutes D. 1 hour 40 minutes |
| Q11. Which one of the five choices makes the best comparison? PEACH is to HCAEP as 46251 is to: | A. 25641 B. 26451 C. 12654 D. 15264 |
| Q12. Mary, who is sixteen years old, is four times as old as her brother. How old will Mary be when she is twice as old as her brother? | A. 20 B. 24 C. 25 D. 26 |
| Q13. Which one of the numbers does not belong in the following series? 2 - 3 - 6 - 7 - 8 - 14 - 15 - 30 | A. 3 B. 7 C. 8 D. 15 |
| Q14. Choose the number that is 1/4 of 1/2 of 1/5 of 200. | A. 2 B. 5 C. 10 D. 25 |
| Q15. John needs 13 bottles of water from the store. John can only carry 3 at a time. What's the minimum number of trips John needs to make to the store? | A. 3 B. 4 C. 4 ⅓ D. 5 |

This is a comprehension check. What best describes what you will do next? (*you must answer this question correctly within two attempts*)

- ○ I will take a logic test.
- ○ I will estimate the score of my matched person on the logic test.
- ○ I will estimate the probability that my matched person was ranked in the TOP 10 on the logic test.

NEXT

Powered by Qualtrics

For this question, you can potentially earn a bonus payment of **50 tokens** based on the accuracy of your answer, *if Part 1 is selected for payment.*

**The most important thing to remember is that the more accurate your estimate is, the higher the chances will be of earning this bonus.**

To see the formula which determines this chance, click below (*opens in a new window*).

Click if you wish to see the exact payment formula.

**What is your estimate of the probability (from 0% to 100%) that your matched person was ranked in the TOP 10?**

*(Click on the gray line below to activate the slider)*

0 10 20 30 40 50 60 70 80 90 100

%

50

SUBMIT ESTIMATE

Powered by Qualtrics

PART 2 OF 2

In Part 1, your estimate of the probability that your matched person was ranked in the TOP 10 was **50%**.

For Part 2, you now have an opportunity to revise and **make a new estimate**.

*If Part 2 is selected for payment*, the determination of your bonus payment of **50 tokens** is still the same as before. However, now you will earn **side payments the *more you change* your estimate**.

- You will receive **1 token for every 5 percent points that you change your estimate away from 50%.**
- This extra amount applies equally, whether you increase or decrease your estimate.
- Note that you are not paid for changes less than 5 percent (e.g. a 14 percent point change would be paid 2 tokens)

⚠ **In summary, you face a trade-off:**

The chances of receiving the 50 token bonus payment are **increasing in your accuracy**.

However, the more that you **revise away from your initial estimate of 50%**, the more guaranteed side payments you will earn.

Click for reminder of exact payment formula.

**What is your *revised* estimate of the probability (from 0% to 100%) that your matched person was ranked in the TOP 10?**

*(Click on the gray line below to activate the slider)*

0 10 20 30 40 50 60 70 80 90 100

%

60

SUBMIT ESTIMATE

Powered by Qualtrics

**Your initial estimate in Part 1 of the probability of the person being ranked in the TOP 10 was 50%.**

**How certain** are you that the optimal estimate in Part 1 was somewhere between 49.0% and 51.0%?

| very uncertain | | | | | | | | | | | | | | | | | | | completely certain |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0% | 5% | 10% | 15% | 20% | 25% | 30% | 35% | 40% | 45% | 50% | 55% | 60% | 65% | 70% | 75% | 80% | 85% | 90% | 95% | 100% |
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ◉ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

NEXT

Powered by Qualtrics

**Thank you for your estimates.**

Before proceeding to the final page of the survey, please answer the following questions.

It may sometimes be difficult or taxing to envision uncertain outcomes. How difficult is it (from 1 to 10) for you to imagine that your matched person was ranked in the TOP 10?

| very easy | | | | | | | | | very difficult |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ○ | ○ | ○ | ○ | ◉ | ○ | ○ | ○ | ○ | ○ |

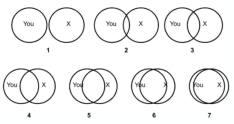
**How warm or cold do you feel towards your matched person?**

- ○ 10 - Extremely warm
- ○ 9 - Very warm
- ○ 8 - Moderately warm
- ○ 7 - Somewhat warm
- ○ 6 - Slightly warm
- ◉ 5 - Neither warm nor cold
- ○ 4 - Slightly cold
- ○ 3 - Somewhat cold
- ○ 2 - Moderately cold
- ○ 1 - Very cold
- ○ 0 - Extremely cold

Now, please **think carefully about the person** you have been matched with.

We now label the person as "**X**".

Please choose the number of the picture below which **best describes your relationship or feelings** towards the person.




| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ◉ | ○ | ○ | ○ |

NEXT

Powered by Qualtrics

**END OF SURVEY**

Thank you for your participation.

Summary of Payment

Part 1 was selected for payment

**The person was ranked in the TOP 10.**

Your estimate in Part 1 was **50%**. According to the formula your percent chance of earning the 50 tokens was **75%**. The computer conducted a simulated lottery where the chances of winning were 75%. The outcome was that **you did NOT receive these tokens**.

Summary of Total Payment:

Your total tokens including the 50 completion tokens is 50.

**Thus your total payment is £0.50.**

Please rate the clarity of the instructions of this survey.

| Very unclear | Somewhat unclear | Somewhat clear | Very clear |
|---|---|---|---|
| ○ | ○ | ○ | ◉ |

Please rate your understanding of this survey.

| I understood little | I understood some parts | I understood most | I understood everything |
|---|---|---|---|
| ○ | ○ | ○ | ◉ |

Please let us know what was confusing, unclear, or any errors. You can also give additional feedback here which helps us improve the survey for others.

**Thank you for participating.**

**Click SUBMIT SURVEY to finalize this survey.**

SUBMIT SURVEY

Powered by Qualtrics

# References

Awad, Edmond, Sydney Levine, Max Kleiman-Weiner, Sohan Dsouza, Joshua B. Tenenbaum, Azim Shariff, Jean François Bonnefon, and Iyad Rahwan (2020) "Drivers are blamed more than their automated cars when both make mistakes," *Nature Human Behaviour*, Vol. 4, No. 2, pp. 134–143.

Bayes, Marjorie A. (1972) "Behavioral cues of interpersonal warmth," *Journal of Consulting and Clinical Psychology*, Vol. 39, No. 2, pp. 333–339.

Enke, Benjamin and Thomas Graeber (2022) "Cognitive Uncertainty," *Working Paper*.

Larrick, Richard P., Katherine A. Burson, and Jack B. Soll (2007) "Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not)," *Organizational Behavior and Human Decision Processes*, Vol. 102, No. 1, pp. 76–94.

Moore, Don A. and Deborah A. Small (2007) "Error and bias in comparative judgment: On being both better and worse than we think we are.," *Journal of Personality and Social Psychology*, Vol. 92, No. 6, pp. 972–989.

Spencer, Steven J., Christine Logel, and Paul G. Davies (2016) "Stereotype Threat," *Annual Review of Psychology*, Vol. 67, No. 1, pp. 415–437.