Journal of Economic Behavior and Organization 207 (2023) 172–202

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Demand for information by gender: An experimental study☆

Karmini Sharma*, Alessandro Castagnetti

Social Sciences Building, University of Warwick, Coventry, CV47AL, United Kingdom

ARTICLE INFO

Article history:
Received 16 November 2020
Revised 21 November 2022
Accepted 19 December 2022
Available online 25 January 2023

JEL classification:
D83
D91
C91

Keywords:
Gender
Information acquisition
Updating

ABSTRACT

We study preferences for seeking information about one's own ability within a male-typed task. Using a lab experiment, we study how women and men might differ in their preferences for information and how these preferences are affected by features of the information environment. We consider a setting where men and women solve a male-typed task, and then choose to receive information about their performance from less or more informative feedback structures. We then vary the costs and benefits of seeking the more informative feedback. In a between-subjects experiment, we make the informative feedback i) private to the subject ii) public iii) subjectively judgemental, and iv) strategic. We find that women are less likely to opt for more informative feedback than men by about 20 p.p. in the first three treatments but this gap vanishes in the strategic treatment. Interestingly, however, this is driven entirely by men who change their behaviour due to strategic incentives. Our results have implications for how men and women learn differently in male-typed domains.

1. Introduction

Occupational segregation of men and women is one of the leading causes of the gender gap in earnings (Altonji and Blank, 1999; Blau and Kahn, 2017). Male dominated industries and sectors tend to be more remunerative and competitive (Azmat and Petrongolo, 2014). Closely linked to this segregation are gender gaps in beliefs about one's ability to perform in gender incongruent domains. Women, in particular, are documented to hold low beliefs about their ability or performance in male stereotypical fields, like maths and science (Bordalo et al., 2019; Coffman, 2014). These beliefs can have adverse impact on gender gaps in self-promotion (Exley and Kessler, 2019), contribution of ideas (Coffman, 2014) and job applications (Coffman et al., 2019b). In this paper, we study the role that endogenous information seeking about own performance plays in shaping gender gaps in beliefs in a male dominated domain, modelled through a male stereotypical task. Information about one's own past performance can help to have well-calibrated beliefs, induce higher effort, increase productivity, social learning and self-confidence (Dobrescu et al., 2021; Azmat and Iriberri, 2010; Blanes i Vidal and Nossol, 2011; Banerjee et al., 2020; Falk et al., 2020). This is important in the light of gender wage gaps that persist even after controlling for occupation and human capital differences between men and women (Goldin, 2014). Stereotypical tasks might affect individuals' demand for performance feedback depriving them of such positive effects of information on (future) performance. In this

☆ Financial support by CAGE, ESRC Grant Ref ES/L011719/1, is gratefully acknowledged. University of Warwick Economics Department IRB approval obtained on October 2019. We thank James Fenske, Christopher Roth, Renke Schmaker and Daniel Sgroi for their feedback and comments. All errors remain our own.

* Corresponding author.
E-mail addresses: karmini3@stanford.edu (K. Sharma), S.castagnetti@warwick.ac.uk (A. Castagnetti).

https://doi.org/10.1016/j.jebo.2022.12.012
0167-2681/

paper, we use a laboratory experiment in which subjects perform a male-typed task and get the choice to receive more or less information about their own performance within different information environments.

In our experiment, all participants perform a male-stereotypical task and are ranked according to their performance relative to all other subjects in the same experimental session. They are then asked to state their prior beliefs of being in the top-half. Subjects then choose to receive information, in the form of signals, from one out of two information structures: one being more informative than the other. They are informed that their choice will be payoff-relevant since they would be asked to report their posterior beliefs, which were elicited in an incentive-compatible fashion. Our experiment features a between-subjects design in which we vary the costs and benefits of receiving information from the most informative information structure. In this way, we can study how information preferences are shaped by the environment in which it is sought. In the first treatment, the private feedback treatment (T1), both information structures disclose information privately. This means that only the subject gets to receive feedback about her own performance. In this sense, there is only an incentive to seek more information. This helps to first understand whether there are initial differences in information seeking preferences.

Then, motivated by common economic and social settings, we designed a public feedback only treatment (T2) where another person in the session observes the feedback that the subject received; while in the public feedback with judgment treatment (T3) this other person also sends a judgment to the subject about her performance. These treatments help to understand whether the publicness or subjectivity of performance feedback can affect gender gaps in information seeking behavior if they exist. For instance, in an academic setting, (Jones et al., 2014) show that women are less likely to want to present their work in academic conferences and less likely to present for higher amount of time. Finally, in the public feedback with hiring decision treatment (T4), the other person has the option to hire the subject if the latter chooses the more informative public feedback. Hiring the subject is optimal if she is in the top half, whereas the subject profits if she is hired regardless of her relative ranking. Thus, T4 introduces incentives related to seeking public information. T4 is modelled after environments where the publicness of information comes with the opportunity to be hired for future projects or opportunities.[1]

Since we incentivized subjects to receive information, we expect that a subject should demand for information unless they have a severe aversion to receiving that information. This aversion could be a means of managing their self-confidence or protecting their ego (Castagnetti and Schmacker, 2020; Möbius et al., 2022). A person who thinks they are of high ability will be more likely to opt for feedback if it is made public, if they expect it to come with subjectively good judgement or if they think that it can get them better opportunities (strategic incentives). These incentives are reversed if the subjects have a low belief about themselves or of their own performance.

The results show a stark difference in information seeking behavior across treatments. First, a significant fraction of women choose the less informative information structure in T1. Around 20% of women choose the less informative information structure, whereas only 3.6% of men do so. Notice, that subjects avoid information even though more information is optimal as subjects know that accuracy in posterior beliefs is incentivized. This finding strongly suggests that women are more likely than men to avoid performance feedback. T2 and T3 do not significantly change the shares of male and female subjects choosing the less informative feedback structure although the results for T3 are a bit noisy. Regardless, the magnitudes are insufficient to remove the gender gap as confirmed by our statistical tests.[2] In other words, public feedback (with or without judgement) does not drive information seeking behavior in our experiment. In contrast, we find that T4 increases the overall share of subjects who choose the less informative structure. Importantly, this is driven mainly by men thereby removing the gender gap in information seeking. Share of men choosing the less informative information structure increases to 21.9% in T4 which is six times the share in T1. This suggests that men are more likely to avoid performance feedback when there are monetary repercussions for revealing low feedback to others. This removes the gender gap in seeking and revealing performance feedback. These results hold when we control for prior beliefs and actual performance of the participants.

Our experiment focuses on male-typed domains. Gender wage gaps in male-typed sectors or domains are especially large. For instance, (Michelmore and Sassler, 2016) show that gender wage gaps are higher in more male-typed sectors within STEM like engineering. Male dominated industries also tend to be more remunerative and competitive (Azmat and Petrongolo, 2014). Exley and Kessler (2018) also focus on a male-typed task in their study due to similar reasons. Large perceived gender differences in ability in male-typed domains can exacerbate selection into such domains. For instance, men perceive themselves to be of a higher ability in finance and also dominate the financial industry (Barber and Odean, 2001). Differences in men and women's performance in maths and science has been the focus of many studies like (Niederle and Vesterlund, 2007) and (Buser et al., 2014). Even after holding ability constant, studies have found that the gender gap in beliefs in male-typed domains can have substantial consequences for academic, financial or career track choices (Buser et al., 2014; Barber and Odean, 2001; Flory et al., 2015). Our results inform discussion about how men and women might learn differently about their own ability in a male dominated environment and how that may affect gender gaps in beliefs. Indeed, we find that demanding less information results in women having more inaccurate beliefs about their own ability than men

[1] Academic settings or modern workplaces are some common avenues where public gatherings (e.g. presentations) are useful not only to receive performance feedback but also for finding potential collaborations and partnerships.

[2] This is discussed in detail in Section 3.3.

in the male typed task in our experiment. This has important repercussions for all downstream effects of inaccurate beliefs we mentioned earlier.

Our paper contributes to several streams of literature. First, we contribute to the literature on information preferences and, more specifically, on information avoidance.[3] Our T1 experiment builds on (Castagnetti and Schmacker, 2020) who study information choices between structures that vary in informativeness. Separately, using T2, T3, and T4 we study the role of costs and benefits of receiving information to study their implications for demand for information which has not been studied before. Further, our task is male stereotypical allowing us to study the influence of male stereotypical nature of the task on information preferences. Relatedly, we contribute to the literature on information structure selection where we focus on information related to one's own performance. This literature studies individual preferences for information that either has instrumental value (Charness et al., 2018; Montanari and Nunnari, 2019; Ambuehl and Li, 2018; Solda et al., 2020) or not (Falk and Zimmermann, 2016; Masatlioglu et al., 2017; Nielsen, 2018; Zimmermann, 2014).

Different papers have shown that individuals have a preference to avoid information that may carry negative news even if the information is costless. In the context of performance feedback, experimental studies have found that a significant fraction of individuals avoid information about their relative rank (Eil and Rao, 2011; Möbius et al., 2022). We contribute to this literature in several ways. First, we study how these preferences are shaped by different costs and benefits such as publicness, judgement and strategic incentives connected to demanding (more) information. Next, our experimental task is purposely chosen to be male stereotypical as opposed to the above-mentioned literature. For instance, Eil and Rao (2011) provide information about two variables: the physical attractiveness of the subjects and their IQ rank while (Möbius et al., 2022) provide information about subjects' performance in an IQ task. In our experiment, the subjects performed a task that included questions on assembling objects, sports, maths, general sciences, and mechanical comprehension which are perceived to be male stereotypical, in particular, maths, science and sports (Bordalo et al., 2019). Results of our paper are also related to the results in (Solda et al., 2020), who show that subjects were more likely to seek biased information about their performance to maintain confidence in order to persuade others. Although they do not test for gender differences, our results in strategic feedback treatment indicate that their results might also be driven by men.[4]

Some of the papers on information demand or belief updating have studied the heterogeneity in demand for information by gender in different contexts. For instance, Castagnetti and Schmacker (2020) look at heterogeneity by gender in demanding information about one's own performance in a gender-neutral task but do not find such difference. Similarly, Eil and Rao (2011) do not find a significant gender difference in willingness to pay to obtain information about one's IQ and beauty ranks in a lab experiment.[5] On the contrary, Möbius et al. (2022) find that women are more averse to feedback than men about their own performance in an IQ task. Our experiment complements them while focusing on a male stereotypical domain and tests whether different features of the information environment like publicness, subjective judgement, and strategic incentives to reveal or not affect this gender gap.

Second, we contribute to a large, relatively recent and thriving literature on gender stereotypes. This literature emphasizes the role of stereotypes in both self-assessments and the evaluation of others (Alan et al., 2018; Bohnet et al., 2016; Bordalo et al., 2019; Carlana, 2019; Coffman, 2014; Coffman et al., 2019a; Milkman et al., 2013). It shows that male stereotypical tasks lead individuals to believe that women's performance is below average. We add to this literature by studying how men and women acquire information about their performance differently. If people hold biased beliefs about their own performance in stereotypical tasks, demanding information might be crucial to correct these beliefs. However, we find that women are less likely to select into receiving more information about their own performance reducing the accuracy of their beliefs. This is different from uniquely studying how individuals react to information that is exogenously provided to them as in (Coffman et al., 2019a). Importantly, we also shut down any channel of anticipated gender discrimination by others since no information about subjects' gender is revealed in public feedback treatment.[6] Relatedly, we add to the literature on gender differences in preferences by providing novel evidence of gender differences in preferences for information.[7] This literature has shown that men and women differ in their preferences for risk (e.g., see Charness and Gneezy, 2012), leadership (e.g., Alan et al., 2020), competition (Niederle and Vesterlund, 2007), among other dimensions.

The rest of the paper is organized as follows. In Section 2, we describe our experimental design. In Section 3, we report the results of our experiment. In Section 4, we discuss our results and in Section 5 we conclude.

2. Experimental design

Key features of the experiment are as follows. First, we asked subjects to perform a test. Second, we asked them to choose between two different information structures from which to receive information about their own relative performance. The only difference between the information structures is how much information they reveal. Third, we varied the implications

[3] See (Golman et al., 2017) for a review of this literature.

[4] For examples in the applied literature on information avoidance in financial settings see Karlsson et al. (2009) and Sicherman et al. (2015); for examples in medical contexts see Ganguly and Tasoff (2016) and Oster et al. (2013).

[5] They find some asymmetry that men want more information when they believe themselves to be at the top of the distribution, and that women need to be paid to get more information when they believe themselves to be at the bottom but don't find these to be significant.

[6] Moreover, in T1, T2 and T3, we shut down any possibility of partners affecting one's payoffs.

[7] For a comprehensive review see Croson and Gneezy (2009).

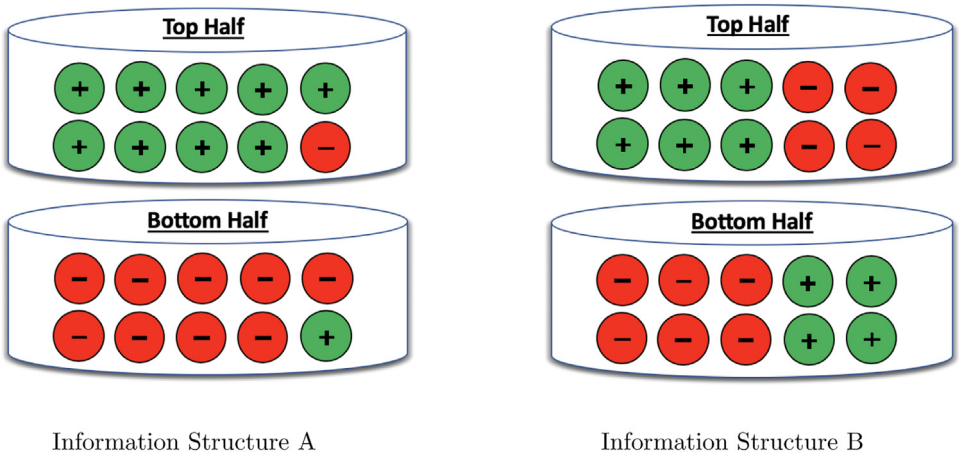


Fig. 1. Information Structures in the Experiment.

of choosing the most informative information structure in a between-subject design. That is, whether it provides the information: a) privately; b) publicly (to another person in the session); c) publicly with possibility of receiving good versus bad judgment from another person; d) publicly with possibility that another person makes a hiring decision based on subject's performance, which monetarily affects the subject.[8] As explained in Section 2.5, subjects signed up for different sessions and the sessions were then assigned to different treatments. Thus the treatment assigned to different subjects is as good as random. In our results later, we also control for baseline controls that were collected to minimize any differences between sessions that occur by chance. We show later that our sample is balanced across all treatments for all except one baseline variable. We eventually show our results with and without controls.

2.1. The quiz and prior beliefs

At the outset of the experiment, participants were asked to solve a quiz. It consisted of 25 questions. The questions were equally split into one of these five categories: assembling objects, general science, maths, mechanical comprehension, and sports. Subjects were given 10 min to answer, in any order, as many questions as they could. We incentivized subjects to exert effort in the quiz. We paid them £2.00 per correct answer out of three randomly chosen questions. Since any three questions could be randomly selected, it helped to ensure against concerns that subjects answered only a certain number of questions to reach a target level of earnings. On average, the subjects answered 13 questions correctly, and the distribution by treatment is provided in Fig. A.17.

We then elicited subjects' beliefs about the likelihood with which they thought that their score in the quiz fell in the top half among all participants in the session.[9] This belief elicitation stage was incentive compatible. We paid subjects either £6.00 or nothing based on the accuracy of their answers following the probabilities matching mechanism (Karni, 2009). We chose this method because it does not rely on risk preferences nor on subjective expected utility.[10] While we explained to subjects the main implications of this method (i.e., that it was in their best interest to report truthfully what they really believed), we did not explain the mechanics of the procedure. We did this as withholding the description of the mechanism increases truthful reporting (see Danz et al., 2020).[11]

2.2. Choices of information structures

Participants were then randomly matched with one other subject in their session. Subsequently, they were asked to choose between two different information structures for receiving information (feedback) about their performance. Irrespective of the choice, performance feedback came in the form of noisy but informative signals. It was informative such that if the subject was in the top (bottom) half of the distribution, she was more likely to correctly receive information about her being in the top (bottom) half. It was noisy because one of the information structure was more accurate than the other.

Figure 1 shows the two information structures. Once the subject chose one information structure, a signal (a ball), corresponding to the subject's relative position in the test, was drawn from the urn. The subject was then shown the signal drawn (a green or a red ball). A green ball with a '+' sign meant that the subject was in the top half of the distribution

[8] Screenshot of the experiment are provided in Appendix A.
[9] In case of ties, they were broken randomly. Subjects were informed about this feature.
[10] For a detailed description of this method, we refer the interested reader to Karni (2009).
[11] However, if any subject demanded more information about the method, they could click on a button to read a full description of the elicitation method.

Table 1
Main Features of the Treatments.

Treatment	Public Signal	Judgment	Hiring Decision
(T1) Private Feedback			
(T2) Public Feedback without Judgment	✓		
(T3) Public Feedback with Judgment	✓	✓	
(T4) Public Feedback with Hiring Decision	✓		✓

and a red ball with a '-' sign meant they were in the bottom half of the performance distribution. As it can be seen, the information structure on the left is more informative than the one on the right. This is because the precision of information structure A is higher than B's. If the subject was in the top half then information structure A provided them with a green ball 90% of the time and information structure B showed green ball only 60% of the time. If the subject was in the bottom half then probabilities were the same but now for a red ball instead of green.

At the time subjects made this choice, they were informed that they would be asked their (posterior) beliefs about their relative ranking again and would be paid according to how accurately they predict it. This meant that acquiring information about their performance was payoff-relevant as it could increase the accuracy of their reported posterior beliefs. This meant that we incentivized the subjects to take-up information structure A. However, we varied the implications of choosing information structure A in a between-subjects design. In particular, our experiment features four conditions: (T1) private feedback, (T2) public feedback without judgment, (T3) public feedback with judgment, and (T4) public feedback with a hiring choice. We now explain them in turn.

2.2.1. (T1) Private feedback

In this treatment, the signal received, irrespective of the information structure chosen, was privately seen by the individual. Hence, if subjects choose the less informative information structure (B) in this treatment, this can be attributed to a preference for avoiding performance feedback. We will focus on gender differences in this demand.

2.2.2. (T2 & T3) Public feedback without and with judgement

In the public feedback without judgment, choosing information structure A implied that the signal received by the subject is also revealed to one randomly chosen person in the subject's session.[12] In the public feedback with judgment, on top of the publicness of the signal from information structure A, the matched partner was asked to send the subject a written message. In particular, the partner had to choose between two predetermined messages: 1) "Your performance must have been really good relative to others", and 2) "Your performance must have been awful relative to others". These messages were deliberately personal to add more subjectivity than the baseline feedback.

In these treatment variations we aim to capture individuals' (un-)willingness to demand performance feedback when it implies that others have access to it too. Moreover, we can also learn whether these preferences are shaped when others could make subjective remarks about the performance.

2.2.3. (T4) Public feedback with hiring decision

In this treatment, the signal from information structure A was not only revealed to the partner but also allowed the partner to make a hiring decision. In particular, after seeing the subject's signal, her partner could choose whether to "hire" her or not. If the subject was hired, then the subject earned £6.00 and nothing otherwise. Importantly, the partner had an incentive to hire a subject who would be in the top half. If he hired a participant who was in the top (bottom) half of the distribution, he would earn £6.00 (£0.00). If instead the partner chose not to hire her he would then earn £2.50 for sure. On the other hand, if the subject chose the information structure B, then the partner was not asked to make a hiring decision and both players earned £2.50 in this part of the experiment. This mimics the real world where not showing one's performance to others might reduce hiring opportunities.[13] This treatment, thus, added a monetary implication of choosing the most informative information structure. The screenshot of how this information was communicated is provided in appendix Figs. A.14, A.15 and A.16.

This treatment allows us to shed light on another force that might prevent people from demanding performance feedback. That is, individuals might have a preference for avoiding performance feedback if that increases their expected returns. This is similar to Solda et al. (2020) however, we focus on a setting where accuracy of posterior beliefs was explicitly incentivized rather than persuasion.

In Table 1 we provide a summary of the treatment variations and we indicate the main features of each treatment.

Subjects were asked to complete a comprehension questionnaire consisting of five questions to ensure that they understood the main features of the information structures and what they entailed for informativeness of the feedback, the

[12] Remember that subjects were randomly matched in pairs of two after the prior belief elicitation.

[13] Again taking the example of academia, it is generally considered costly if researchers do not present their research to the scientific community in terms of future collaborations or job opportunities.

publicness of the signals and their partner's set of actions. They could not proceed to the information structure selection until they answered these questions correctly.

2.3. Signal received and posterior beliefs

Subjects then received the signal from their chosen information structure. They were then asked to state their belief about being in the top half in their session. We again incentivized subjects to report their beliefs truthfully with the same belief elicitation procedure explained in Section 2.1.

Please note that since the two members in the pair were facing the same decisions, it meant that in the public feedback treatments, each subject was shown the signal of her partner if she/he chose information structure A. In the public feedback with judgment she was asked to send the partner one of the two predetermined messages as well; while in the public feedback with hiring decision, she was asked to decide whether to hire the partner or not as well.

Across all treatments, subjects were first shown the signal for their own performance as per their chosen information structure and asked about their posterior beliefs (which was incentivized) as shown in Fig. A.13. They were then shown the signal of their partner (if the partner chose the public feedback) and asked to send the judgement statement (in T3) or to make the hiring decision (in T4). Then finally they were shown the judgement statement chosen by their partner (in T3) or the hiring decision made by him/her (in T4) if the subject had chosen information structure A.

2.4. Debriefing

Participants were then asked to answer some debriefing questions. First, we asked subjects to explain in their own words why they chose one information structure over the other. We paid them £0.50 for their answers. Next, we elicited, in an incentive compatible fashion, participants' risk preferences with the (Gneezy and Potters, 1997) risk elicitation task. We then asked subjects to complete a questionnaire that measures individuals' overconfidence in terms of over precision.[14] It is based on (Block and Harper, 1991) and it asks participants to specify a confidence interval to their answers such that there is a 90% chance that each answer falls inside it according to them. Finally, we asked them to answer a questionnaire. It included demographic questions such as participants' age, country of origin, and gender. We also asked them whether they had participated in experiments before and if they knew anyone in the session. Participants were finally asked a general willingness to take risk question (Dohmen et al., 2011).

To avoid hedging motives across the different payoff-relevant parts of the experiment (including the quiz payments), we randomly selected one part of the experiment to count for final payments. This means that the participants were paid for either the quiz, the prior-posterior belief elicitation or the debriefing. We informed participants about this feature of the payment scheme at the outset of the experiment. This helps to alleviate the concern that participants may exert effort in only one or two parts of the experiment once they think they have performed well or reached a certain financial goal within the experiment.

2.5. Implementation

The experiment was conducted in fall 2019 at the Economics laboratory of Warwick University. Overall, 344 participants, recruited through the SONA recruitment software, took part in the experiment. We conducted 24 sessions of about 45 min each. Participants earned an average payment of £11.00, including the show-up of £5.00. We programmed the experiment with oTree (Chen et al., 2016). Descriptive statistics of the sample are provided in Appendix B.

Screenshots of the experiment instructions are provided in Appendix figures A1 to A14.

3. Results

We first discuss the summary statistics of our sample and then look at the prior beliefs. We then look at the information structure choice by treatment and gender. In this analysis, we also perform econometric regressions to quantify our results and show the robustness of our findings. Finally, we look at posterior beliefs to study how subjects learnt from the signals and formed their posterior beliefs. We then discuss the overall results and implications for efficiency.

3.1. Descriptive statistics

The summary of the characteristics of our sample is provided in appendix Table B.1. The summary statistics show that between 52% to 61% of the experimental subjects were women for a mean of 56%. The average age of the subjects was 21 (since these were university students), majority of the students had a quantitative education (depending on which department you belonged to), and most of the subjects are in the middle of the risk preference elicitation range. That is, most perceive themselves as not that willing to take risks[15] and subjects were willing to bet close to 50% of the amount in a

[14] See Moore and Healy (2008) for a detailed description of the different ways in which individuals can be overconfident.

[15] Higher the number, higher is the willingness to take risks

Table 2
Gender differences in the choice of information structure B by treatment.

	Proportion Men	Proportion Women	All
(T1) Private Feedback	0.04	0.24	0.15
(T2) Public Feedback without Judgment	0.06	0.17	0.13
(T3) Public Feedback with Judgment	0.10	0.19	0.15
(T4) Public Feedback with Hiring Decision	0.22	0.23	0.22
All Treatments	0.11	0.20	0.16

The table provides the proportion of participants who choose information structure B by treatment and gender.

risky lottery. On average, the subjects think that there is a 53% chance that they will be in top half of the performance distribution, close to 50% had a previous experience with doing lab experiments and on average 29% knew others in the lab. These characteristics are balanced across treatments for all variables except for the share who knew others in the lab. Close to 60% knew each other in sessions that were assigned to treatment 1 while it was close to 22% in the other sessions. In any randomized experiment there can be imbalances by chance, but to make sure that these differences are not driving our results, we show results when controlling for all these demographic variables together. The variables we include in the regressions are : age, education (year of study in the university), department you are enrolled in, , country of origin, ethnicity, mother tongue, previous experience with experiments, share of other subjects they know in the experiment, their own perception of risk aversion, risk choice, and their prior belief of their own performance. We show the questions we asked for demographics in the experiment in appendix Section B.1.

We depict the distribution of priors in Fig. A.20, of posteriors in Fig. A.21, and of actual performance in Fig. A.22 by gender and treatment in each case. In all treatments, men's priors on average seem much higher than that of the women. Their distribution also seems shifted more to the right than that of the women. Comparing the posterior with the prior distribution, we find that both men and women's distributions become more dispersed after the treatments. Looking at the performance, there is no significant difference between men's and women's performance in any of the treatments. We show the average differences in appendix Table B.3 where indeed the differences are small.

3.2. Accuracy of prior beliefs

To measure prior beliefs, we asked subjects their belief about being in the top 50% of the performance distribution in the session. The mean prior belief across all individuals is 53.15 (s.d. 22.23). However, as expected due to the gender stereotype of the task, prior beliefs for men are higher than those for women. While for men the average belief is 61.66 (s.d. 20.46), it is only 46.49 (s.d. 21.32) for women. We control for prior beliefs given these gender differences. In this way, we can also assess whether gender effects in information structure choices are driven by prior beliefs.

Importantly, we can also look at how accurate prior beliefs are. To do this, we can study the distance between the prior belief and participants' true rank. For those in the top half, this means the distance between 100% and their prior belief, whereas for those that are in the bottom half, the distance is exactly their prior belief. That is, for subjects in the top half the (lack of) accuracy is *[(I(In the top half)=1)-Prior Belief]* and for those in the bottom half the error is *[(I(In the top half)=0)-Prior Belief]* where prior belief is the probability with which participant thinks that he/she is in the top half of the distribution in their session. The average distance is 42.82 (s.d. 21.28) and it is essentially the same across men and women: 42.39 (s.d. 22.30) for men and 43.16 (s.d. 20.49) for women.

If subjects were fully calibrated in their beliefs, the average distance should be zero; whereas if they knew nothing it should be around 50%. We see that the mean distance is somewhere in the middle but substantially closer to 50%. In sum, what this analysis strongly suggests is that there is scope for learning and the subsequent choice on information structures is relevant.

3.3. Information structure choice

We start by providing an overview of information structure choices by treatment and gender. Table 2 shows the proportion of subjects who chose information structure B over A. Interestingly, men and women react very differently to the treatment conditions. For men we can see that as we move from (T1) to (T4) the proportion of subjects choosing B over A increases from 3.6% to 21.9%. While for women the proportion of those who choose information structure B is around 20.0% in all treatments.

We now perform the regression analysis to quantify our experimental results. We start by looking at information structure choices in (T1) and by gender. In particular, we start with the following specification:

$$Y_i = \alpha + \beta_1 Female_i + X_i' \beta_2 + \epsilon_i \quad (1)$$

Y_i is a dummy variable which takes the value 1 if participant i chooses information structure B. $Female_i$ is a dummy taking the value 1 if participant i is a female. X_i refers to control variables for individual i. Here, β_1 is the difference in take up of the less informative information structure between men and women in (T1). This coefficient informs us on

Table 3
Choice of Feedback Mode B in (T1).

	OLS			Probit		
	(1) Info B	(2) Info B	(3) Info B	(4) Info B	(5) Info B	(6) Info B
Female	0.201**	0.220**	0.238**	0.201***	0.210***	0.175**
	(0.079)	(0.085)	(0.115)	(0.079)	(0.079)	(0.073)
Prior Belief		✓	✓		✓	✓
Demographics			✓			✓
R-Squared	0.077	0.098	0.205	0.105	0.131	0.313
N	66	66	66	66	66	66

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

gender differences in take up of the less informative information structure. Robust standard errors are reported in the parenthesis.

Table 3 shows the results. In Columns (1)-(3) we report the results of OLS regressions, while in columns (4)-(6) we report those for probit regressions.[16] In Columns (2) and (5) we add subjects' prior beliefs as control variable, while in Columns (3) and (6) we further add the demographic control variables. We report marginal effects for probit specifications for ease of interpretation.

We can see that women are 20 percentage points more likely to choose the less informative information structure (B). The coefficients are stable across all specifications after including the prior beliefs and the rest of the controls mentioned previously. In sum, in (T1) there is a highly significant gender difference in take-up of information. We can also test whether the proportion of individuals that chose information structure B over A for each gender is equal to zero. While for men we find that it is not statistically significantly different from zero (F-test, p-value=0.326), for women it is statistically different from zero (F-test, p-value=0.013). This shows that, despite more informative signals being monetarily incentivized, women have a preference for avoiding information about their own performance in the quiz.

In appendix Table B.5, we show results with session fixed effects. Although there were only 24 sessions, and the fixed effects necessarily reduce the degrees of freedom, especially for treatment 1 which has a total of 66 subjects, we find that our results remain significant (although now are marginally significant). The coefficients are slightly lower than before but indicate that women are close to 18 p.p more likely to demand less informative feedback than men. We also clustered the standard errors at the session level and find that the results stay qualitatively similar to the above for both linear and the probit models although for OLS we lose the power to detect the effects when we add all the controls. These results are in appendix Table B.8.

Now, to understand the impact of the *publicness* of the feedback (signals) on information seeking behavior, in appendix Table B.4, we show the results by augmenting Eq. (1) to include variables $T2$ and $T3$, which are dummy variables taking the value of 1 if the treatment was (T2) or (T3), respectively. We also include their interaction with the female dummy (T1 $\times$ Female and T2 $\times$ Female). We do not find that there is an additionally significant impact of treatments (T2) and (T3) on information structure selection between men and women as depicted by the insignificant coefficients on the dummies or their interactions with gender. We find that women stay between 10.9 to 8 p.p more likely to opt for less informative feedback. These results suggest that making a signal public (with or without judgement) does not affect the preference of women for less informative feedback.[17]

We study the impact of adding strategic considerations in (T4) on information preferences. Before running any regression, in Fig. 2 we provide graphical evidence of information structure choice by gender. In the figure, we pool together the results from treatments (T1), (T2), and (T3). From the figure, we can see that gender differences in the choice of information structures are stark between men and women in the non-strategic treatments while in the strategic environment (T4) the gender gap fully disappears. The difference shrinks from 0.122 (t-test p-value=0.006) to 0.005 (t-test p-value=0.956).

[16] Throughout this section we conduct probit regressions along with a linear probability regression model since our outcome variable is dichotomous.

[17] The results with session fixed effects are in the appendix Table B.6 and with clustered standard errors are in appendix Table B.9. With session fixed effects, our coefficients go down slightly and become marginally significant but our results stay similar qualitatively. This is intuitive since session fixed effects affect our degrees of freedom and reduce the power. With standard clustering our results stay the same, but with wild cluster bootstrap we find that our OLS estimates either become marginally significant or lose it but stay below significance of 15%. On the other hand, our probit estimates stay significant much like the session fixed effects. Overall, we conclude that our results indicate that T2 does not have any impact on the gender gap which remains significantly different from zero while T3 has some effect in reducing the gender gap but only to a smaller extent.

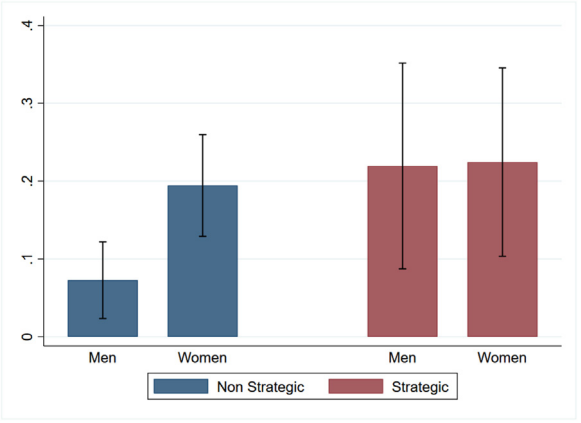


Fig. 2. Take Up of Information Structure B by Strategic/Non Strategic Environment. *Notes:* Gender differences in take up of private feedback in the strategic and not strategic treatments.

Table 4
Overall difference in take up of less informative feedback, all treatments.

	OLS			Probit		
	(1) Info B	(2) Info B	(3) Info B	(4) Info B	(5) Info B	(6) Info B
Female	0.201**	0.198**	0.2011**	0.237**	0.235**	0.242**
	(0.078)	(0.080)	(0.083)	(0.100)	(0.101)	(0.095)
T2	0.025	0.026	0.040	0.062	0.064	0.088
	(0.055)	(0.055)	(0.059)	(0.148)	(0.149)	(0.150)
T3	0.066	0.067	0.082	0.138	0.139	0.175
	(0.056)	(0.056)	(0.060)	(0.142)	(0.142)	(0.141)
T4	0.184**	0.186**	0.198***	0.294*	0.299*	0.356**
	(0.074)	(0.075)	(0.075)	(0.159)	(0.159)	(0.157)
T2 × Female	-0.092	-0.096	-0.118	-0.094	-0.098	-0.111
	(0.103)	(0.102)	(0.104)	(0.099)	(0.096)	(0.075)
T3 × Female	-0.114	-0.117	-0.110	-0.124*	-0.127*	-0.126**
	(0.105)	(0.104)	(0.104)	(0.075)	(0.073)	(0.058)
T4 × Female	-0.196*	-0.202*	-0.227*	-0.162***	-0.164***	-0.173***
	(0.119)	(0.118)	(0.121)	(0.051)	(0.049)	(0.094)
Prior Belief		✓	✓		✓	✓
Demographics			✓			✓
F-test $\beta_1 + \beta_7 = 0$	0.955	0.963	0.775	0.954	0.938	0.979
R-Squared	0.033	0.034	0.0375	0.042	0.043	0.103
N	344	344	344	344	344	344

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

To quantitatively analyze the impact of the strategic treatment on information structure choices, we now run the following regression:

$$Y_{ij} = \alpha + \beta_1 Female_{ij} + \beta_2 T2_j + \beta_3 T3_j + \beta_4 T4_j + \beta_5 T2 \times Female_{ij} + \beta_6 T3 \times Female_{ij} + \beta_7 T4 \times Female_{ij} + X'_{ij}\beta_8 + \epsilon_{ij} \quad (2)$$

Here we look at the additional impact that the strategic environment in (T4) has on the choice of the less informative information structure compared to (T1) for men. β_4 provides this information. Moreover, β_7 captures how and whether the strategic incentives (of being hired) affect the gender-gap in information-seeking behavior.

The results are shown in Table 4.[18] We find that (T4) leads to a large increase in take up of private feedback for men compared to T1 (the omitted category). This increase amounts to approximately 18 to 19 percentage points in columns 1,

[18] In Columns (1)-(3) we report the results of OLS regressions, while in columns (4)-(6) we report those for probit regressions. In Columns (2) and (5) we add subjects' prior beliefs as control variable, while in Columns (3) and (6) we add, on top, the demographic control variables.

2 and 3. T4 additionally does not have a significant impact on women's take up of information, which is same as that in T1. As we saw before, there is no differential impact of (T2) or (T3) on gender gap in take-up of private feedback, but we find that T4 removes the gender gap entirely. The coefficients against $T4 \times Female$ are almost equal to the coefficient against *Female* (20 p.p). An *F*-test for whether the sum is zero of the coefficients against the female dummy and the interaction with (T4) cannot be rejected. This shows that the there is no additional impact of T4 on women's take up information. As we saw in Table 2 and Fig. 2, this result is entirely driven by men (also apparent from the estimates for the β_4 coefficient, which gives the effect of strategic considerations on men's take-up of private information). This shows that the strategic incentive treatment increases take-up of private (less informative) feedback only for men indicating men's sensitivity to (strategic) incentives. Probit estimates show a similar story.[19] The coefficients against (T4) are highly significant for probit specifications. Women, on the other hand, continue to demand less informative feedback and do not respond to such incentives.

We provide results with session fixed effects and clustered standard errors in appendix Tables B.7 and B.10. With session fixed effects our coefficient on gender stays significant although marginal and the value becomes smaller. We still cannot reject that the gender gap is different from zero as given by the *F*-test at the bottom which is our key finding in the main results above. However the fixed effects reduce the power for us to detect significance on the strategic treatment for men while the coefficients stay large. With clustering we find, that most of our coefficients stay marginally significant. Crucially we cannot reject the null that overall the gender gap gets removed under the strategic interactions treatment using wild cluster bootstrap on the sum of coefficients ($\beta_1 + \beta_7$).

We also show the differences in any of the social demographics of those who chose to get feedback from different information structures in appendix Table B.2. It shows that apart from gender, the two groups were similar along all other dimensions. 69% of those who chose information structure B were women, while it only 53% for those who chose information structure A. This confirms that other than gender, the other controls are not driving our results despite there being imbalance for one of our controls at the baseline.

4. Discussion

4.1. Posterior beliefs

We now study individuals' posterior beliefs. The mean posterior belief is 53.36 (s.d. 29.93). While it is 55.89 (s.d. 31.67) for men, it is 51.37 (s.d. 28.42) for women.

We now look at the accuracy of participants' posterior beliefs. That is, whether beliefs are closer to 100% (0%) for participants in the top (bottom) half. We look at this distance depending on the information structure chosen. The average distance is now 38.94 (s.d. 28.01);[20] it is 38.08 for those who chose the most informative information structure and it is 43.33 (s.d. 25.20) for those who chose the less informative one. Comparing this distance to that found in their priors, we see that the distance has significantly reduced for those subjects in the first group (Δ=4.66, p-value=0.005), while it has not for those in the latter group (Δ=1.75, p-value=0.503). To understand gender differences, we find for T1, T2 and T3 the average distance is 38.7 for women and 31.4 for men indicating that women maintain less accurate beliefs about themselves because of less information seeking. This is significantly different (t-test p-value=0.06). Thus, women continue to believe that they are of low ability (despite no prior difference in performance) due to less information seeking. On the other hand, men and women learn less about themselves but equally so in T4 with a distance of 50.6 and 46.4 respectively.

This analysis shows that participants that chose information structure B do not learn as much compared to those that chose information structure A. Thus, subjects who chose the most informative information structure ended up with more accurate posterior beliefs. This highlights that the way that individuals select into receiving information has strong implications in terms of how much they learn and, subsequently, on how accurate their (posterior) beliefs are. Gender differences in take-up of less versus more informative feedback, thus, implies gender differences in accuracy of these beliefs about own ability which has severe downstream effects on women's outcomes.

4.2. Implications for efficiency

In terms of efficiency, we can see that in treatments (T1) to (T3) the rational payoff maximizing choice is to receive information from information structure A. In fact, choosing information structure B is inefficient as it provides less information and, as we have seen, it leads to more inaccurate posterior beliefs. We find that women are more likely to choose the inefficient choice avoiding information on their true rank.

Recall in T4 it was crucial whether one is revealed to be above or below the median to one's partner (if information

[19] Recall that these estimates keep the values of the controls at their mean levels and hence the magnitudes are larger than OLS.

[20] Remember that it was 42.82 for the prior beliefs.

structure A is chosen) for payoffs. Therefore, to analyse the efficiency implications of choosing one information structure over the other, we divide our sample of men and women into two groups: above median and below median performers. In particular, we look at above median performers for whom choosing more informative information structure is the efficient choice both in terms of learning one's true rank and of being hired by the partner.[21] We find that above median performers seek lesser information under strategic than non-strategic environments implying that inefficiency rises in this context. We see lower number of above average performers seeking more public and informative feedback for their performance. This has implications not just for the those seeking feedback, but those who would have a chance to see above average performers' signal of performance. Quantitatively, take up of more information goes down for above median performers by 14 p.p. in the strategic environment (T4) compared to the non strategic treatments (T1-T3). As discussed above, however, strategic incentives do remove the gender gap between men and women but only by increasing men's take up of lesser information. In appendix Fig. A.19 we show the take up of less informative feedback by gender and quartiles of performance. Men in the fourth quartile were less likely to reveal information about themselves and also learn about themselves. Thus, a reduction in gender gap comes at the cost of increased inefficiency overall, and specifically for men.

4.3. Overall results

In a between-subjects design, we asked experimental subjects to perform a male stereotypical task and then to choose feedback about their performance from more or less informative information structures. We find that women are less likely to choose to receive more informative feedback about their own performance than men. This gender gap is of about 20 p.p. and remains stable across most of our specifications.

Women's demand for less informative feedback stays at approximately 20% whether we make more informative feedback public, subjective or strategic indicating a sticky and low demand for information. We cannot rule out null effects of introducing publicness to the feedback or subjective judgement to the more informative feedback. We can rule out any effect of only public feedback treatment without (T2) or with judgment (T3) on the gender gap in demand for information although our results for T3 are larger but not sufficient to remove the gender gap. When we introduce strategic incentives, we detect significant effects of the strategic incentives treatment in closing the entire gender gap in demand for information. However, this is driven by an increase in percentage of men demanding less informative feedback from a 3% to 11% between T1 to T4. Overall thus, there is strong evidence that strategic incentives reduce the gender gap.

The literature on gender stereotypes has shown that stereotypes can affect self-assessments and self-confidence in gender incongruent domains and drive gender differences in behaviour. Controlling for prior beliefs in our specifications does not change our results. On the other hand, when we asked the subjects in the debriefing of why they chose the private feedback, most women answered that they thought that they were in the bottom half and they did not want others to see it in Fig. A.18. Women were also more likely to state that they were less confident about their performance while none of the men mention that. This indicates that prior beliefs might be affecting women in other ways.

To explore beliefs further, we study whether there is any evidence that men and women incorporate prior beliefs differently in their decision to demand information. We additionally control for prior beliefs interacted with gender prior beliefs in appendix Table B.11. We find suggestive evidence that there might indeed be such a gender difference although the evidence is not robust. We find that women and men have significantly different response to their prior beliefs. We find indication of the same for T2 and T3 in appendix Table B.12. This remains a plausible explanation for gender differences we observe in information seeking in T1,T2 and T3. The tables suggests that women and men might respond differently to priors in how they demand feedback about themselves. While a full investigation of these patterns is beyond the scope of this paper, we believe that further research is needed to understand why this might be the case.

Our experiment tests for gender differences only in a male-typed domain. An experiment with female stereotypical task will allow researchers to understand whether the pattern holds only for a male stereotypical domain or not. To conjecture what might happen, we use qualitative answers we collected from our participants in the debriefing discussed above. Both men and women who chose less informative feedback did so because they thought they were of low ability and did not want others to know (in case of T2, T3 or T4). But only women reported that they were not confident in their ability and we also have some indication that women might incorporate prior beliefs differently than men. If men and women indeed incorporate their beliefs differently and if a female stereotypical task affects their beliefs as strongly as the male stereotypical task does, then we expect that men might be less information seeking in a female stereotypical task than women.

Exley and Kessler (2018) and Bordalo et al. (2019) did not find gender differences in beliefs about ability in a verbal ability task to model female stereotypical domain. Coffman (2014), however, looks at gender differences in tasks that involved quiz questions about art & literature and entertainment & pop culture which are conjectured as more female-typed tasks. The prior differences in beliefs of women's advantage over men in these was lower than that of men over women in male-typed domain captured through sports.[22] Thus depending on the strength of the effect of the female stereotyp-

[21] For below median performers, efficiency will be determined by a trade-off between learning about one's own true rank and the probability of being hired by the partner.

[22] For instance when answers were elicited on scale of [-1,1] where -1 is 'Women know more', 1 is 'Men know more' and 0 is labelled as ' no gender difference', men's average for art is -0.317 and for women is -0.419. For sports however, men's average is 0.643 and for women is 0.571.

ical task on beliefs and if men and women continue to respond differently to them, the results in our study might be reversed.

5. Concluding remarks

In this paper, we study whether there are gender gaps in information seeking about own performance in a male stereotypical domain. We find that women are nearly 20 p.p. more likely to choose less informative feedback despite more information being incentivized. We do not find strong evidence that it is affected by adding publicness. We have weak evidence that adding judgement to the feedback might reduce the gender gap. However, we find strongest evidence that the gender gap disappears when we introduce a strategic consideration of being hired by an experimental partner. This effect is driven entirely by men who increase the take-up of the less informative information structure.

This research provides a new perspective to the literature on gender differences in economic behavior. In particular, it shows that there are gender differences in preferences for information, and that they are shaped by the strategic environment in which it is sought. This, in turn, affects how much someone learns about past performance. This is not without consequences. In fact, this may affect future performance levels through different channels. For example, getting information about your own performance can help you in taking corrective actions for improvement (e.g. higher effort). But it can also help in keeping your ego intact or manage your self-confidence.

Future research should enrich the current analysis at least in two dimensions apart from what we mentioned above. First, new research should more closely study the (potentially) many implications of differences in information-seeking behavior by gender. Specially, in male dominated environments in which gender differences in beliefs and outcomes (e.g. wages) are higher. Second, while we provide a careful analysis on how the environment shapes information preferences by gender, we believe that a step forward would be to look at these preferences in real-world settings. Indeed, this may prove particularly fruitful in terms of understanding how the publicness of information shape preferences. In real-world scenarios publicness of the feedback may play a more decisive role compared to what we find here.

Declaration of Competing Interest

None.

Appendix A. Screenshots of the experiment

Here we provide the screenshots for T2. Screenshots for the other experimental treatments are available upon request.

A1. Welcome page

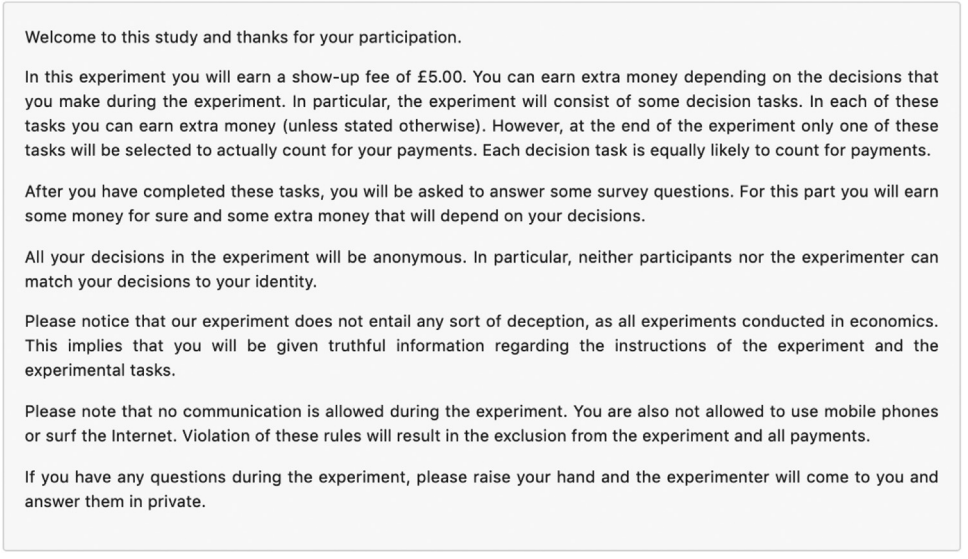


Fig. A.1. Screenshot of the Welcome Page Template.

A2. Instructions – the quiz

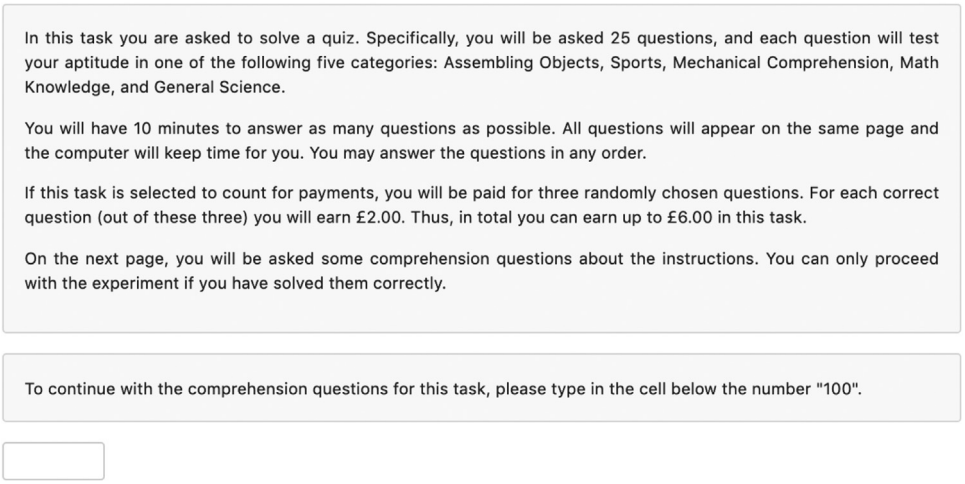


Fig. A.2. Screenshot of the Instructions Template for the Quiz.

A3. Comprehension questionnaire – the quiz

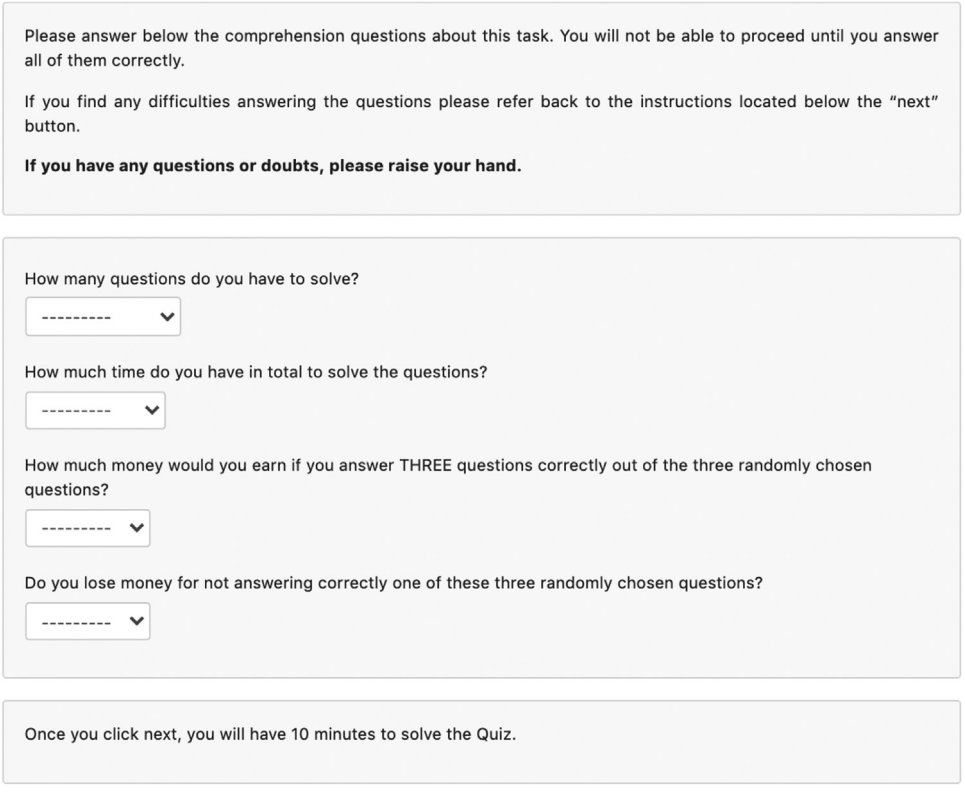


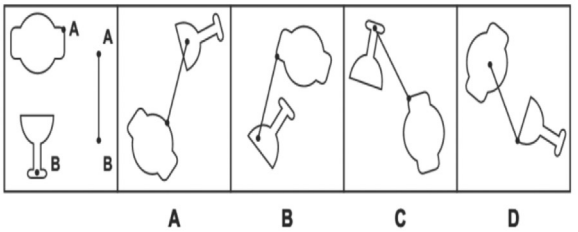
Fig. A.3. Screenshot of the Comprehension Questionnaire Template for the Quiz.

A4. The quiz

Please, answer the questions below. You have 10 minutes (600 seconds) to answer as many questions as possible.

Category: Assembling Objects

Question 1. Which figure best shows how the objects in the left box will touch if the letters for each object are matched?



Question 1 Answer:

Question 2. Which figure best shows how the objects in the left box will appear if they are fit together?

Fig. A.4. Screenshot of the Quiz.

A5. Instructions – Belief elicitation

In the next screen, you will be asked to assess your own performance in the Quiz as compared to other people's performance in this lab session.

That is, your task is to guess whether your performance (the number of correctly solved questions) in the Quiz is in the top half of the distribution of these participants. Ties are broken randomly. In particular, we will ask you to state the probability with which you think that your score is in the top half (that is top 50%) among these participants.

Specifically, your score will be compared to all other participants' performances in today's experiment, except for the score of one randomly chosen participant. The score of this participant will not be part of the distribution.

If this task is selected for payment, you have the opportunity to earn £2.00 or £6.00 depending on your reported belief about you being in the top half and whether you are actually in the top half.

The payment mechanism is such that reporting your true belief about your performance will maximize your chance of earning the higher payment of £6.00. (If you would like to understand why it is in your best interest to tell us your true belief, please click on the button "Detailed Explanation" that you can find below the next button).

To continue please type in the cell below the number "30".

Fig. A.5. Screenshot of the Instructions Template for the Prior Belief Elicitation.

A6. Prior belief elicitation

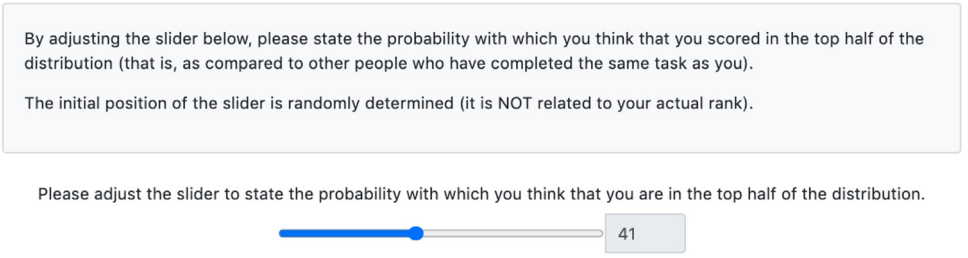


Fig. A.6. Screenshot of the Prior Belief Elicitation.

A7. Instructions – Information structure choice (T1)

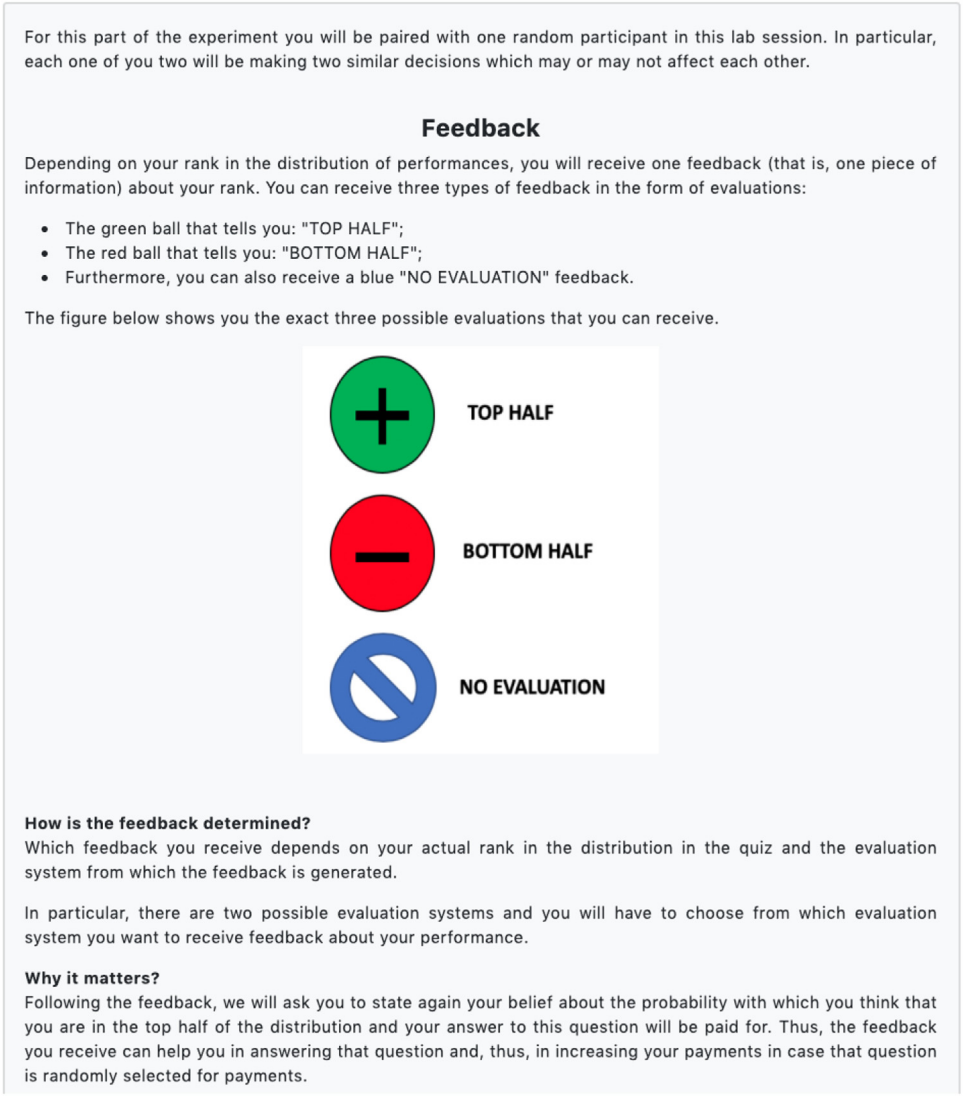


Fig. A.7. Screenshot of the Instructions Template for the Information Structure Choice – Part I.

The different components of the evaluation systems are as follows:

1. Precision of evaluation

You will always receive your feedback with some probability of error. This probability of error depends on the precision of the evaluation system. In particular, the precision of the Public evaluation system is 90%, while that of the Private evaluation system is 60%. This means that the information regarding your rank that you receive from the Public evaluation system is correct 90% of the times, while it is correct 60% of times for the private evaluation system.

To better understand how the precision works, please refer to the picture below. If you are in the top half and you choose the public evaluation system, then you will receive the green (correct) evaluation 9 out of 10 times. Similarly, if you are in the bottom half, you will receive the red (correct) evaluation 9 out of 10 times. Thus, in each case you will get the wrong evaluation 1 out of 10 times. If you instead choose the private evaluation system and you are in the top half then you will receive the green (correct) evaluation 6 out of 10 times. Similarly, if you are in the bottom half, you will receive the red (correct) evaluation 6 out of 10 times. Thus, in each case you will get the wrong evaluation 4 out of 10 times.

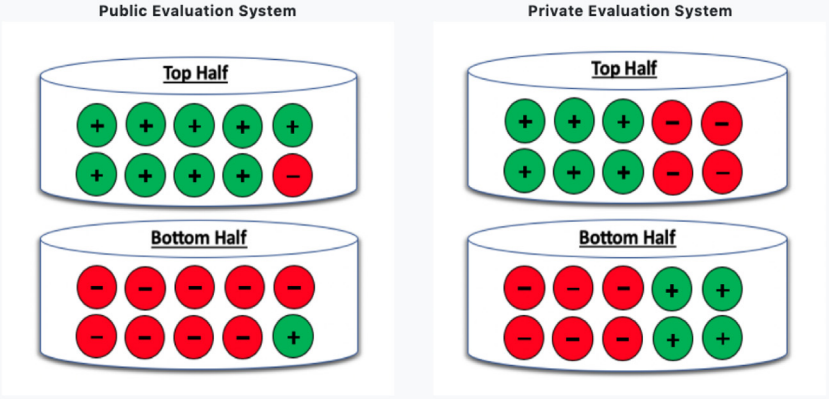


Fig. A.8. Screenshot of the Instructions Template for the Information Structure Choice – Part II.

2. Mode of evaluation

In the Public evaluation system your evaluation will also be shown to the participant you are paired with in this task. That is, if you get a green evaluation, your partner will also get to see this evaluation about your performance. Similarly, if you get a red evaluation, your partner will get to see it too. In the private evaluation system, instead, the evaluation will be privately communicated to you. Therefore, the participant you are paired with will not get to see your evaluation. In fact, regardless of the evaluation that you receive, he/she will see the following picture:

Your partner's feedback

Importantly, while you are making all these choices, the participant you are matched with will be making the same choices. Thus, if he/she chooses the Public evaluation system, then you will get to see his/her evaluation. If, instead, he/she opts for the Private evaluation system, then you will not see his/her feedback.

Summary

In sum, the Public evaluation system is more informative about your actual performance than the Private evaluation system since it has higher precision. Also, in contrast to the Private evaluation system, your paired participant gets to see your feedback in the public evaluation system (i.e., the evaluation you receive).

Similarly, the chosen evaluation system by your partner will determine whether you see his/her feedback.

To continue with the comprehension questions for this part of task 2, please type in the cell below the number "40".

Fig. A.9. Screenshot of the Instructions Template for the Information Structure Choice – Part III.

A8. Comprehension questions – Information structure choice

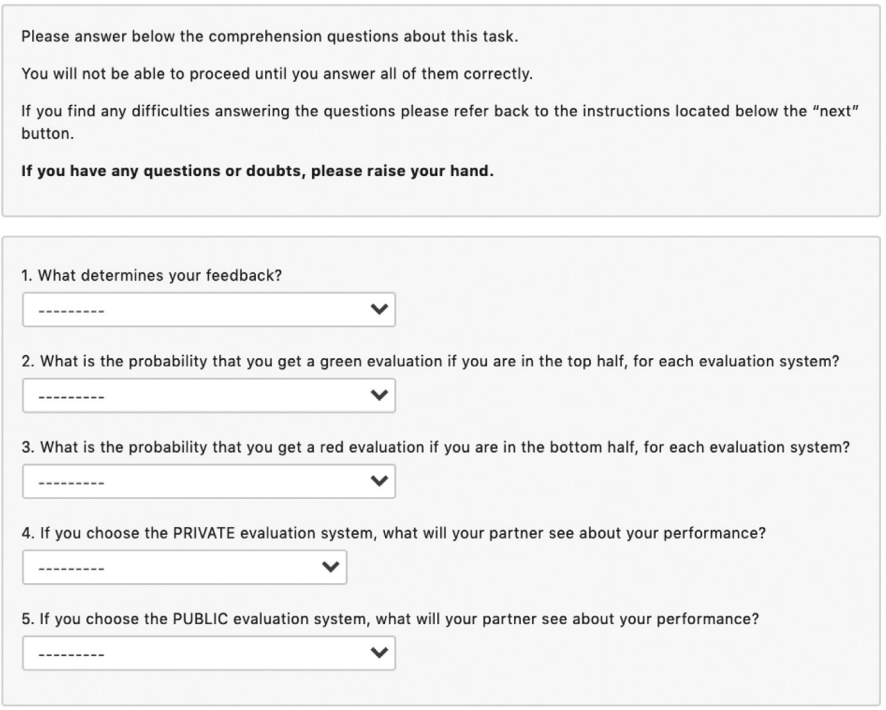


Fig. A.10. Screenshot of the Comprehension Questionnaire Template for the Information Structure Choice.

A9. Information structure choice

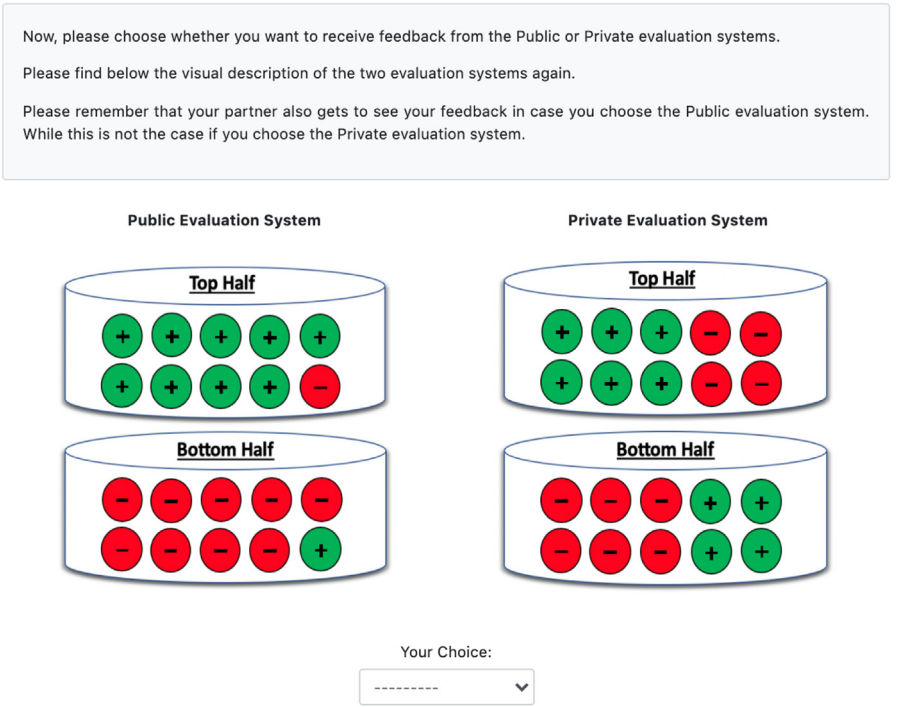


Fig. A.11. Screenshot of the Information Structure Choice.

A10. Partner'S feedback

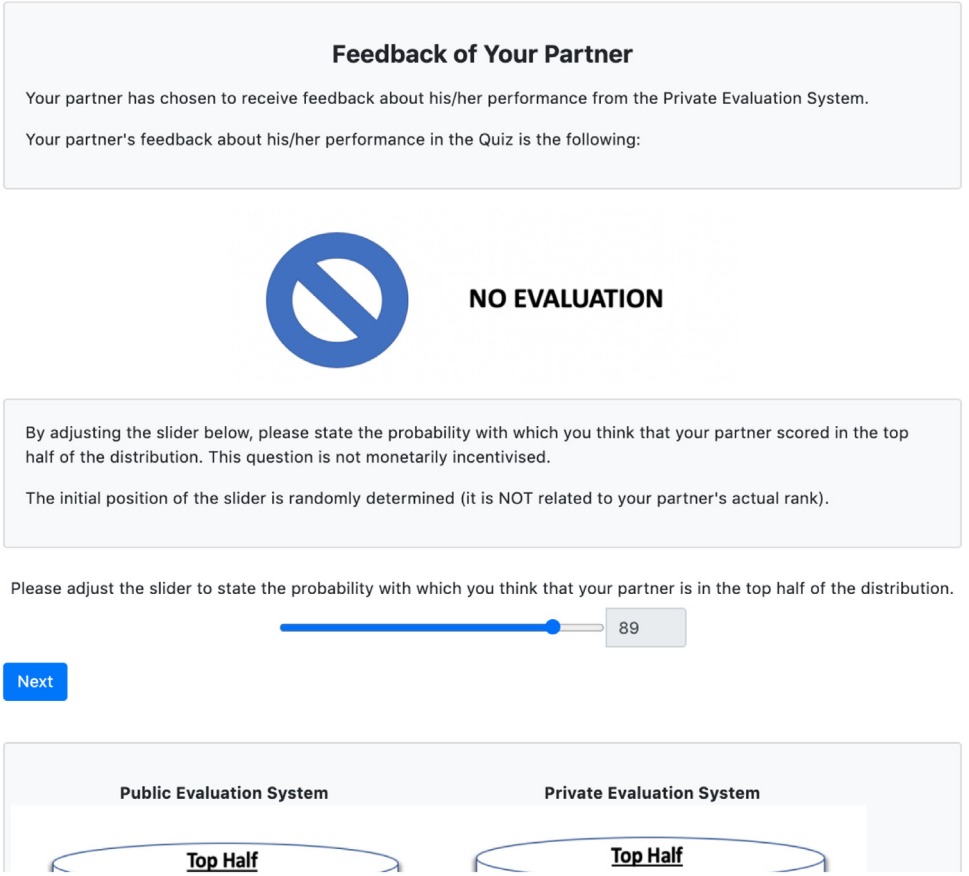


Fig. A.12. Screenshot of the Feedback the Subject Received about her Partner.

A11. Posterior belief

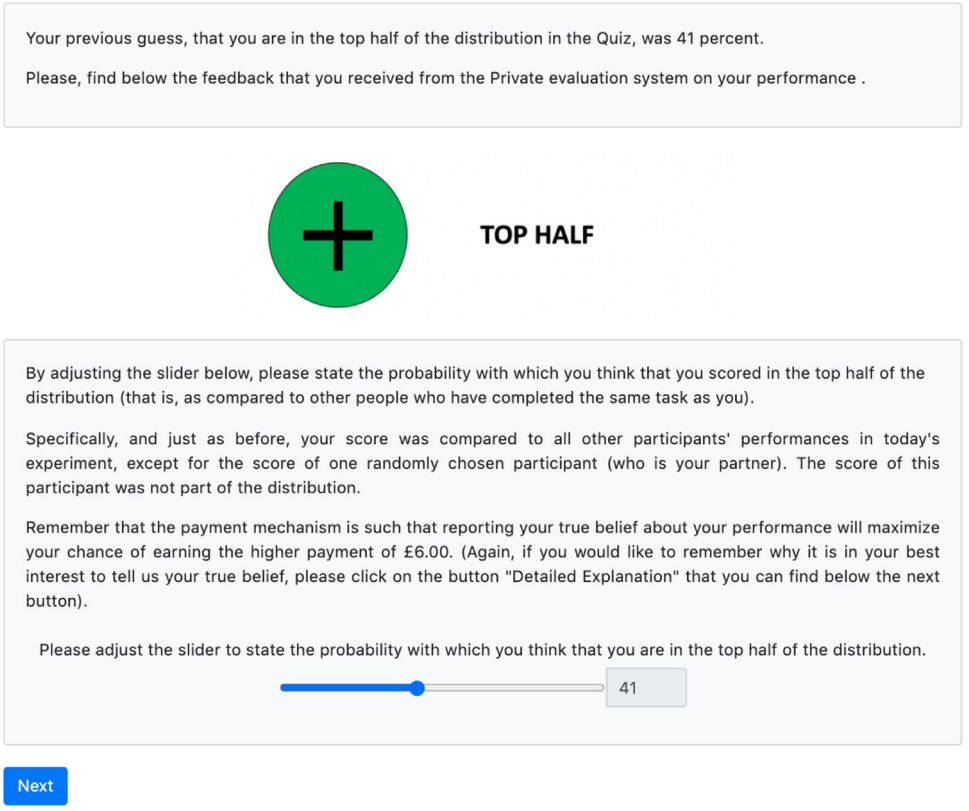


Fig. A.13. Screenshot of the Feedback Received and Elicitation of Posterior Belief.

A12. Instructions for strategic treatment (T4)

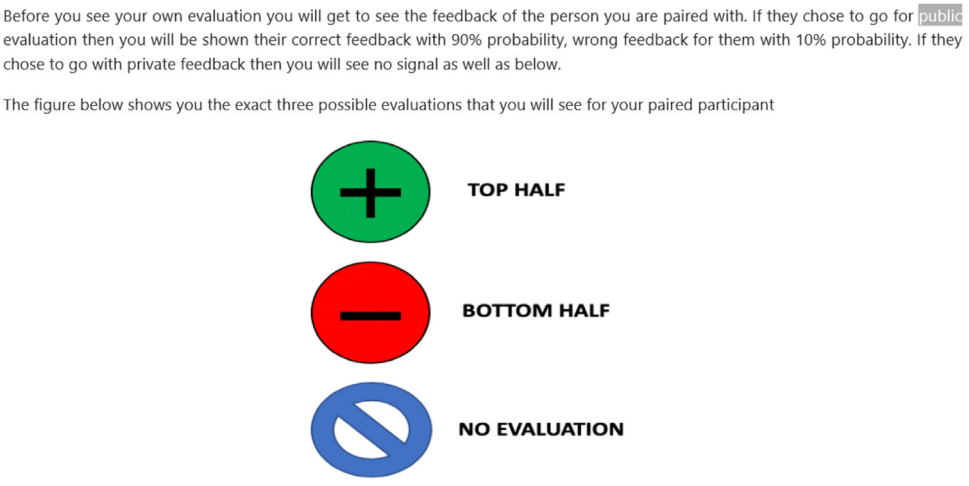
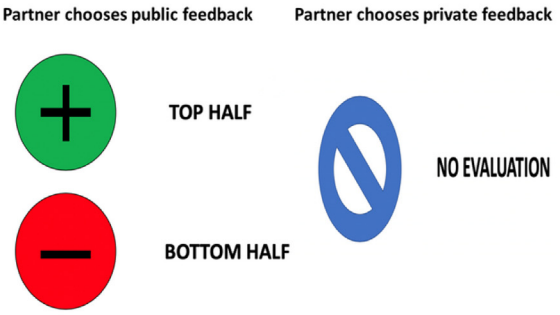


Fig. A.14. Screenshot of information about payments for Strategic Treatment.

A13. Instructions for strategic treatment (T4) continued

The figure below shows you what kind of feedback you will see depending on the evaluation system chosen by your paired participant.



Once you get to see their feedback, you have to make the decision on whether you would like to **partner** with them. On the basis of your decision to partner with your paired participant will affect your payoffs in the following manner in this part of the experiment. In particular:

- If you decide to partner with them and they are really in the top half of the distribution, then you will be paid **£6.**
- If you decide to partner with them and they are really in the bottom half of the distribution, then you will be paid **0.**
- If you decide to not partner with them and instead go solo, then you will be paid **£2.5.**

Fig. A.15. Screenshot of information about payments for Strategic Treatment.

A14. Instructions for strategic treatment (T4) continued

Thus, it is in your best economic interest to partner with him/her if you think your paired participant is in the top half, whereas it is best for you not to partner with him/her if he/she is in the bottom half on the basis of the feedback you see for them. Please note whether you get to see their feedback or not depends on whether you paired participant chose for public or private feedback.

Importantly, the score of your randomly matched participant was not part of the distribution of scores your performance has been compared to. This implies that whether he/she is in the top half does not imply that you are less likely to be in the top half. Similarly, if he/she is in the bottom half, you are not more likely to be in the top half.

Your paired participant's Hiring Decision

Similarly, the participant you are paired with will also make a partnering decision about you. And before he/she makes the decision, they will also get to see your evaluation if you chose the public evaluation system. They will also see no feedback if you chose private evaluation system. This is to help them take the decision for partnering with you in the same way as we show you their feedback to help you make the same partnering decision.

Importantly, if your paired participant decides to partner you, you will earn **£2.5** irrespective of your rank (and if this part of the task is randomly selected to count for you earnings). If not, you will earn 0. Please note though, your chances of being chosen for partnering with them are better if they think that you are in the top half of distribution. Their partnering incentives are just like yours.

Thus, it is in your best economic interest that your paired participant decides to "partner" with you.

The following are your payoffs if your paired participant decides to partner with you

Your decision to partner or not with them	Your partner is truly in the top half	Your partner is truly in the bottom half
Partner with the paired participant	£6	0
Do not partner with the paired participant	£2.5	£2.5

Fig. A.16. Screenshot of information about payments for Strategic Treatment.

A15. Distribution of questions answered correctly

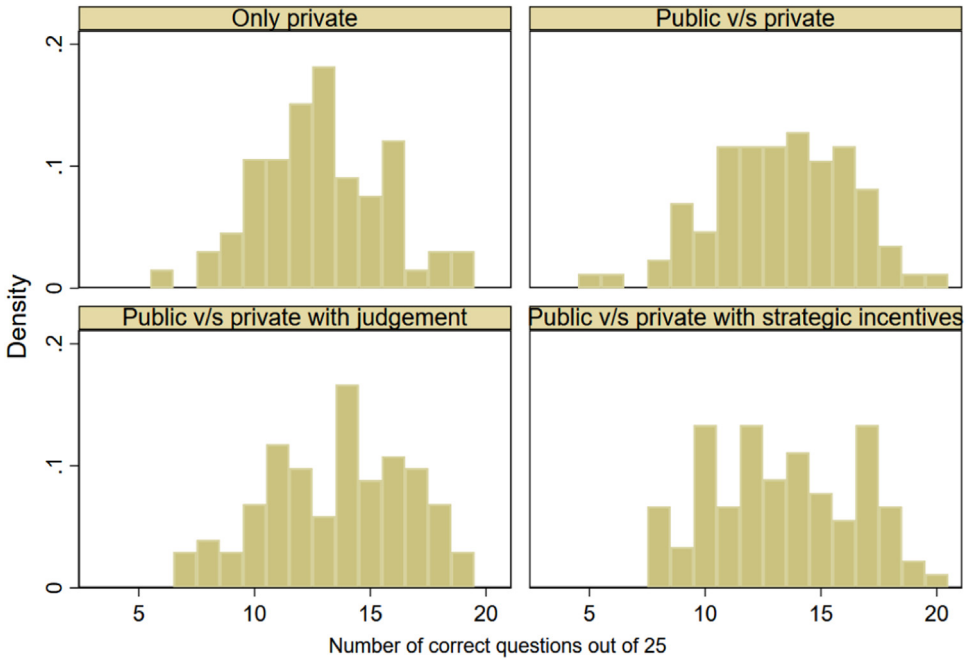


Fig. A.17. Distribution of number of questions answered correctly by treatment. *Notes:* The graph provides the distribution of the number of questions that subjects answered correctly by the treatment they were assigned to.

A16. Qualitative answers

Fig. A.18. Qualitative answers to reasons behind choosing less informative feedback. *Notes:* The graph provides the different reasons for which men and women chose less informative feedback out of those who chose less informative feedback.

A17. Efficiency versus equity

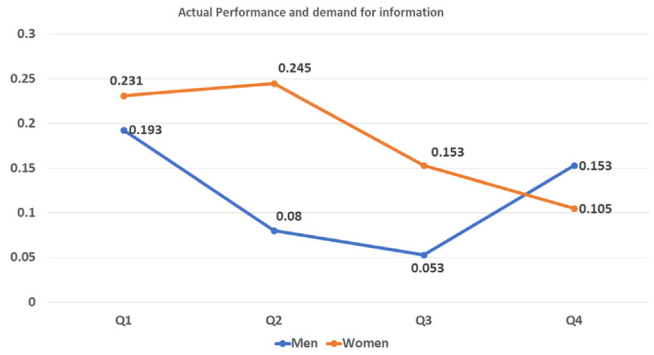


Fig. A.19. Efficiency versus equity by performance quartiles. *Notes:* The graph provides percentage of men and women who take-up less informative feedback in T4. Y axis provides the percentage that choose information structure B and X-axis depicts the quartiles of performance for men and women.

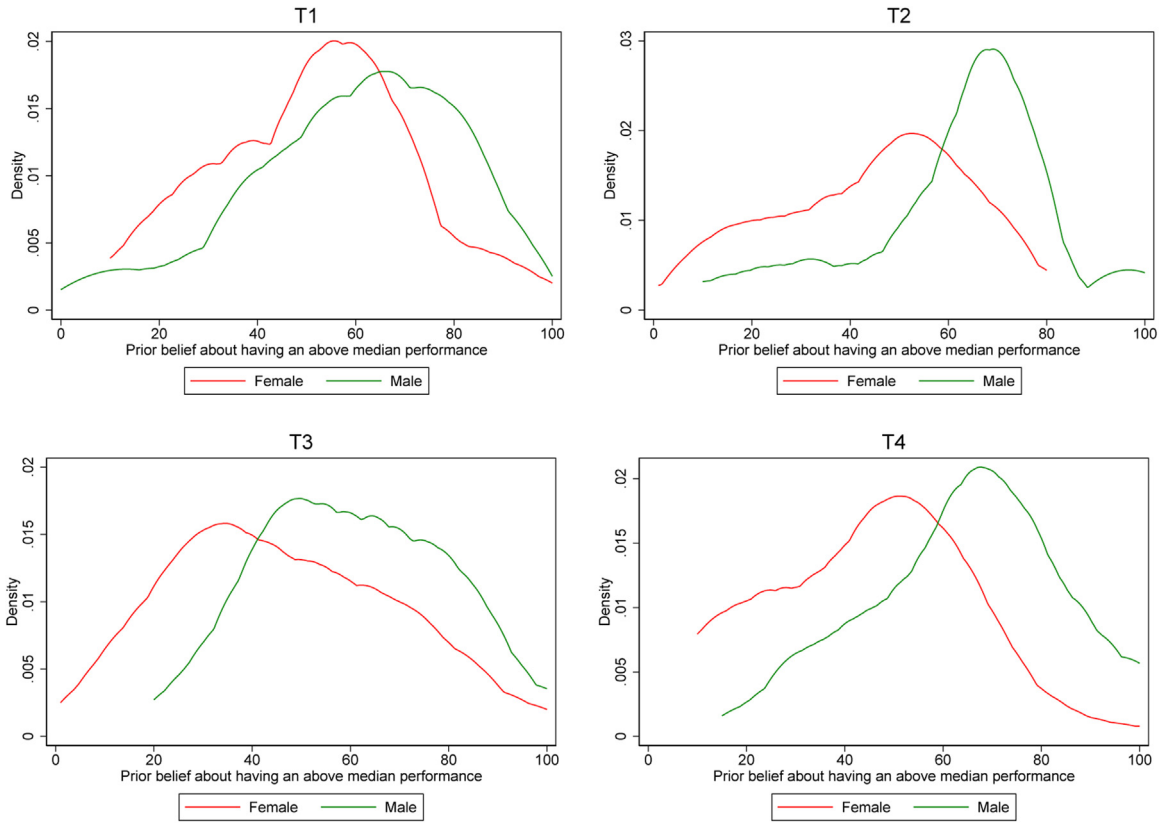


Fig. A.20. Distribution of priors by gender and treatment. *Notes:* Kernel density plots depicting distribution of priors by gender and treatment.

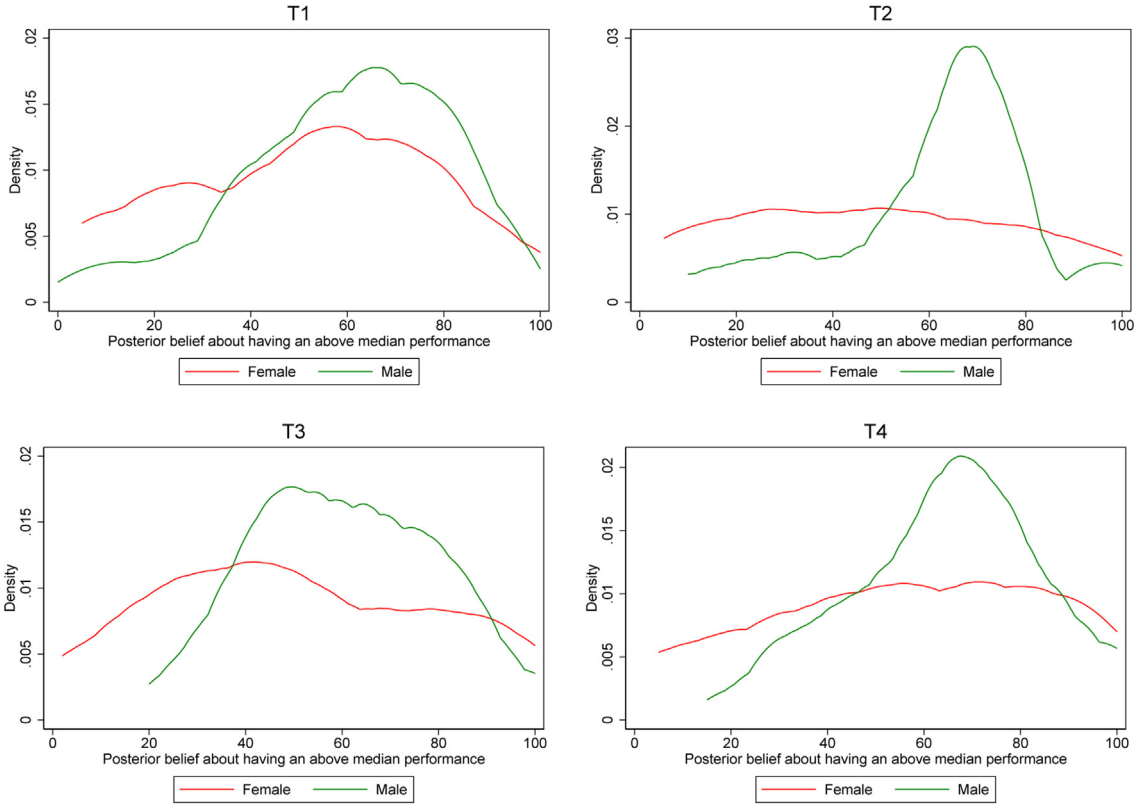


Fig. A.21. Distribution of posteriors by gender and treatment. *Notes:* Kernel density plots depicting distribution of posteriors by gender and treatment.

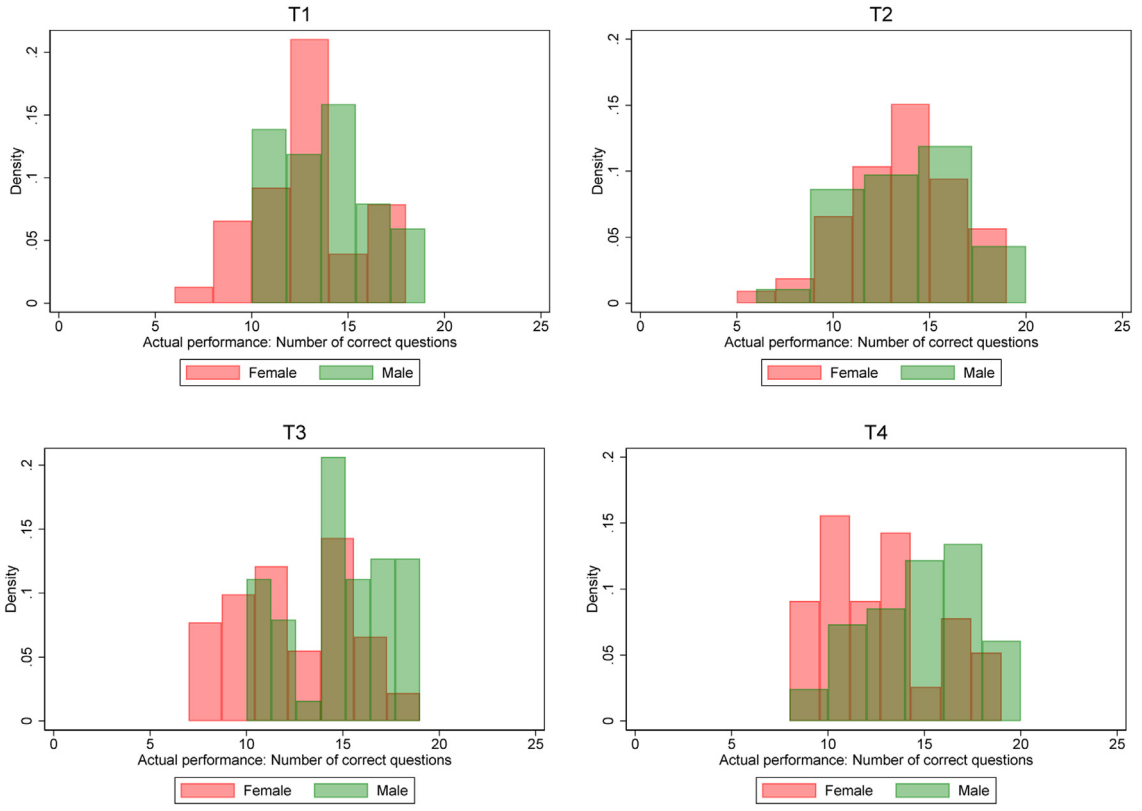


Fig. A.22. Distribution of performance by gender and treatment. *Notes:* Histogram of performance in the quiz by gender and treatment.

Appendix B. Summary statistics

B1. Demographic questions

The questions we asked are:

- What is your gender?(Male, Female, Other)
- What is your age? (Text answer)
- What is your mother tongue? (Text answer)
- What is your country of origin? (Text answer)
- What is your ethnicity? (Text answer)
- What is your year of study? (1st, 2nd, 3rd, 4th, +5 years, Other (I am not a student))
- In which department are you currently enrolled? Text answer
- Have you taken part in experiments before? (0,1,2,3,4, +4 experiments)
- Do you know one or more participants in the experiment today? (0,1,2,3,4,+4 participants)
- How willing are you to take risks? (0 (not willing at all), 1,2,3,4,5,6,7,8,9,10 (very willing))

Table B.1
Descriptive statistics.

	T1	T2	T3	T4
Female (Share)	0.57	0.61	0.52	0.54
	(0.49)	(0.48)	(0.50)	(0.50)
Age (Mean)	21.59	21.27	21.14	21.65
	(3.65)	(4.88)	(4.22)	(4.11)
Year of study	1.47	1.56	1.59	1.62
	(0.50)	(0.50)	(0.49)	(0.49)
Quantitative Education (Share)	0.69	0.61	0.63	0.63
	(0.46)	(0.48)	(0.48)	(0.48)
Risk Preferences (Mean)	5.93	6.15	5.65	5.76
	(2.08)	(1.86)	(1.83)	(1.94)
Risk choice (Mean)	48.13	50.98	51.79	45.01
	(31.82)	(30.91)	(29.70)	(30.59)
Score (Mean)	12.87	13.27	13.54	13.37
	(2.72)	(2.96)	(3.05)	(3.13)
Prior (Mean)	54.77	51.26	53.37	53.53
	(21.65)	(21.98)	(22.51)	(22.82)
Experience with experiments (Share)	0.50	0.52	0.50	0.46
	(0.50)	(0.50)	(0.50)	(0.50)
Friends in the experiment (Share)	0.60	0.22	0.27	0.16
	(0.49)	(0.41)	(0.44)	(0.37)
Number of subjects (per session)	13.20	14.33	14.57	15.00
	(2.68)	(1.96)	(2.22)	(2.09)
N	66	86	102	90

Notes: the table shows descriptive statistics (in means or in shares) of our experimental subjects by treatment. Female is the share of female subjects. Age is the reported age of the subject. Quantitative education is a dummy variable equal to 1 if the subject's course of study is mainly quantitative, year of study is the year of undergraduate study they are in if they are an undergraduate. Risk preferences and Risk choice are two variables that capture subjects' risk preferences. The first is the Dohmen et al. (2011) risk elicitation question, while the second is the Gneezy and Potters (1997) risky choice question. Score is the subject's score in the quiz and Prior is the subject's prior belief about her relative performance in the quiz. Experience with experiments is a dummy variable that equals 1 if the subject has taken part to at least 3 other experiments in the past. The Friends in the experiment dummy variable is equal to 1 if the subject knows at least 1 other participant in her session. Standard deviations are in parentheses.

Table B.2
Descriptive statistics by information structure chosen.

	Information structure A	Information structure B
Female (Share)	0.55	0.69
	(0.49)	(0.46)
Age (Mean)	21.33	21.71
	(4.35)	(3.77)
Year of study	1.57	1.51
	(0.49)	(0.50)
Quantitative Education (Share)	0.69	0.61
	(0.46)	(0.48)
Risk Preferences (Mean)	5.86	5.87
	(1.90)	(2.01)
Risk choice (Mean)	49.17	48.82
	(31.18)	(27.93)
Score (Mean)	12.87	13.27
	(2.72)	(2.96)
Prior (Mean)	53.70	50.35
	(22.90)	(23.27)
Experience with experiments (Share)	1.50	1.48
	(0.50)	(0.50)
Friends in the experiment (Share)	0.29	0.30
	(0.45)	(0.46)
N	288	56

Notes: the table shows descriptive statistics (in means or in shares) of our experimental subjects by the information structure they chose to get the feedback from for all treatments together. Female is the share of female subjects. Age is the reported age of the subject. Quantitative education is a dummy variable equal to 1 if the subject's course of study is mainly quantitative, year of study is the year of undergraduate study they are in if they are an undergraduate. Risk preferences and Risk choice are two variables that capture subjects' risk preferences. The first is the Dohmen et al. (2011) risk elicitation question, while the second is the Gneezy and Potters (1997) risky choice question. Score is the subject's score in the quiz and Prior is the subject's prior belief about her relative performance in the quiz. Experience with experiments is a dummy variable that equals 1 if the subject has taken part to at least 3 other experiments in the past. The Friends in the experiment dummy variable is equal to 1 if the subject knows at least 1 other participant in her session. Standard deviations are in parentheses.

Table B.3
Gender differences by treatment.

	Average Men	Average Women	All
Prior			
(T1) Private Feedback	0.59	0.51	0.54
(T2) Public Feedback without Judgment	0.61	0.44	0.51
(T3) Public Feedback with Judgment	0.60	0.46	0.53
(T4) Public Feedback with Hiring Decision	0.64	0.44	0.53
Posterior			
(T1) Private Feedback	0.57	0.50	0.53
(T2) Public Feedback without Judgment	0.54	0.48	0.51
(T3) Public Feedback with Judgment	0.58	0.50	0.54
(T4) Public Feedback with Hiring Decision	0.53	0.55	0.54
Performance			
(T1) Private Feedback	13	12	12
(T2) Public Feedback without Judgment	13	13	13
(T3) Public Feedback with Judgment	14	12	13
(T4) Public Feedback with Hiring Decision	14	12	13

Notes: The table shows average differences between men and women in their stated priors, posteriors (incentivized), and their actual performance (that is the number of correct questions solved, incentivized). The differences are shown by treatment. Priors and posteriors are depicted as means to the question: 'What is the likelihood with which you think that your performance is in the top half of the performance distribution'.

Table B.4
Choice of Information Structure B in (T1)-(T3).

	OLS			Probit		
	(1) Info B	(2) Info B	(3) Info B	(4) Info B	(5) Info B	(6) Info B
Female	0.201**	0.197**	0.197**	0.213**	0.210**	0.203**
	(0.078)	(0.081)	(0.084)	(0.090)	(0.091)	(0.081)
T2	0.025	0.026	0.066	0.055	0.057	0.114
	(0.055)	(0.055)	(0.061)	(0.129)	(0.129)	(0.132)
T3	0.066	0.067	0.092	0.118	0.121	0.149
	(0.056)	(0.056)	(0.062)	(0.117)	(0.117)	(0.111)
T2 × Female	-0.092	-0.097	-0.126	-0.087	-0.092	-0.097
	(0.103)	(0.102)	(0.105)	(0.094)	(0.090)	(0.062)
T3 × Female	-0.114	-0.118	-0.112	-0.115	-0.118	-0.096
	(0.105)	(0.104)	(0.103)	(0.073)	(0.071)	(0.054)
Prior Belief		✓	✓		✓	✓
Demographics			✓			✓
F-test $\beta_1 + \beta_4 = 0$	0.02	0.02	0.09	0.03	0.04	0.06
F-test $\beta_1 + \beta_5 = 0$	0.17	0.24	0.11	0.19	0.28	0.06
R-Squared	0.036	0.037	0.104	0.049	0.049	0.160
N	254	254	254	254	254	254

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include:age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.5
Choice of Feedback Mode B in (T1) with session fixed effects.

	OLS		
	(1) Info B	(2) Info B	(3) Info B
Female	0.164*	0.182*	0.217*
	(0.092)	(0.098)	(0.127)
Prior Belief		✓	✓
Demographics			✓
Session FE	✓	✓	✓
R-Squared	0.098	0.117	0.214
N	66	66	66

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.6
Choice of Information Structure B in (T1)-(T3) with session fixed effects.

	OLS		
	(1) Info B	(2) Info B	(3) Info B
Female	0.1644	0.1599	0.1724
	(0.0918)*	(0.0935)*	(0.0951)*
T2	0.0195	0.0160	0.1118
	(0.1139)	(0.1148)	(0.1231)
T3	0.1150	0.1131	0.1201
	(0.1425)	(0.1433)	(0.1469)
T2 × Female	-0.0659	-0.0713	-0.1205
	(0.1117)	(0.1104)	(0.1120)
T3 × Female	-0.0924	-0.0967	-0.1162
	(0.1249)	(0.1242)	(0.1212)
Prior Belief		✓	✓
Demographics			✓
Session FE	✓	✓	✓
R-Squared	0.0671	0.0682	0.1323
N	254	254	254

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include:age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.7
Overall difference in take up of less informative feedback, all treatments with session fixed effects.

	OLS		
	(1) Info B	(2) Info B	(3) Info B
Female	0.1644*	0.1608*	0.1715*
	(0.0917)	(0.0930)	(0.0949)
T2	0.0195	0.0167	0.0720
	(0.1139)	(0.1145)	(0.1185)
T3	0.1150	0.1135	0.1127
	(0.1424)	(0.1430)	(0.1523)
T4	0.1229	0.1257	0.1476
	(0.1380)	(0.1384)	(0.1457)
T2 × Female	-0.0659	-0.0702	-0.1036
	(0.1116)	(0.1106)	(0.1123)
T3 × Female	-0.0924	-0.0958	-0.1216
	(0.1248)	(0.1243)	(0.1236)
T4 × Female	-0.1830	-0.1886	-0.2299*
	(0.1291)	(0.1290)	(0.1318)
Prior Belief		✓	✓
Demographics			✓
Session FE	✓	✓	✓
F-test $\beta_1 + \beta_7 = 0$	0.837	0.765	0.537
R-Squared	0.057	0.058	0.099
N	344	344	344

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.8
Choice of Feedback Mode B in (T1) with standard errors clustered at the session level.

	OLS			Probit		
	(1) Info B	(2) Info B	(3) Info B	(4) Info B	(5) Info B	(6) Info B
Female	0.2011*	0.2202*	0.2387	0.201**	0.210***	0.175*
	(0.0851)	(0.0891)	(0.1369)	(0.084)	(0.081)	(0.095)
Prior Belief		✓	✓		✓	✓
Demographics			✓			✓
WC bootstrap p values	0.0625	0.0625	0.1875	0.0625	0.0625	0.0625
R-Squared	0.077	0.098	0.205	0.105	0.131	0.313
N	66	66	66	66	66	66

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Standard errors are clustered at the session level. *p*-values from wild cluster bootstrap is reported in the table. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.9
Choice of Information Structure B in (T1)-(T3) with clustered standard errors.

	OLS			Probit		
	(1) Info B	(2) Info B	(3) Info B	(4) Info B	(5) Info B	(6) Info B
Female	0.201**	0.197**	0.197**	0.213**	0.210**	0.203**
	(0.079)	(0.080)	(0.088)	(0.010)	(0.104)	(0.093)
T2	0.025	0.026	0.066	0.055	0.057	0.114
	(0.050)	(0.050)	(0.047)	(0.123)	(0.123)	(0.108)
T3	0.066	0.067	0.092	0.118	0.121	0.149
	(0.058)	(0.059)	(0.070)	(0.125)	(0.125)	(0.118)
T2 × Female	-0.092	-0.097	-0.126	-0.087	-0.092	-0.097
	(0.089)	(0.088)	(0.093)	(0.093)	(0.091)	(0.060)
T3 × Female	-0.114	-0.118	-0.112	-0.115	-0.118	-0.096
	(0.099)	(0.104)	(0.101)	(0.099)	(0.081)	(0.061)
Prior Belief		✓	✓		✓	✓
Demographics			✓			✓
WC bootstrap *p*-values	0.10	0.12	0.13	0.0781	0.100	0.085
R-Squared	0.036	0.037	0.104	0.049	0.049	0.160
N	254	254	254	254	254	254

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include:age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Standard errors are clustered at the session level and reported in parentheses. Wild cluster bootstrap p values are reported as well. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.10
Overall difference in take-up of less informative feedback, all treatments with clustered standard errors.

	OLS			Probit		
	(1) Info B	(2) Info B	(3) Info B	(4) Info B	(5) Info B	(6) Info B
Female	0.201**	0.198**	0.2011**	0.237**	0.235**	0.242**
	(0.078)	(0.080)	(0.089)	(0.115)	(0.115)	(0.111)
T2	0.025	0.026	0.040	0.062	0.064	0.088
	(0.049)	(0.049)	(0.050)	(0.141)	(0.142)	(0.132)
T3	0.066	0.067	0.082	0.138	0.139	0.175
	(0.058)	(0.058)	(0.066)	(0.149)	(0.150)	(0.149)
T4	0.184***	0.186***	0.198***	0.294*	0.299**	0.356**
	(0.044)	(0.045)	(0.052)	(0.159)	(0.152)	(0.148)
T2 × Female	-0.092	-0.096	-0.118	-0.094	-0.098	-0.111
	(0.103)	(0.102)	(0.104)	(0.099)	(0.096)	(0.073)
T3 × Female	-0.114	-0.117	-0.110	-0.124*	-0.127	-0.126**
	(0.105)	(0.104)	(0.104)	(0.082)	(0.081)	(0.065)
T4 × Female	-0.196*	-0.202*	-0.227*	-0.162***	-0.164***	-0.173***
	(0.119)	(0.118)	(0.121)	(0.051)	(0.049)	(0.038)
Prior Belief		✓	✓		✓	✓
Demographics			✓			✓
WC bootstrap p-values						
β_1	0.07	0.09	0.16	0.07	0.10	0.11
β_4	0.10	0.11	0.11	0.07	0.06	0.05
β_7	0.03	0.026	0.02	0.02	0.02	0.01
$\beta_1 + \beta_7$	0.90	0.96	0.71	0.92	0.94	0.67
F-test $\beta_1 + \beta_7 = 0$	0.932	0.948	0.726	0.936	0.916	0.626
R-Squared	0.036	0.036	0.365	0.042	0.043	0.103
N	344	344	344	344	344	344

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Standard errors clustered at the session level are reported in parentheses. Wild cluster bootstrap p values are added for the key coefficient estimates. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.11
Gender difference in responding to priors for T1 only.

	(1) Info B	(2) Info B	(3) Info B	(4) Info B
Female	0.220**	-0.103	0.238**	-0.241
	(0.085)	(0.152)	(0.115)	(0.192)
Prior	0.003	-0.001	0.002	-0.004*
	(0.002)	(0.001)	(0.002)	(0.002)
Prior_gender		0.006*		0.009**
		(0.003)		(0.003)
Demographics			✓	✓
N	66	66	66	66
F-test (Prior+PriorXGender=0)		0.080		0.059

The results above are depicted for subjects in T1 only. The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

Table B.12
Gender difference in responding to priors for T2 and T3 only.

	(1) Info B	(2) Info B	(3) Info B	(4) Info B
Female	0.06	0.31**	0.05	0.30**
	(0.04)	(0.12)	(0.05)	(0.12)
Prior	-0.001	0.001	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Prior_gender		-0.01**		-0.01**
		(0.002)		(0.002)
Demographics			✓	✓
N	188	188	188	188
F-test (Prior+PriorXGender=0)		0.06		0.06

The results above are depicted for subjects for T2 and T3 only. The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses. Asterisks denote significance: $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

References

Alan, S., Ertac, S., Kubilay, E., Loranth, G., 2020. Understanding gender differences in leadership. Econ. J. 130 (626), 263–289.
Alan, S., Ertac, S., Mumcu, I., 2018. Gender stereotypes in the classroom and effects on achievement. Rev. Econ. Stat. 100 (5), 876–890.
Altonji, J.G., Blank, R.M., 1999. Race and gender in the labor market. Handb. Labor Econ. 3, 3143–3259.
Ambuehl, S., Li, S., 2018. Belief updating and the demand for information. Games Econ. Behav. 109, 21–39.
Azmat, G., Iriberri, N., 2010. The importance of relative performance feedback information: evidence from a natural experiment using high school students. J. Public Econ. 94 (7–8), 435–452.
Azmat, G., Petrongolo, B., 2014. Gender and the labor market: what have we learned from field and lab experiments? Labour Econ. 30, 32–40.
Banerjee, R., Gupta, N.D., Villeval, M.C., 2020. Feedback spillovers across tasks, self-confidence and competitiveness. Games Econ. Behav. 123, 127–170.
Barber, B.M., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Q. J. Econ. 116 (1), 261–292.
Blanes i Vidal, J., Nossol, M., 2011. Tournaments without prizes: evidence from personnel records. Manage. Sci. 57 (10), 1721–1736.
Blau, F.D., Kahn, L.M., 2017. The gender wage gap: extent, trends, and explanations. J. Econ. Lit. 55 (3), 789–865.
Block, R.A., Harper, D.R., 1991. Overconfidence in estimation: testing the anchoring-and-adjustment hypothesis. Organ. Behav. Hum. Decis. Process. 49 (2), 188–207.
Bohnet, I., Van Geen, A., Bazerman, M., 2016. When performance trumps gender bias: joint vs. separate evaluation. Manage. Sci. 62 (5), 1225–1234.
Bordalo, P., Coffman, K., Gennaioli, N., Shleifer, A., 2019. Beliefs about gender. Am. Econ. Rev. 109 (3), 739–773.
Buser, T., Niederle, M., Oosterbeek, H., 2014. Gender, competitiveness, and career choices. Q. J. Econ. 129 (3), 1409–1447.
Carlana, M., 2019. Implicit stereotypes: evidence from teachers' gender bias. Q. J. Econ. 134 (3), 1163–1224.
Castagnetti, A., Schmacker, R., 2020. Protecting the Ego: Motivated Information Selection. Working Paper.
Charness, G., Gneezy, U., 2012. Strong evidence for gender differences in risk taking. J. Econ. Behav. Organ. 83 (1), 50–58.
Charness, G., Oprea, R., Yuksel, S., 2018. How do People Choose Between Biased Information Sources? Evidence from a Laboratory Experiment. Working Paper.
Chen, D.L., Schonger, M., Wickens, C., 2016. oTree – an open-source platform for laboratory, online, and field experiments. J. Behav. Exp. Finance 9, 88–97.
Coffman, K.B., 2014. Evidence on self-stereotyping and the contribution of ideas. Q. J. Econ. 129 (4), 1625–1660.
Coffman, K.B., Collis, M., Kulkarni, L., 2019. Stereotypes and Belief Updating. Working Paper.
Coffman, K.B., Collis, M., Kulkarni, L., 2019. When to Apply? Working Paper. Harvard Business School.
Croson, R., Gneezy, U., 2009. Gender differences in preferences. J. Econ. Lit. 47 (2), 448–474.
Danz, D., Wilson, A.J., Vesterlund, L., 2020. Belief Elicitation: Limiting truth Telling with Information or Incentives. CESifo Working Paper.
Dobrescu, L.I., Faravelli, M., Megalokonomou, R., Motta, A., 2021. Relative performance feedback in education: evidence from a randomised controlled trial. Econ. J. 131 (640), 3145–3181.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: measurement, determinants, and behavioral consequences. J. Eur. Econ. Assoc. 9 (3), 522–550.
Eil, D., Rao, J.M., 2011. The good news-bad news effect: asymmetric processing of objective information about yourself. Am. Econ. J. Microecon. 3 (2), 114–138.
Exley, C.L., Kessler, J.B., 2018. Motivated Errors. Technical Report.
Exley, C.L., Kessler, J.B., 2019. The Gender Gap in Self-Promotion. Technical Report. National Bureau of Economic Research.
Falk, A., Kosse, F., Schildberg-Hörisch, H., Zimmermann, F., 2020. Self-assessment: The Role of the Social Environment. Working Paper.
Falk, A., Zimmermann, F., 2016. Beliefs and Utility: Experimental Evidence on Preferences for Information. IZA Discussion Paper No. 10172.
Flory, J.A., Leibbrandt, A., List, J.A., 2015. Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. Rev. Econ. Stud. 82 (1), 122–155.
Ganguly, A., Tasoff, J., 2016. Fantasy and dread: the demand for information and the consumption utility of the future. Manage. Sci. 63 (12), 4037–4060.
Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. Q. J. Econ. 112 (2), 631–645.
Goldin, C., 2014. A grand gender convergence: its last chapter. Am. Econ. Rev. 104 (4), 1091–1119.
Golman, R., Hagmann, D., Loewenstein, G., 2017. Information avoidance. J Econ Lit 55 (1), 96–135.
Jones, T.M., Fanson, K.V., Lanfear, R., Symonds, M.R., Higgie, M., 2014. Gender differences in conference presentations: aconsequence of self-selection? PeerJ 2, e627.
Karlsson, N., Loewenstein, G., Seppi, D., 2009. The ostrich effect: selective attention to information. J. Risk Uncertain. 38 (2), 95–115.
Karni, E., 2009. A mechanism for eliciting probabilities. Econometrica 77 (2), 603–606.
Masatlioglu, Y., Orhun, A.Y., Raymond, C., 2017. Intrinsic Information Preferences and Skewness. Working Paper.

Michelmore, K., Sassler, S., 2016. Explaining the gender wage gap in stem: does field sex composition matter? RSF Russell Sage Foundation J.Soc. Sci. 2 (4), 194–215.
Milkman, K., Akinola, M., Chugh, D., 2013. Discrimination in the Academy: A Field Experiment. SSRN, Working Paper.
Möbius, M.M., Niederle, M., Niehaus, P., Rosenblat, T.S., 2022. Managing self-confidence: theory and experimental evidence. Manage. Sci..
Montanari, G., Nunnari, S., 2019. Audi Alteram Partem: An Experiment on Selective Exposure to Information. Working Paper.
Moore, D.A., Healy, P.J., 2008. The trouble with overconfidence. Psychol. Rev. 115 (2), 502.
Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Q. J. Econ. 122 (3), 1067–1101.
Nielsen, K., 2018. Preferences for the Resolution of Uncertainty and the Timing of Information. Working Paper.
Oster, E., Shoulson, I., Dorsey, E., 2013. Limited life expectancy, human capital and health investments. Am. Econ. Rev. 103 (5), 1977–2002.
Sicherman, N., Loewenstein, G., Seppi, D.J., Utkus, S.P., 2015. Financial attention. Rev. Financ. Stud. 29 (4), 863–897.
Solda, A., Ke, C., Page, L., Von Hippel, W., 2020. Strategically delusional. Exp. Econ. 23 (3), 604–631.
Zimmermann, F., 2014. Clumped or piecewise? Evidence on preferences for information. Manage. Sci. 61 (4), 740–753.