ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Entering a gender-neutral workplace? College students' expectations and the impact of information provision

Francesca Barigozzi
José J. Domínguez
Natalia Montinari

Quaderni - Working Paper DSE N°1188

Entering a gender-neutral workplace? College students' expectations and the impact of information provision*

Francesca Barigozzi†, José J. Domínguez‡ and Natalia Montinari§

Abstract: Although women often outperform men in school and college, they still face higher unemployment rates and lower wages when employed. Are prospective workers aware of these inequalities, or do they expect to enter a gender-neutral workplace? This paper investigates college students' expectations and the effect of information provision about gender gaps in academic performance and early labor market outcomes on the two sides of the labor market. Our lab experiment comprises a questionnaire to elicit students' beliefs about academic performance and labor market inequalities, a demand-side game, and a supply-side game. In the demand-side game, subjects act as employers and are asked to hire three candidates and assign them to tasks that differ in complexity and profitability. In the supply-side game, we elicit individual willingness to compete. Information provision takes the form of feedback on the elicited beliefs. Our treatments vary the timing of the feedback: subjects in the feedback treatment received feedback before facing the other two games, while subjects in the priming and the control treatments only received feedback at the end of the experiment. First, our findings indicate that participants are largely unaware of gender gaps. Second, while information provision doesn't substantially alter employers' hiring decisions, it increases the likelihood of assigning women to challenging tasks. Third, while feedback enhances willingness to compete among job market candidates, it does not significantly alter the gender gap in competitiveness. Overall, our experiment suggests potential positive effects of information provision on women's labor market outcomes.

Keywords: Gender Gaps, Expectations, Information provision, Competition, Hiring, Task assignment.
JEL Codes: D03, C91, J71.

*We thank Federico Lenzi and Francesca Madeddu for their excellent research assistance in preparing the Questionnaire and the Feedback treatment. We are grateful to the participants at the ESA meeting 2022 for their useful comments and suggestions. Francesca Barigozzi and Natalia Montinari acknowledge funding from the University of Bologna RFO.
†University of Bologna, Department of Economics, Piazza Scaravilli 2, 40126 Bologna, Italy. E-mail: francesca.barigozzi@unibo.it.
‡ University of Granada, Department of Economic Theory and History, Campus Universitario La Cartuja, 18071 Granada (Spain). E-mail: jjdominguez@ugr.es.
§University of Bologna, Department of Economics, Piazza Scaravilli 2, 40126 Bologna, Italy. E-mail: natalia.montinari2@unibo.it.

Non-technical abstract

The study focuses on understanding college students' awareness of gender gaps in academic performance and early labor market outcomes and the influence of information provision on their labor market decisions. Although in many countries women often outperform men in school and college, they still face higher unemployment rates and lower wages when employed. Are prospective workers aware of these inequalities, or do they expect to enter a gender-neutral workplace? We study college students' expectations and the effect of information provision about gender gaps in academic performance and early labor market outcomes on their behavior in two sides of the labor market. To this end we design a lab experiment composed by three parts: i) a questionnaire to elicit students' beliefs about gender differences academic performance and labor market inequalities after one and five years from graduation: ii) a demand-side game, and iii) a supply-side game. In the demand-side game, all participants act as employers and are asked to hire three candidates from a pool and then to assign them to tasks that differ in complexity and profitability. The aim of this game is to analyze the impact of information provision on gender stereotypes about the ability of male and females candidates to perform a math based real effort task. In the supply-side game, we elicit individual willingness to compete using the experimental protocol introduced by Niederle and Vesterlund (2012) where participants have to decide whether to compete or not on a math-based task. Information provision takes the form of feedback on the elicited beliefs. Our treatments vary the timing of the feedback: subjects in the feedback treatment received feedback after the belief elicitation and before facing the other two games, while subjects in the priming and the control treatments only received feedback about the correctness of their beliefs at the end of the experiment. The findings revealed that most students lacked awareness of gender gaps in academic performance and early labor market outcomes. This unawareness persists despite recent media coverage of gender disparities. Women, particularly in Engineering, demonstrated a higher awareness of these gaps. Regarding behavior impact, providing information did not influence employers' hiring decisions but increased the likelihood of assigning female candidates to challenging tasks. This effect was attributed to information highlighting women's superior academic performance. On the supply side, information provision increased the willingness to compete for both genders, with no change in the gender gap in competitiveness. The information served to counter biases and boost confidence, especially among women, without reducing their competitiveness. Results from our study suggest that providing information about gender gaps does not prevent women from competing but instead empowers them with awareness, potentially influencing their career trajectories. This insight holds significant implications for policy and practice, highlighting the importance of accurate information dissemination in addressing gender inequalities in the labor market. Overall, this paper offers a nuanced understanding of the interplay between information, beliefs, and behavior in the context of gender gaps in labor market outcomes, providing valuable insights for future interventions and policies aimed at fostering a more equitable labor market.

1. Introduction

In most countries, girls outperform boys in school (Salvi del Pero and Bytchkova, 2013). Later in life, women also attain higher levels of education than men, on average, and perform at least as well as men in college. Nevertheless, women are still less likely than men to be employed, and when they are employed, they earn less (Bertrand, 2020). An important question is whether prospective workers are aware of these figures or, conversely, they expect to enter a gender-neutral workplace.

Workers' expectations largely influence labor market outcomes (Kim et al., 2019; Balleer et al., 2023; Mueller and Spinnewijn, 2023). Specifically, success in the workplace is influenced by investments and signals related to productivity, which prospective workers undertake based on their expectations of future job prospects. However, it remains uncertain to what extent the gender gap in labor market outcomes can be attributed to gender differences in beliefs. As an example, women's decision to opt out of competitive job selection procedures might be rational if they believe a gender gap in employment exists. The extensive media debate that ensued after movements like #MeToo# and the extended media coverage regarding the evidence of heightened gender gaps in the labor market due to the Covid-19 pandemic has highlighted the presence of gender inequalities in the last years. Then, we wonder whether workers about to enter the labor market (e.g., college students) hold correct expectations of early gender inequalities in the labor market and whether those expectations affect their labor market outcomes.

Suppose that college students' expectations are not accurate. Would providing information be an appropriate intervention? There is no evidence on whether information provision on gender gaps can impact the behavior of men and women in the labor market, nor is it clear if such an intervention would be desirable. In theory, accurate information about future career prospects enables job seekers to make more efficient choices (Chen and He, 2021). For instance, having accurate information empowers job candidates to decide if and to what extent they should compete with their peers during job selection procedures and throughout their employment. However, enlightening women about persistent gender gaps in labor market outcomes carries the risk of triggering a self-fulfilling prophecy. This could potentially lead women to feel discouraged and ultimately accept the prevailing unequal labor market conditions, inadvertently perpetuating the status quo. Accurate information may also help employers make more efficient choices when hiring and assigning workers to different tasks. Specifically, information on academic performance might increase allocation efficiency and simultaneously suggest that gaps in labor market outcomes are not driven by a gap in academic ability. Overall, it is important to study how information provision affects labor market behavior and whether it is an effective intervention to reduce the gender gap in labor market outcomes.

In this study, we design a laboratory experiment on 312 college students to examine i) expectations of early gender gaps in academic performance and labor market outcomes and ii) the impact of information provision on labor demand and

supply. Notably, our information provision offers both positive and negative evidence. Female students receive positive news about their relative college performance, while negative news pertains to early gender gaps in market outcomes. The reverse applies to males. Overall, the information provided conveys that early labor market disparities aren't due to academic ability differences.

Our research aims to address the following questions: Do individuals on the brink of entering the labor market possess accurate beliefs regarding academic performance and job market opportunities? If not, what kinds of biases are present? Does information about the gap in academic performance and early gender gaps in market outcomes influence employers and job market candidates? And if it does, does this information improve women's prospects in the labor market?

We start by collecting students' expectations on gender gaps in academic performance and labor market outcomes. These outcomes include the early employment rate, wages one year and five years after graduation, and the probability of self-employment among graduates in the respondents' same field of study. We provide field-specific gender gap information to students in Economics and Management, Engineering, and Law using data from students who graduated from the University of Bologna (Italy) a few years before. These fields differ in male/female enrollment ratios, early employment, and wage gaps. The elicitation task is non-incentivized.

Next, we implement two different games to analyze the effect of information provision on students' behavior. Note that the gender gap in early-career outcomes results from the interaction of employers and job seekers, i.e., the interaction of demand and supply. For example, employers might be less likely to select women, especially as leaders (Reuben et al., 2014), and women might refrain from competitive environments (Buser et al., 2007). Hence, it is important to study the impact of information on the two sides of the labor market. Our Demand-side game exploits the task assignment game designed by Domínguez and Montinari (2021). All participants assume the role of employers and must choose three workers from a mixed-gender pool of candidates. Among these candidates, two are designated for a simple task (such as summing up three-digit numbers). At the same time, one is meant for a challenging task involving mathematical problem-solving (such as quantitative multiple-choice questions). Employers' earnings are contingent on the performance of the workers they select. These employers are provided with information about the candidates' age, gender, and field of study and receive a performance-related signal. We employ this game to discern how employers make hiring and task assignment decisions on a gender basis. The Supply-side game follows the approach of Niederle and Vesterlund (2007). Participants engage in the same simple task as earlier (summing up three-digit numbers). Here, they choose between receiving compensation through a non-competitive (low) piece rate or opting for a competitive tournament with a potentially higher payoff. This game is utilized to determine if information plays a role in influencing gender-related disparities in willingness to compete.

To assess the effects of information provision, we implemented three distinct experimental treatments. These treatments varied in two key aspects: first, the timing of eliciting participants' beliefs about labor market inequalities (to identify potential de-biasing mechanisms), and second, the point at which information was provided—either immediately following belief elicitation or at the study's conclusion. This structured comparison across treatments enables us to derive causal inference regarding the influence of information about gender disparities. Specifically, we examine its impact on individuals' decisions to engage in competitive scenarios (addressed in the Supply-side game) and on their choices related to hiring and task allocation (examined in the Demand-side game). Additionally, by accounting for participants' pre-existing awareness of gender gaps and implicit biases, as measured by the Implicit Association Test (IAT), we aim to discern whether our intervention's effectiveness stems from a de-biasing mechanism or merely from the provision of information.

We have derived three key sets of results from our study. In relation to awareness of early gender gaps, we have found that most participants lack awareness of gender gaps in academic performance and labor market outcomes. Despite widespread media coverage, it appears that many Italian college students still hold misconceptions about early gender gaps in the labor market. While these students may be aware of gender disparities among more established members of the workforce, our results document that they ignore the early emergence of these gaps, which occur as soon as one year after graduation. Women tend to exhibit a higher awareness of early gender gaps in labor market outcomes than men, especially in Engineering.

Exploring the impact of information on the behavior of employers and job market candidates, we observe the following. For employers (Demand-side game), the information does not influence hiring decisions. In other words, considering all candidates assigned to both tasks, the proportion of women hired remains consistent between the treatment and control groups. However, information does increase the likelihood of female candidates being assigned to the Challenging Task. This effect appears to be driven by information highlighting women’s superior academic performance.

For job market candidates (Supply-side game), we find that information provision raises willingness to compete for both males and females so that the gender gap in competitiveness does not change. This outcome is driven by individuals who are already aware of gender gaps and women who hold implicit gender biases. In essence, our feedback treatment serves as a mechanism for countering biases and boosting confidence among female participants. As mentioned before, the feedback treatment also positively influences men’s willingness to compete, particularly among those aware of gender gaps. This phenomenon suggests that men naturally lean toward competition, particularly after being 'reminded' of women's superior college performance. Importantly, information provision does not diminish women’s willingness to compete. Hence, our experiment does not support concerns about the potential negative consequences of providing information about existing gender inequalities in the labor market.

The paper is organized as follows. Section 2 describes the related literature. Section 3 explains the experiment design and execution. Section 4 derives our three sets of results. Section 5 discusses the mechanisms driving our results. Section 6 offers policy implications.

2. Literature review

A couple of empirical papers used survey data to elicit students' expectations of labor market outcomes and are thus related to the first part of our lab experiment aimed at eliciting students' awareness about gender gaps. Filippin and Ichino (2005) investigated whether male and female college students hold similar expectations about future wages and whether the gender wage gap is correctly anticipated. They explore the extent to which second-year students at Bocconi University, a prestigious Business and Management private college in Milan (Italy), anticipate the gender wage gap. In addition, wage expectations of (N= 1154) students are matched with the actual wages of a sample of similar Bocconi graduates. The authors find that the gender gap implied by students' expectations one year after graduation is consistent with the gender gap implied by the earnings of their elder counterparts. There is, instead, a misperception of the gender gap later in the career after graduation because students expect the gender gap to be roughly constant, while realizations indicate an increasing gap with experience.

Our elicitation of students' awareness about the existence of a gender wage gap is similar to Filippin and Ichino's "explicit expected gap" elicitation because we both ask students whether they think a man or a woman with their same background is more likely to earn a given wage level. While Filippin and Ichino elicit expectations about the gender wage gap also indirectly by comparing personal wage expectations with the actual wage of similar graduates, we elicit students' expectations about other gender gaps such as academic performance and unemployment rate for graduates in their same field of study.

Only partially related to our paper is Sipe et al. (2009), who elicited students' expectations of gender discrimination in the labor market, instead of students' expectations about gender gaps in labor market outcomes as we do. The authors analyze survey data (N= 1,373) collected among undergraduate students enrolled in a Business major in the College of Business Administration in a large public university in the USA. Their findings suggest that students underestimate the extent and consequences of gender discrimination and tend to consider themselves immune to gender discrimination.

Several experimental studies have focused on the possible mechanisms that explain the gender gap in employment and wages from both sides of the labor market. Among them, we report below the papers that somehow inspired the design of the Demand-side and Supply-side games of our lab experiment.

On the supply side, gender differences in taste for competition have been hypothesized as one of the potential predictors of labor market outcomes (Buser et al., 2014; Dohmen and Falk, 2011; Heinz et al., 2016). If women have a stronger aversion to competitive workplaces than men of similar ability, the low share of women participating in competitive sectors could explain the gender gap in labor market success (Niederle and Vesterlund, 2007; Flory et al., 2014). The Supply-side game of our experiment is drawn from those papers, especially Niederle & Vesterlund (2007).

Career achievements do not, however, depend only on the supply side of the labor market. On the demand side, factors such as employers' gender beliefs and discrimination can also partly explain wage gaps. Notwithstanding equal performance by female and male workers, employers' gender biases may result in the assignment of less profitable activities to women in the organization (De Pater et al., 2010). The experimental evidence on task assignment shows that discrimination against women operates not only with respect to the possibility of promotion in the vertical hierarchy, women being less likely to be selected as team leaders (Reuben et al., 2012; Peterle & Rau, 2017), but also horizontally, having less likelihood of being assigned to high-promotability tasks within the organization, compared to men (Babcock et al., 2017a; von Bieberstein et al., 2020).

In a recent lab experiment, Domínguez and Montinari (2021) investigate how gender quotas in hiring affect the allocation of workers into different tasks within an organization. Participants in the role of employers were asked to hire a team of six workers from a pool of 15 and assign them to one of two tasks, which differed in complexity and profitability. Employers had information about workers' age, gender, and university major as well as a signal of performance. Workers assigned to the Simple Task had to complete additions, while workers assigned to the Challenging Task had to solve mathematical problems. Even if the focus and objective of our paper is different from Domínguez and Montinari (2021), our Demand-side game is inspired by their setting and differentiates between simple and challenging tasks.

3. Experimental Design and Procedures

We conducted a laboratory experiment comprising three distinct components: 1) a Questionnaire (Q), 2) a Demand-side game (D), and 3) a Supply-side game (S). In the following paragraphs, 3.1, 3.2, and 3.3, we provide a detailed description of each part. Section 3.4 outlines the treatments used in the experiment, Section 3.5 elucidates the feedback content provided, and Section 3.6 presents the experimental procedures.

3.1. Belief Elicitation: Questionnaire (Q)

Participants, who were students from three schools—Economics and Management, Law, and Engineering—were requested to complete a concise questionnaire. They were encouraged to consider their responses carefully. Each respondent in the questionnaire was asked to provide their expectations regarding performance disparities within their specific school and the labor market for graduates from that particular school. Specifically, the questionnaire mentioned

three hypothetical individuals: "*Giovanni, Angela, and Hassan"* who *"just graduated from the School of Economics and Management (*or *Engineering,* or *Law) of the University of Bologna.*" Respondents were initially asked to estimate the probability of each individual achieving a graduation grade equal to or higher than a specific threshold, as indicated in Question A:

- **Question A:** "*Based on your own experience and information, what is the probability that Giovanni (*then *Angela,* and *Hassan) graduates with a grade at least as high as 105 (*or *106,* or *102) out of 110?"*

The grade mentioned in Question A corresponds to the *average graduation grade* achieved by graduates from one of the Master' programs in the field of Economics and Management offered by the University of Bologna in the years 2010, 2011, and 2012. The average graduation grades for Engineering and Law are provided in parentheses.[1] Importantly, the questions in the questionnaire were field-dependent, meaning that subjects answered questions specifically concerning individuals who graduated in the same field of study as themselves.

We chose the names 'Giovanni' and 'Angela' as representative of typical Italian names, reflecting a Caucasian man and woman, respectively. In addition, we introduced the name 'Hassan' to represent an Arab man. This was done to reduce the emphasis on gender when presenting the questions. The objective of this question was to gauge respondents' beliefs regarding the presence of a gender gap in academic performance within their field of study, operating under the assumption that they may possess better knowledge about other students who attended the same master's program. So, we compare the probabilities assigned to *Angela* and *Giovanni* in each field of study.

The subsequent section of the questionnaire aims to elicit expectations regarding gaps in labor market outcomes following graduation for the same three hypothetical individuals. Specifically, six questions are presented, following the same structure as Question A. Each question begins with the following information: "*Giovanni/Angela/Hassan just graduated from the School of Economics and Management/Engineering/Law at the University of Bologna with a grade of 105/106/102 out of 110.*" Participants were then asked to express their expectations by assigning probabilities to specific labor market outcomes relevant to their field of study. As an illustration, we provide the questions with threshold wages that are applicable to graduates from the School of Economics and Management. These wage figures were extracted from the wage distribution of graduates in the three fields of study who graduated between 2010 and 2012,

[1] The University of Bologna provides a wide range of Master's programs, including eighteen programs in the field of Economics and Management, three programs in the field of Law, and twenty-eight programs in Engineering. At the University of Bologna, the average graduation grade for Engineering is 106/110, while for Law, it is 102/110. These average graduation grades have been computed based on data collected from individuals who graduated in 2010, 2021, and 2012.

one year and five years after graduation. We obtained this data from AlmaLaurea (https://www.almalaurea.it/), the statistical office of Italian public universities.

In your view, which is the

- *probability that Giovanni (*then *Angela* and *Hassan) will be employed on a permanent contract?* **Question B1 (permanent contract).**
- *probability that Giovanni (*then *Angela* and *Hassan) will earn 1170 Euro per month one year after graduation?* **Question B2 (median wage after 1 year).** [2]
- *probability that Giovanni (*then *Angela* and *Hassan) will earn 1420 Euro per month one year after graduation?* **Question B3 (high-stake wage after 1 year).**
- *probability that Giovanni (*then *Angela* and *Hassan) will earn 1460 Euro per month five years after graduation?* **Question B4 (median wage after 5 years).**
- *probability that Giovanni (*then *Angela* and *Hassan) will earn 1760 Euro per month five years after graduation?* **Question B5 (high-stake wage after 5 years).**
- *probability that Giovanni (*then *Angela* and *Hassan) will be self-employed five years after graduation?* **Question B6 (self-employment after 5 years).**

Participants received a fixed payment of 10 Euros upon completing the questionnaire, regardless of the correctness of their answers. We decided to offer payment for the questionnaire without incentivizing their beliefs to minimize experimental demand. We were concerned that if participants were explicitly incentivized, they might infer an expectation of differential treatment for Angela and Hassan compared to Giovanni. By not linking payment to their beliefs, we aimed to mitigate potential biases in their responses and maintain the experiment's integrity.

3.2. Demand-side Game (D)

In the demand-side game, all participants in a session assumed the role of employers for a firm. Employers were tasked with making two consecutive decisions following the approach of Domínguez and Montinari (2021). Firstly, they had to select three candidates from a pool of six to hire for their firm. Subsequently, employers had to assign the chosen candidates to two different tasks that varied in difficulty and profitability for the firm. Specifically, each employer had

[2] The wages of questions B2 and B4 correspond to the median wage of graduates from the University of Bologna with a Master's degree in the field of Economics and Management one year and five years after graduation, respectively. For graduates with a Master's degree in Engineering and Law, the corresponding wages are 1280 and 660 Euros after one year, respectively, and 1460 and 1270 Euros after five years. The high-stake wages of questions B3 and B5 represent the upper bound of the third quartile in the wage distribution of graduates with a Master's degree in Economics and Management one year and five years after graduation. For graduates with a Master's degree in Engineering and Law, the corresponding wages are 1450 and 1130 Euros after one year, respectively, and 1920 and 1560 Euros after five years.

to select two candidates for the Simple Task (ST) and one candidate for the Challenging Task (CT). The ST involved summating as many three-digit numbers as possible within a 6-minute timeframe. On the other hand, the CT required solving as many mathematical problems as possible within a 10-minute timeframe. Employers had access to information such as the candidates' age, gender, and field of study. Additionally, they received a performance signal indicating the productivity of the candidates. This performance signal was a noisy measurement based on the number of correct answers provided by the candidate during the first three minutes of the 6-minute ST.

Each employer participated in four rounds, encountering a different group of candidates in each round. The candidate profiles used in the game were derived from the candidate pools analyzed by Domínguez and Montinari (2021).[3] The pools of candidates were randomly rotated across rounds to ensure that all employers evaluated the same pool of candidates by the end of the experiment. No information was provided to employers after each round. During each round, employers were compensated based on the performance of their assigned candidates. They received 10 cents for each addition correctly solved by candidates assigned to the ST and 1 Euro for each problem correctly solved by the candidate assigned to the CT. Before the decision stage, employers participated in a non-rewarded version of the ST and CT. This allowed them to familiarize themselves with the tasks before making their decisions in the game.[4]

3.3. Supply-side Game (S)

The supply-side game comprises three stages, following the protocol by Niederle and Vesterlund (2007). Subjects were asked to perform an adding task (similar to the ST) in all stages. They had three minutes to solve as many additions of five two-digit numbers as possible. As it is well known, even if there are no gender differences in the average ability to perform this specific task, men's ability to solve mathematical tasks is generally perceived as relatively higher than women's. The payment scheme differed at each stage. In Stage 1 (*Piece rate*), subjects earned 0.50 Euro per each correct calculation. In Stage 2 (*Tournament*), subjects were allocated into groups of three male and three female students. The gender composition of the groups was common knowledge, but the identity of the group members was not revealed. The payments in Stage 2 depended on the subjects' relative performance. The tournament winners were the two subjects who correctly solved the largest number of additions. Winners received 1.50 Euro per addition correctly solved, while other subjects in the group did not receive any payment. In Stage 3 (*Choice*) of the game, subjects decided on the preferred payment scheme. They were paid 0.50 Euro per correct addition if the piece rate was selected. If the tournament was

[3] In Domínguez and Montinari (2021), 60 subjects performed ST and CT, but only 56 provided valid answers. No significant gender differences exist in those individuals' absolute performances in ST and CT. The information provided to employers in the current experiment was derived from a subsample of 24 candidates extracted from their pool.

[4] Table A1a and Table A1b in Appendix provide a detailed overview of the pool of candidates included in the groups and their performance in ST and CT. Note that the information presented in the last two columns of Table A1a is not observable to the employers who only learn a signal on the candidates' performance in the ST (corresponding to the column "Signal"). The last two columns of Table A1a, showing candidates' performance, explain the payoff received at the end of this game by employers choosing the corresponding candidate either for ST or CT. Table A1b in the Appendix shows no significant gender differences in the performances in ST and CT of those 24 individuals.

selected, the individual's performance in Stage 3 was compared to that of other members of the same group in Stage 2. Specifically, participants in the tournament won if they solved correctly at least as many additions as the winner ranking second in Stage 2. Again, winners received 1.50 Euro per calculation correctly submitted. In both Stage 2 and Stage 3, ties were broken randomly. Information on the relative performance of subjects in each group in Stages 2 and 3 was not disclosed until the end of the study. At the end of Stage 3, we elicited individual beliefs about subjects' relative ranking for all subjects and same-gender subjects in their group and gender differences in performance under tournament (Stage 2). Beliefs were incentivized. Subjects received 1 Euro for each of the three guesses if the answer was correct and 0 otherwise.

3.4. Treatments

To study the impact of feedback (F) provision on supply and demand side behaviors, we implemented three treatments varying the feedback timing: *Baseline*, *Salience*, and *Feedback*. We considered a between-subjects design where subjects did not participate in more than one treatment. In the *Baseline Treatment (BT),* the Demand-side or the Supply-side game was played first. The Questionnaire was introduced after the games to avoid potential biases in the decisions. The Feedback was provided after the Questionnaire, which also corresponded to the end of the study. In the *Feedback Treatment (FT),* the Questionnaire was administered in the first part; then, Feedback was provided immediately before implementing the S and D games. More details on the Feedback provision are provided in Section 3.5.

Note that *FT* modifies two aspects of *BT:* the timing of the Questionnaire and the timing of the Feedback provision. To control for the effect of filling out the Questionnaire (without feedback) before the decisions in the supply game and demand game, we introduced the *Salience Treatment* (*ST*), which is identical to the Feedback Treatment, except that Feedback was provided only at the end of the study. Playing the questionnaire first in ST may make the gender dimension salient, even if no Feedback is provided.

Within treatments, the order of the Demand-side and Supply-side games was randomized to ensure the consistency of the behavior in each game. Table 1 summarizes the treatments and specifies the order of their parts (Q = Questionnaire, D = Demand-side game, S = Supply-side game, F = Feedback) and the number of observations per treatment, with male and female participants being perfectly balanced.

Table 1. Treatments, sequences of games, and participants.

Treatment	*Order*	*% Female*	*N*
Baseline (BT)	D-S-Q-F S-D-Q-F	50%	72
Salience (ST)	Q-D-S-F Q-S-D-F	50%	132
Feedback (FT)	Q-F-D-S Q-F-S-D	50%	108

Note: The second column indicates the order in which the parts of each treatment were presented. D = Demand-side game; S = Supply-side game; Q = Questionnaire; F = Questionnaire feedback.

3.5. Feedback's Content

The feedback used data from three cohorts of graduates from the University of Bologna collected from Almalaurea. We calculated the exact probabilities requested in each question for both men and women.[5] The overall information is qualitatively similar for the three fields of study. In a nutshell, women like Angela outperform men like Giovanni and Hassan in college, while a gap exists in labor market outcomes in favor of men. Specifically, the average graduation grade is higher for women (Question A), but women are less likely to be employed in a permanent contract (question B1); to earn a high wage one year after graduation (questions B2 and B3), to earn a high wage five years after graduation (questions B4 and B5); to be self-employed five years after graduation (question B6). Thus, the feedback provides one piece of information representing good news for women (feedback on Question A) and another that is bad news (feedback on Questions B1-B6). And the opposite for men. Notably, the two pieces of information suggest that gender inequality in the labor market *is not* explained by differences in academic performance. Table 2 reproduces, for each question, the exact probabilities shown to the participants in the feedback stage by gender (Giovanni/Angela) and school.

[5] An example of the feedback provided to students enrolled in a master's program offered by the School of Economics and Management is reproduced in the Appendix. Interestingly, descriptive statistics indicate that those male graduates who, before enrolling in the University of Bologna, obtained a high school or a BA degree outside the European Union perform equally well in college and in the labor market as male native students. This is why, in the feedback, Giovanni and Hassan have the same overall performance.

Table 2. Summary of the information contained in the Feedback treatment by question, school, and gender.

	School of Economics and Management		School of Engineering		School of Law	
	Men	Women	Men	Women	Men	Women
Question A (average graduation grade)	104/110	107/110	105/110	106/110	99/110	103/110
Question B1 (% in permanent contract, 1year)	17.50%	17.10%	18.6%	13.6%	18.4%	16.9%
Question B2 (% above median wage, 1 year)	65%	44%	53%	42%	60%	43%
Question B3 (% above high wage, 1 year)	37%	13%	27%	20%	34%	18%
Question B4 (% above median wage, 5 years)	58%	39%	53%	40%	68%	44%
Question B5 (% above high wage, 5 years)	43%	11%	27%	18%	40%	19%
Question B6 (% self-employed, 5 years)	19%	16.6%	12.1%	11.6%	56.4 %	50.3%

Note: Those probabilities were extracted from the wage distribution of graduates in 2010, 2011, and 2012 one year and five years after graduation in the three fields of master's programs. AlmaLaura (https://www.almalaurea.it/), the statistical office of Italian public Universities, provided us with the data.

3.6. Experimental Procedures

The experiment was conducted at BLESS (the University of Bologna Experimental Laboratory) using zTree (Fischbacher, 2007). The sample consists of N=312 individuals recruited via ORSEE (Greiner, 2015) between October 2019 and April 2021.[6] Participants are students enrolled in Economics and Management, Engineering, or Law programs at the University of Bologna, either in the last year of a BA or a Master's program. The sample was divided into 15 sessions that lasted about 90 minutes each. Subjects were informed that one randomly selected part of the experiment would become relevant for payments at the end of the experiment. Average earnings were 17 Euro, including a 5 Euro show-up fee. Subjects were also asked to complete a post-experimental questionnaire and an online version of the Implicit Association Test about Science and Gender (IAT, Greenwald et al., 1998).[7] The Appendix reproduces an English version of the instructions, while the original experiment was run in Italian. We focused on the three mentioned schools to have subsamples with different gender segregations: Engineering, with most male students (75%); Economics and Management, whose students' pool is gender balanced (50%); and Law, where most students are female (64%). As displayed in Table 2, female students outperform male students in all schools but are less successful in the labor market. Table A2 in the Appendix reports summary statistics of our participants by field of study.

4. Results

In this section, we present the results of the experiment. First, we focus on the subjects' expectations about ability (Question A) and labor market inequalities (Question B1-B6). Second, we analyze the decisions in the Demand-side and the Supply-side games, focusing on the impact of our treatment. We aim to understand i) students' awareness of gender

[6] Data collection has been interrupted for about one year because of Covid-19 restrictions. To account for possible variations in beliefs and behavior around labor market prospects, we checked and observed no systematic changes before and after the pandemic. See Table A3 in the Appendix for details.

[7] We used the IAT Gender – Science, which often reveals a relative link between "Liberal arts" and "Females" and between "Science" and "Males".

gaps by gender and field of study; and whether, compared to the control, Feedback provision iia) increases the probability of success of female candidates in both hiring and task assignment (Demand-side game) and iib) increases women's willingness to compete in a math task in which men are generally perceived as better performers (Supply-side game). Moreover, we use the Questionnaire's answers to control whether information provision has a differential impact depending on the subjects' awareness of the existing gender inequalities.

4.1. Awareness of gender gaps in ability and labor market outcomes

To understand whether information provision affects subjects' behaviors in the two games, we first analyze awareness of gender gaps in our sample. To do so, we focus on replies to questions A and B1-B6.[8] We focus on the *relative magnitude* of the probabilities assigned to women (*Angela*) and men (*Giovanni*). A subject is then defined as "aware of a gender gap" if the relative magnitude of the probabilities assigned to men and women is correct according to descriptive statistics reproduced in the Feedback (see Table 2). For example, in question A (gender gap in ability), aware subjects assigned a higher probability to Angela.[9] The proportion of aware subjects for each question is summarized in Table 3, which also reports p-values from a set of Mann-Whitney tests of gender differences in accuracy for each question and school.

-

At the aggregate level, Table 3 shows that most students are unaware of the existing early gender gaps, especially in academic performance (captured by Question A). There are no significant gender differences in awareness among subjects from Economics and Management. On the contrary, there exists a significant gender difference in awareness of gaps in academic performance among Law students: men are more accurate in predicting that women obtain higher grades. Considering gaps in labor market outcomes, gender differences in awareness are detected among subjects from Engineering. Women in Engineering are more accurate than men in assessing probabilities of reaching all wages, although the difference in *high-stake wage 1* is not significant. Interestingly, more than 50 percent of men and women in Engineering predicted correctly that men and women are equally likely to be self-employed after five years from graduation. We also observe no major differences between schools with respect to awareness. Overall, women are more accurate than men in predicting early wage gaps, but this evidence is mainly driven by the high level of accuracy observed for women in Engineering.

[8] Tables A4 and A5 in the Appendix summarize the average probabilities assigned by our participants to Angela and Giovanni in each question and compare them.

[9] As a second example, for question B, the reports show that, 5 years after graduation, men in Economics and Management have a higher probability than women to earn more than 1760 Euro per month. A subject from Economics and Management is accurate in question B5 (High-stake wage 5) if she assigns a higher probability to Giovanni than Angela.

Table 3. Questionnaire: percentage of aware subjects for each question.

	Economics			Engineering			Law			Total		
Questions	Women	Men	*p*	Women	Men	*p*	Women	Men	*p*	Women	Men	*p*
A: Ability												
Ability	0.491	0.400		0.317	0.338		0.383	0.654	0.021	0.404	0.410	
B1-B6: Labour market outcomes												
Permanent contract	0.364	0.480		0.610	0.300	0.001	0.350	0.308		0.423	0.358	
Median wage 1	0.509	0.420		0.463	0.275	0.039	0.317	0.192		0.423	0.308	0.035
High-stake wage 1	0.436	0.420		0.488	0.313		0.300	0.192		0.397	0.327	
Median wage 5	0.400	0.340		0.561	0.300	0.005	0.367	0.269		0.429	0.308	0.026
High-stake wage 5	0.436	0.380		0.659	0.275	0.000	0.367	0.308		0.468	0.314	0.005
Self-employed	0.436	0.320		0.561	0.538		0.450	0.308		0.474	0.429	

Note: The table presents the percentage of aware subjects by question, field of study, and gender. In question A (gender gap in ability), aware subjects assigned a higher probability to Angela (i.e. women). In all questions B (gender gaps in labour market outcomes), aware subjects assigned a higher probability to Giovanni (i.e. men). Columns *p* report p-values from Mann-Whitney tests of gender differences in case gender differences are significant.

At the individual level, we distinguish between awareness of gender gaps in academic performance (hereafter awareness of gaps in ability) and labor market outcomes (hereafter awareness of gaps in market outcomes). Table 3 shows that 41% of the sample is aware of gender gaps in ability (question A) without significant gender differences.

As for awareness of gaps in market outcomes (questions B1-B6), on one side, we find that 92 out of 312 subjects (30%) were accurate in just one reply, while only 5% (N=16) were accurate in replies to all six questions. Sixty subjects were accurate in none of the blocks. In what follows, we define awareness of gaps in early labor market outcomes as follows: an individual is aware of gaps in labor market outcomes if they are correct in their reply to three or more questions among B1-B6.

Our results are the following. Using our definition, students aware of gaps in labor market outcomes represent 41% of the sample (N=127/312). Women are significantly more aware of gender gaps than men, 51% against 32% (Mann-Whitney test Z=3.329, p=0.001). This finding is driven by women in Engineering, who were more aware of gender gaps in market outcomes than men in the same field, 68% against 32% (Z= 3.733, p=0.001). Moreover, women in Engineering are more aware of gender gaps in market outcomes than women in Law (41%) (Z=2.618, p=0.008) and women in Economics and Management (47%), but this last difference is not statistically significant. Our findings are summarized in Result 1. Finally, it is also interesting to stress that only 20% of subjects were aware of gaps in academic performance *and* labor market outcomes, and 38% were unaware of both.

Result 1: Students' awareness. *Only 41% of our participants are aware of gender gaps in academic performance, and the same proportion is aware of early gender gaps in labor market outcomes. Women exhibit a higher awareness of gender gaps in labor market outcomes than men, particularly when enrolled in Engineering.*

4.2. Demand-side game: hiring and task assignment decisions under information

In this section, we analyze the effect of Feedback Provision on the probability of female candidates being hired and assigned to the CT (i.e. the most complex and profitable task). In other words, we show how information affects the gender gap in employment and leadership. Figure 1 summarizes the proportion of male and female candidates hired and assigned to CT within the treatment variation. Overall, women represent 38% of the candidates selected by employers to be part of their firm and 29% of the candidates assigned to the CT.

Inspection of Figure 1 reveals that the proportion of women among hired candidates does not differ significantly across treatments, meaning that feedback does not impact hiring decisions.[10] For this reason, in what follows, we focus on assignment decisions. We find that the proportion of female candidates assigned to the CT is significantly higher in *Feedback* Treatment (32%) than in *Benchmark* Treatment (25%) (Mann-Whitney Z=-2.007, p=0.044),[11] while the share of women assigned to the CT in *Salience* Treatment (29%) is not significantly different from the *Benchmark* one, suggesting that the *Feedback* Treatment has an impact on the Demand side game. However, the significant difference between *Benchmark* Treatment and *Feedback* Treatment needs deeper scrutiny. Decisions in the Demand-side game are not only contingent on gender; thus we cannot claim a significant presence or absence of treatment effects on task assignment decisions until controlling for several factors that affect employers' decisions. To account for such heterogeneity, Table 4 analyzes the probability of candidates being assigned to the CT and presents the marginal effects of different specifications of a probit model. The models regress a dummy variable (*Task*) that equals one if the candidate is assigned to CT and 0 otherwise on the candidates' gender (*Female* candidate), age, school, and the relative ranking of the candidate in her group according to the signal of performance (*Ranking*, continuous). In addition, we control for employers' characteristics such as gender, awareness of gender gaps in academic performance and market outcomes, and implicit gender stereotypes in the association of women and science gathered from the implicit association test (captured by the variable Dscore)[12]. The models also consider a set of covariates that account for the effect of the decision environment, such as the order in which the pools are displayed, the number of women in each group, the number of women in top-1 according to the signal, the position in which candidates are presented in the decision screen and the period[13]. We also control for the order of the experimental parts, i.e. whether the Demand game was played before or after the Supply game.

[10] Table A6 in the Appendix shows regressions on the probability of candidates being hired and reports the same result.
[11] The significant increase in the percentage of women assigned to CT from BT to FT is also supported by the rank order statistic Somers' D (provided by the `somersd' package in Stata). The test looks at the proportion of women assigned to CT accounting for the presence of clusters at the level of employers in the data (D=0.081, p=0.039).
[12] 218 out of 312 subjects (70%) obtained a score strictly higher than zero in the Implicit Association Test. A positive score translates into subjects holding implicit gender-stereotyped beliefs on mathematical ability. The same number of men and women (109) scored positive on the IAT.
[13] A formal specification of the variables used in the analysis of Table 4 is reproduced in the Appendix.

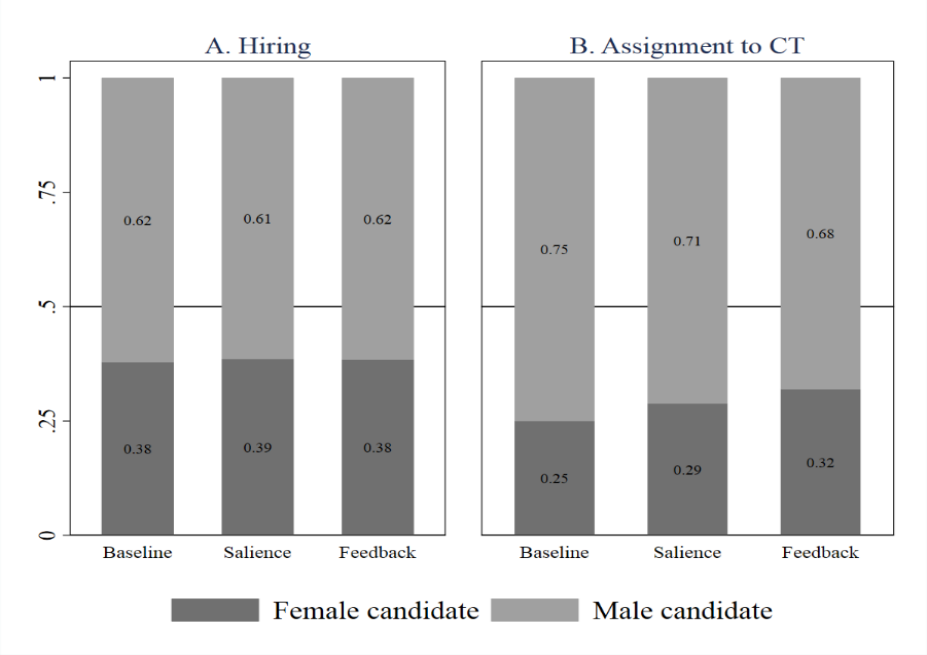


Figure 1. Employers' hiring and task assignment decisions in the demand-side game with CT = Challenging Task.
Note: The graph shows the proportion of female and male candidates among all candidates hired by employers (N=3744, 312 employers x 3 candidates hired x 4 rounds; 864 in BT; 1584 in ST; 1296 in FT) and among all candidates assigned to CT (N=1248, 312 employers x 1 candidate x 4 rounds; 288 in BT; 528 in ST; 432 in FT).

Models (1) and (2) in Table 4 consider the entire sample of employers. In Model (1), we observe that female candidates have a lower probability of being assigned to CT than male candidates. In Model (2), we complete the analysis by introducing the interactions between Female candidates and Treatments. First, the interaction between Female candidate and the *Feedback* Treatment presents a positive and significant effect (0.054), meaning that female candidates in the *Feedback* Treatment have a higher probability of being assigned to CT than female candidates in the *Baseline* Treatment. The *Feedback* Treatment coefficient in Model (2) is still significant but changes its sign (from 0.003 to -0.015), suggesting that feedback provision negatively affects male candidates. This result shows that a gender gap in favor of men exists in the *Baseline* Treatment (as shown by the negative coefficient of Female candidate in Model (2)) and that a replacement between male and female candidates in the assignment to CT is observed in the *Feedback* Treatment. The probability of female candidates being assigned to CT in the *Salience* treatment is not significantly different from that observed in the *Baseline* Treatment, as observed by the non-significant effect of the interaction between Female candidates and *Salience* Treatment. In this case, filling out the questionnaire without receiving feedback before the Demand-side game holds the decisions in the game unaffected.

Models (3) to (6) split the sample considering whether employers are aware of gaps in academic performance and early gaps in market outcomes. For all these models, the coefficient of Female candidate is negative and significant, indicating a detrimental effect for women in the baseline treatment. *Feedback* Treatment slightly improves the assignment of

women to CT in comparison to the *Baseline* Treatment among employers unaware of gaps in ability, represented by the positive but only weakly significant (at 10%) effect of the interaction between Female candidates and Feedback in Model (3).

The same result for such interaction is obtained in Model (5), which refers to participants unaware of gaps in labor market outcomes. For employers aware of gaps in ability (Model 4) and market outcomes (Model 6), *Feedback* Treatment does not change the women's assignment to CT. To sum up, Models (3) and (5) suggest that the positive effect of *Feedback* Treatment on women's assignment to the Challenging Task is driven by individuals unaware of gaps in labor market outcomes and ability.

Table 4. Treatment effect: Assignment to Challenging Task in the Demand-side game.

Dependent variable: Pr (Task=1)						
Estimation technique: Probit regression						
Population:	All	All	Unaware of the ability gap	Aware of the ability gap	Unaware of outcomes gap	Aware of outcomes gap
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Female candidate (0-1)	-0.110***	-0.141***	-0.145***	-0.132***	-0.164***	-0.103***
	(0.015)	(0.020)	(0.024)	(0.037)	(0.024)	(0.035)
Feedback (0-1)	0.003**	-0.015*	-0.014	-0.015	-0.014	-0.012
	(0.001)	(0.008)	(0.010)	(0.013)	(0.010)	(0.014)
Feedback x Female candidate		0.054**	0.055*	0.050	0.059*	0.035
		(0.025)	(0.032)	(0.041)	(0.032)	(0.040)
Salience (0-1)	0.002*	-0.007	-0.007	-0.005	-0.012	0.003
	(0.001)	(0.007)	(0.009)	(0.012)	(0.008)	(0.014)
Salience x Female candidate		0.028	0.031	0.019	0.050*	-0.006
		(0.023)	(0.029)	(0.040)	(0.028)	(0.040)
Ranking (Signal, continuous)	-0.013**	-0.013**	-0.016**	-0.009	-0.013	-0.014
	(0.007)	(0.007)	(0.008)	(0.011)	(0.008)	(0.011)
Age (continuous)	-0.011***	-0.011***	-0.011***	-0.012***	-0.009***	-0.016***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
STEM (0-1)	0.238***	0.238***	0.237***	0.240***	0.224***	0.257***
	(0.014)	(0.014)	(0.018)	(0.022)	(0.018)	(0.020)
Male evaluator (0-1)	-0.001	-0.001	-0.003*	0.002	-0.003**	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Dscore>0 (0-1)	-0.001	-0.001	-0.000	-0.002	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Aware of outcomes gap (0-1)	0.000	0.000	-0.001	0.003*		
	(0.001)	(0.001)	(0.001)	(0.001)		
Aware of ability gap (0-1)	0.000	0.000			-0.001	0.002
	(0.001)	(0.001)			(0.001)	(0.001)
Additional Controls	✓	✓	✓	✓	✓	✓
Observations	7,488	7,488	4,440	3,048	4,392	3,096

Note: The table shows the marginal effects of different probit estimations of the probability of candidates being assigned to the Challenging Task. The overall number of observations corresponds to 7488 (312 evaluators x 6 candidates x 4 rounds). Unaware of inequalities are 183 subjects (183 evaluators x 6 candidates x 4 rounds =4392 observations). Aware of inequalities are 129 subjects (129 evaluators x 6 candidates x 4 rounds=3096 observations). Unaware of ability are 127 185 subjects (185 subjects x 6 candidates x 4 rounds=4440 observations). Aware of ability are 127 subjects (127 subjects x 6 candidates x 4 rounds = 3048 observations). Additional controls include the order of the tasks, candidates' position in the decision screen, number of women in the pool, number of women among the top-1 performers according to the signal, and period. Standard errors, clustered at the employer level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The main findings from this section are summarized in Result 2.

Result 2: Demand-side game (hiring and tasks assignment). *Feedback Treatment does not affect hiring decisions, but it increases the probability of female candidates being assigned to the Challenging Task compared to the Baseline Treatment. This effect is driven by employers unaware of the gaps in ability and early labor market outcomes. Salience Treatment, instead, displays no effect compared to the Baseline treatment.*

We complete this Subsection with a remark on the impact of the feedback treatment on efficiency. Differently from Ruebens et al. (2014), in our experiment, the increased assignment of women to the CT generated by the feedback provision does not improve efficiency. To show that, we defined efficiency from two perspectives: i) in terms of employers' earnings and ii) as a percentage deviation from the maximum attainable earnings. Using the two measures, we show that efficiency differences are not significant across treatments (see Table A7 in the Appendix for a more detailed analysis). This result depends on the specific groups' composition (displayed in Table A1a and A1b in the Appendix) and the performances of the groups' members in the two tasks that does not display substantial gender differences.

4.3. Supply-side game: Willingness to compete under feedback provision

In this section, we focus on the supply side of the market. The willingness to compete is expressed through the dichotomous decision in Stage 3 of the supply-side game, where subjects decide whether to select a tournament payment scheme or opt for a non-competitive payment scheme (piece rate) to be applied to their performance in the addition task. In our sample, where the number of men and women is perfectly balanced (N=156/312), the proportion of men who decided to compete (65%) doubles that of women (31%) (Mann-Whitney Z=5.995, p=0.000). When considering the performance, we find that men outperformed women under the piece rate payment (we refer to the piece rate of Stage 1): men correctly solved, on average, 5.22 calculations (sd=2.78), and women solved 4.43 calculations (sd=2.11) correctly (Mann-Whitney test Z= 2.765, p=0.006). When considering Stage 2, the gender differences in performance are not significant in the tournament stage: men correctly solved 5.95 (sd=2.84) calculations, and women correctly solved 5.34 (sd=2.40) calculations. This suggests that women hold the same ability level as men in this task but exert more effort when competing, or that the learning processes for males and females may differ.

Consider now Figure 2. Panel A shows the proportion of subjects who decided to compete in each treatment. It can be noted that the *Feedback* Treatment significantly increases the share of women choosing to compete compared to the *Baseline*, but this does not help to close the gender gap in willingness to compete because men's willingness to compete significantly increases as well. Eventually, the difference in the competing gap under *Baseline* and *Feedback* is not

significant. Panel B displays the decision to compete by gender over the pool of same-sex subjects. Results mirror the ones displayed in panel A: in the feedback treatment, we can see that both male and female subjects increase their competitiveness compared to the Baseline treatment (Women: Z=-1.369, p=0.087 Men: Z=-1.484, p=0.075).

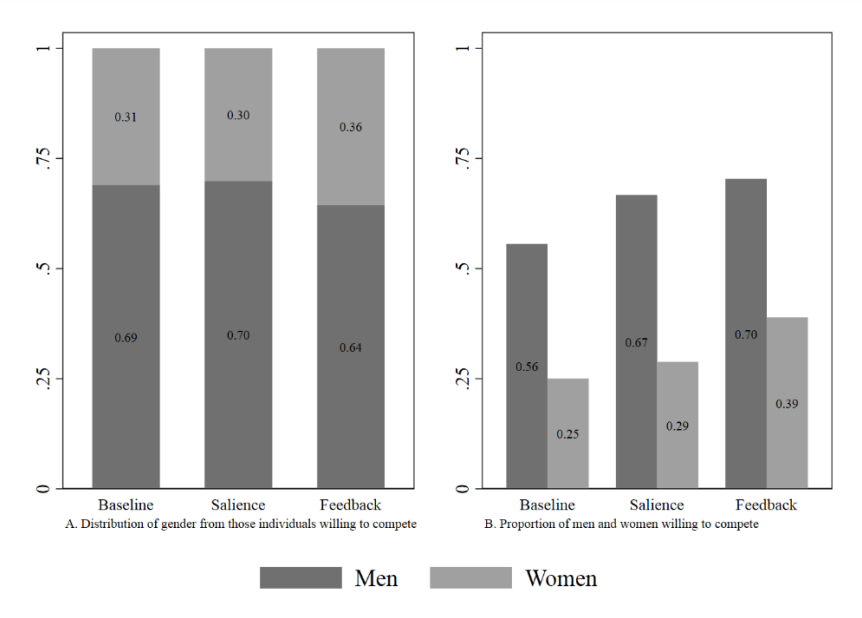


Figure 2. Proportion of subjects willing to compete in the supply-side game.

Note: Panel A shows the proportion of women and men among those subjects who decided to compete (N=29 in Baseline, N=63 in Salience, and N=59 in Feedback). Panel B shows the proportion of men and women who decided to compete over the pool of same-sex subjects (Baseline: n=36, Salience n=66, Feedback: n=54).

Beyond gender and the treatment variation, we expect that some other aspects influence the decision to compete: subjects' self-confidence, the decision environment, and subjects' personal characteristics. We control for these aspects in different specifications of a probit model that regresses the probability of choosing the competitive payment scheme in Stage 3 (*Choice*). The marginal effects of these models are presented in Table 5. We consider, as independent variables, the subjects' gender, the number of correct calculations under tournament in Stage 1 and Stage 2, the field of study, age, the degree of risk-loving, self-confidence[14], whether the subject played the supply-side game before or after the demand game, awareness of the ability gap and the gap in market outcomes, and implicit gender stereotypes gathered from the implicit association test.[15]

[14] 77 subjects (25%) are categorized as underconfident individuals (39 are men and 38 are women). This is done by comparing individuals' expected ranking to their actual ranking in the tournament. 50% of the sample were overconfident, and the remaining 25% were correct in their guess.

[15] A formal specification of the variables used in the analysis of Table 5 is reproduced in the Appendix.

Table 5. Treatment effect by gender: Willingness to compete in the supply-side game.

Dependent variable:	Pr (Choice=1)					
Estimation technique:	Probit regression					
Population:	All	All	Women		Men	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Female (0-1)	-0.302***	-0.283**				
	(0.066)	(0.137)				
Feedback (0-1)	0.181**	0.189	0.220*	-0.152	0.139	0.355*
	(0.087)	(0.121)	(0.117)	(0.187)	(0.106)	(0.208)
Female x Feedback		-0.018				
		(0.174)				
Salience (0-1)	0.057	0.071	0.086	0.053	0.083	0.002
	(0.083)	(0.115)	(0.113)	(0.166)	(0.101)	(0.203)
Female x Salience		-0.029				
		(0.169)				
Performance (Piece-rate)	0.023	0.023	0.016	0.017	0.022	0.018
	(0.017)	(0.018)	(0.023)	(0.023)	(0.020)	(0.021)
Performance (Tournament)	0.051***	0.051***	0.047**	0.053**	0.042**	0.050**
	(0.017)	(0.017)	(0.024)	(0.024)	(0.020)	(0.021)
Engineering (0-1)	0.229***	0.228***	0.416***	0.366***	0.001	0.001
	(0.085)	(0.085)	(0.115)	(0.115)	(0.112)	(0.115)
Economics (0-1)	0.186**	0.185**	0.129	0.103	0.127	0.119
	(0.082)	(0.083)	(0.095)	(0.097)	(0.129)	(0.133)
Age (continuous)	0.008	0.008	0.013	0.013	-0.005	-0.008
	(0.012)	(0.012)	(0.014)	(0.014)	(0.017)	(0.017)
Risk-lover (continuous)	0.007	0.007	0.010	0.006	0.006	0.010
	(0.014)	(0.014)	(0.018)	(0.018)	(0.018)	(0.018)
Underconfident (0-1)	-0.233***	-0.233***	-0.321***	-0.196	-0.140	0.103
	(0.073)	(0.073)	(0.098)	(0.193)	(0.092)	(0.168)
Underconfident x Feedback				-0.121		-0.382*
				(0.236)		(0.225)
Underconfident x Salience				-0.327		-0.297
				(0.266)		(0.235)
Dscore>0 (0-1)	0.038	0.037	-0.007	-0.192	0.121	0.074
	(0.068)	(0.069)	(0.082)	(0.175)	(0.086)	(0.170)
Dscore>0 x Feedback				0.520**		-0.147
				(0.235)		(0.224)
Dscore>0 x Salience				0.093		0.244
				(0.216)		(0.223)
Supply first (0-1)	-0.023	-0.022	-0.053	-0.030	-0.010	-0.006
	(0.063)	(0.063)	(0.078)	(0.079)	(0.080)	(0.081)
Aware of outcomes gap (0-1)	0.002	0.003	-0.080	-0.119	0.029	0.016
	(0.066)	(0.067)	(0.089)	(0.093)	(0.086)	(0.086)
Aware of ability gap (0-1)	0.032	0.031	0.065	0.088	-0.015	-0.022
	(0.065)	(0.065)	(0.085)	(0.085)	(0.085)	(0.087)
Observations	312	312	156	156	156	156

Note: The table shows the marginal effects of different probit estimations of candidates' probability of choosing the competitive payment scheme in Stage 3 of the supply-side game. Standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

In model (1), which considers the entire sample, the drivers that significantly predict the probability of choosing the competition are gender, performance in the tournament, the field of study, and self-confidence. These drivers align with previous literature (Niederle and Vesterlund, 2007; Kamas and Preston, 2012). While tournament performance and studying Engineering or Economics and Management (in comparison to studying Law) affect the choice to compete positively, being female and underconfident affect it negatively. Moreover, we observe in Model (1) that Feedback provision significantly increases the willingness to compete while Salience does not. Model (2) introduces an interaction between being female and the treatments to understand who drives the increase in the general willingness to compete under *Feedback*. The interactions between female and *Feedback* Treatment and the isolated feedback variable are not significant, meaning that the feedback effect observed in the whole population in Model (1) is not driven neither by women nor by men. This confirms that *Feedback* affects women's and men's willingness to compete similarly, and that information provision does not help reduce the gender gap in the willingness to compete. Interestingly, playing the supply-side game before or after the demand game does not affect the willingness to compete.

Models (1) and (2) compare the relationship between gender and the three treatments. Models (3) to (6) split the sample by gender to determine the pure within-gender effect of the treatments. In model (3), which considers only women, we observe that the coefficient of *Feedback* is only significant at 10%, meaning that women significantly (but weakly) increase their willingness to compete when they are provided with feedback. We also observe that tournament performance, the field of study, and confidence play a meaningful role in women's decision to compete. In Model (5), which considers men alone, the field of study and confidence are not relevant anymore, and tournament performance is essentially the only factor playing a key role.

Finally, Models (4) and (6) analyze how the treatments interact with men's and women's beliefs. Specifically, these models introduce interactions between the treatment and beliefs about subjects' ranking and implicit biases (IAT score). Implicit biases can affect the subjective comparison of performance between different genders and may impact willingness to compete. The interaction between treatments and underconfident subjects allows us to see how feedback provision affects those subjects whose actual ranking in the group is above their expected ranking and those with a positive *dscore* (i.e. implicit bias). In Model (4), we observe that *Feedback* and *Salience* Treatments do not affect the decisions of underconfident women since the effect of the interactions between treatments and underconfidence is not significant. Nevertheless, we do observe that feedback provision significantly increases the willingness to compete of women holding stereotyped beliefs (see the positive and significant coefficient of the interaction between *Feedback* and a positive score in the IAT in Model (4). Model (6) confirms that performance in previous stages is the only explanator of men's willingness to compete.

Table 6 repeats the analysis but splits the sample into different subgroups based on subjects' awareness, like in Table 4. Models (1), (3), (5), and (7) in Table 6 show that, irrespective of awareness of gender gaps in ability and market outcomes, women are less likely to choose competition than men. The coefficient for Female is negative and significant in all models. In addition, we observe that only those subjects who were already aware of gender gaps in ability (Model 3) and outcomes (Model 7) were affected by the feedback provision, which increased their willingness to compete. The *Salience* treatment, again, does not affect the decision to compete substantially (except for subjects who were already aware of the gender gap in ability (Model 3).

Table 6. Treatment effect by awareness: Willingness to compete in the supply-side game.

Dependent variable: Pr (Choice=1)								
Estimation technique: Probit regression								
Population:	Unaware ability gap		Aware ability gap		Unaware outcome gap		Aware outcome gap	
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.298***	-0.231	-0.318***	-0.394	-0.325***	-0.257	-0.336***	-0.357
	(0.086)	(0.166)	(0.107)	(0.255)	(0.088)	(0.169)	(0.106)	(0.244)
Feedback	0.108	0.151	0.428***	0.382*	0.103	0.163	0.316**	0.298
	(0.115)	(0.153)	(0.142)	(0.204)	(0.111)	(0.148)	(0.159)	(0.216)
Female x Feedback		-0.097		0.100		-0.149		0.033
		(0.214)		(0.321)		(0.227)		(0.305)
Salience	-0.055	-0.014	0.350**	0.307*	0.026	0.052	0.150	0.141
	(0.108)	(0.149)	(0.137)	(0.179)	(0.103)	(0.139)	(0.140)	(0.202)
Female x Salience		-0.089		0.089		-0.060		0.021
		(0.213)		(0.291)		(0.210)		(0.299)
Performance (Piece-rate)	0.025	0.025	0.016	0.014	0.028	0.031	0.001	0.001
	(0.022)	(0.022)	(0.030)	(0.031)	(0.023)	(0.023)	(0.027)	(0.027)
Performance (Tournament)	0.050**	0.049**	0.073**	0.074**	0.040*	0.038	0.090***	0.090***
	(0.022)	(0.022)	(0.029)	(0.030)	(0.023)	(0.023)	(0.025)	(0.025)
Engineering	0.208*	0.207*	0.215	0.227	0.129	0.122	0.351**	0.350**
	(0.115)	(0.115)	(0.139)	(0.139)	(0.108)	(0.108)	(0.169)	(0.166)
Economics	0.231**	0.230**	0.078	0.088	0.139	0.129	0.214	0.213
	(0.115)	(0.115)	(0.127)	(0.128)	(0.102)	(0.102)	(0.170)	(0.169)
Age	0.016	0.015	-0.004	-0.003	0.014	0.013	0.002	0.002
	(0.016)	(0.016)	(0.019)	(0.019)	(0.016)	(0.016)	(0.020)	(0.021)
Risk-lover (continuous)	0.009	0.010	0.019	0.017	-0.022	-0.021	0.053**	0.053**
	(0.018)	(0.018)	(0.025)	(0.026)	(0.018)	(0.018)	(0.025)	(0.025)
Underconfident (0-1)	-0.238**	-0.238**	-0.281**	-0.280**	-0.300***	-0.297***	-0.155	-0.154
	(0.097)	(0.097)	(0.133)	(0.132)	(0.095)	(0.096)	(0.121)	(0.121)
Dscore>0	-0.040	-0.040	0.183	0.180	-0.038	-0.042	0.075	0.074
	(0.089)	(0.090)	(0.119)	(0.121)	(0.090)	(0.089)	(0.116)	(0.119)
Supply first	-0.015	-0.014	-0.083	-0.088	0.009	0.010	-0.077	-0.078
	(0.083)	(0.084)	(0.107)	(0.107)	(0.083)	(0.083)	(0.103)	(0.103)
Aware of outcome gap (0-1)					0.010	0.009	0.130	0.131
					(0.088)	(0.088)	(0.108)	(0.109)
Aware of ability gap (0-1)	-0.014	-0.010	0.093	0.088				
	(0.089)	(0.089)	(0.108)	(0.109)				
Observations	185	185	127	127	183	183	129	129

Note: The table shows the marginal effects of different probit estimations of the probability of candidates of choosing the competitive payment scheme in the Stage 3 of the supply-side game. Standard errors, clustered at individual level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Models (2), (4), (6), and (8) consider an interaction between Treatments and Female. In none of the cases the interaction is significant, suggesting that, compared to the *Baseline*, the gender gap in willingness to compete does not change under the feedback provision. The previous findings are summarized in Result 3.

Result 3: Supply-side game (willingness to compete). *The Feedback Treatment enhances the willingness of both males and females to compete, yet it does not alter the gender gap in competitiveness. Individuals aware of gaps in academic performance and early labor market outcomes, as well as women holding implicit biases, are more willing to compete under the Feedback treatment.*

5. Discussion

The findings outlined in **Result 1** reveal that Italian college students, regardless of their field of study, generally lack awareness of the persistent gender gaps in the labor market. This lack of awareness persists despite the frequent media discussions following movements like #MeToo# and, for the subset of students invited to the lab in 2021, even amid media reports highlighting increased gender inequalities in the labor market due to the Covid-19 pandemic. Notably, *a significant majority of students (60%) believe that they will be entering a gender-neutral workplace*. While these students may be aware of gender disparities among more established members of the workforce, they either ignore or underestimate the early emergence of these gaps, which occur as soon as one year after graduation.

Upon closer examination of specific academic domains, we observe that female students in the field of Engineering display a greater awareness of gender gaps in labor market outcomes compared to their male counterparts, as well as compared to female students in other fields of study. This heightened awareness among Engineering female students could be attributed to better-informed peers and family members, who might have advised them about the challenges and potential discrimination they could face in the labor market. Alternatively, they might have actively sought information to gain a deeper understanding of possible market outcomes.

Similarly, a general lack of awareness is evident regarding the gender gap in academic performance, a topic that hasn't yet received substantial media attention. Notably, women consistently outperform men across all fields of study in college. Surprisingly, most students (60%) incorrectly believe that male students achieve better grades. An interesting exception is observed among male students in Law, where 65% correctly acknowledge that their female counterparts perform better.

Our *Feedback* Treatment offers insights into the enduring gender gaps within labor market outcomes. Specifically, it addresses disparities in permanent contracts, wage rates one year and five years after graduation, and self-employment. Moreover, the *Feedback* Treatment sheds light on the gender gap in academic performance, consistently favoring

women across all fields of study. Importantly, this second piece of information regarding academic performance serves as a positive affirmation for women. It reminds them that their performance in college is at least as remarkable as that of their male counterparts.

Furthermore, when coupled with information on gender gaps in labor market outcomes, knowledge about the gender gap in academic performance might also serve as a de-biasing intervention. This combined information allows students to comprehend that disparities in labor market outcomes are not driven by differences in academic performance (and might thus be unjust). Consequently, the feedback treatment could potentially evoke a sense of frustration among female students, who may perceive these insights as indicative of discrimination against women. Overall, enlightening women about gender gaps in academic performance and labor market outcomes carries the risk of triggering a self-fulfilling prophecy, leading women to feel disheartened and ultimately acquiesce to the prevailing unequal labor market conditions, thereby perpetuating the status quo. In our experiment, this scenario would pertain to a situation where, under the *Feedback* Treatment, women exhibit a stronger tendency to 'shy away from competition' compared to their responses under the *Baseline* and/or *Salience* Treatment. **Result 3** counters this notion. It reveals that *Feedback* bolsters women's willingness to engage in competition; see discussion pertaining to Result 3 below.

The Demand-side game examines the influence of information on employers' behaviors. **Result 2** reveals that, although the *Feedback* Treatment doesn't increase the overall count of women hired for one of the two tasks, it does raise the proportion of women assigned to the Challenging task by both male and female employers. Our data also suggests that the increase in women being assigned to the challenging task is more pronounced among students who are unaware of gender gaps, though this outcome is moderately significant (10%).

It's important to note that Feedback encourages women's assignment to the challenging task through two mechanisms. First, it provides positive information about women's academic performance. Second, it highlights inequalities in the labor market and functions as a de-biasing intervention, promoting more equitable assignment of women to the challenging ask. These mechanisms could potentially overlap and reinforce each other. As mentioned earlier, Result 2 seems to be driven by individuals unaware of gender gaps in labor market outcomes and academic performances. Additionally, the lack of significance in the IAT test score indicates that students, regardless of whether they hold implicit stereotypes, exhibit similar behaviors. These combined observations emphasize information as the primary mechanism driving Result 2.

Importantly, it's worth noting that *Feedback* does not exert an influence on hiring behaviors, i.e., on the overall percentage of women hired. This finding aligns with *Feedback*'s role as an informative tool. In the context of our study, participants encounter math tasks of varying difficulty levels. It's plausible that awareness of women's capabilities doesn't significantly impact hiring decisions, as college students may accurately perceive that men and women perform

comparably in the Simple Task. Conversely, expectations regarding men's and women's abilities in completing the Challenging Task might be less accurate. Consequently, students might adjust their assignment to the Challenging Task based on the provided information.

The Supply-side game investigates the effects of the *Feedback* Treatment on job market candidates' willingness to compete. **Result 3** indicates a positive impact of feedback provision on women's inclination to compete. This rise in women's willingness to compete is particularly driven by female students who possess implicit biases and are already aware of gender gaps in ability and market outcomes. This suggests that feedback provision serves both as a 'de-biasing treatment' and a 'confidence boost' for women already aware of existing gender disparities. Moreover, the absence of significant effects from the *Salience* Treatment implies that simply 'priming gender gaps' is insufficient to foster a heightened willingness to compete. Instead, a more substantial intervention that instills confidence in women's academic performance is required to drive behavioral changes.

The coexistence of negative news regarding gender gaps in market outcomes and positive news about women's higher academic performance might help explain why *Feedback* Treatment fails to influence the behaviors of women unaware of gender gaps. Our interpretation posits that, for women lacking awareness of gender gaps, the positive and negative pieces of information may have offset each other. Nevertheless, the insightful and policy-relevant takeaway from this lack of reaction to *Feedback* by these women is that information about gender gaps in labor market outcomes doesn't appear to deter (previously unaware) female students from continuing to compete at their previous levels.

Lastly, within the Supply-side game, the *Feedback* Treatment also elicits an increase in men's willingness to compete. Once again, this result is driven by male students who are aware of gender gaps. This outcome was unexpected to us, and we interpret it as follows: Men exhibit a natural inclination toward competition, particularly after being 'reminded' of women's superior performance in college. In essence, upon realizing that women outperform them academically, male students embrace the challenge and seek to demonstrate their competitive prowess within their mixed-gender groups.

6. Conclusion

In our laboratory experiment, we first assessed students' awareness of early gender gaps and then analyzed how providing information about these early gaps influenced their behaviors. Specifically, we examined the impact of this information on students' choices while taking on the roles of job seekers or employers. Our feedback treatment encompassed two pieces of information crucial for making effective decisions in a competitive work environment.

The initial piece of information pertained to college performance, revealing a gap that favors women. The second piece focused on various early labor market outcomes—such as employment probabilities, salary levels, and the likelihood of self-employment—all of which tend to favor men. Collectively, these two pieces of evidence indicate that men's greater success at the beginning of their professional career *cannot be* attributed to their superior academic abilities. Therefore, the *Feedback* Treatment serves a dual purpose: (i) it offers insights into early gender gaps, and (ii) it serves as a de-biasing/priming intervention. By controlling for participants' awareness or lack thereof regarding gender gaps, as well as their adherence to stereotypical beliefs (as measured by the IAT test score), we can differentiate whether our treatment impacts behaviors primarily through the delivery of information or the de-biasing/priming mechanism.

From our findings, we can draw several policy implications. To begin with, it's important to note that students are generally unaware of the existing early gender gaps in the labor market. Result 1 indicates that a significant number of both male and female students believe that they will be entering a workplace where gender equality is prevalent. Consequently, our first policy implication underscores the necessity of offering information about early gender gaps in the labor market to college students despite the ongoing media debate on this issue.

Secondly, Result 2 points to the relevance of interventions targeting employers. In this respect, note that the *Feedback* Treatment does not work as a de-biasing intervention and does not increase the share of hired women. Still, emphasizing women's academic performance helps increase the number of women selected for the Challenging Task. Hence, our findings suggest that priming gender gaps in labor market outcomes and informing about women's academic performance could help increase gender equity inside organizations by decreasing the glass ceiling.

Finally, a primary concern when presenting information about early gender gaps in the labor market is the potential for this knowledge to have a negative impact on women's confidence and motivation. However, Result 3 suggests that providing a positive message about women's academic abilities will counteract this discouragement. Specifically, we observed that women did not reduce their willingness to compete. Interestingly, our intervention had the unintended effect of increasing men's willingness to compete. This suggests that upon realizing women are stronger competitors than previously thought, men might become more eager to compete against women in the workplace.

References

Babcock, L., Recalde, M. P., & Vesterlund, L. (2017a). "Gender Differences in the Allocation of Low-Promotability Tasks: The Role of Backlash." *American Economic Review,* 107(5), 131-35.

Balleer, A., Duernecker, G., Forstner, S., & Goensch, J. (2023). "Biased expectations and labor market outcomes: Evidence from German survey data and implications for the East-West wage gap." CESifo Working Paper, No. 10336.

Bertrand, M. (2020). "Gender in the 21st Century," *American Economic Association Papers and Proceedings*, 2020, 110 (5): 1-24.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). "Gender, competitiveness, and career choices." *The Quarterly Journal of Economics*, 129 (3), 1409-1447.

Chen, Y., & He, Y. (2021). "Information acquisition and provision in school choice: an experimental study." *Journal of Economic Theory*, 197, 105345.

De Pater, I. E., Van Vianen, A. E., & Bechtoldt, M. N. (2010). "Gender differences in job challenge: A matter of task allocation." *Gender, Work & Organization*, 17(4), 433-453.

Dohmen, T., & Falk, A. (2011). "Performance pay and multidimensional sorting: Productivity, preferences, and gender." *American Economic Review*, 101 (2), 556-90

Domínguez, J. J., & Montinari, N. (2021). "Gender Quotas and Task Assignment in Organizations," WP 21/13, Department of Economic Theory and Economic History of the University of Granada.

Filippin, A. and A. Ichino (2005), "Gender wage gap in expectations and realizations," *Labour Economics* 12(1), 125-145.

Fischbacher, U. (2007). "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental economics*, *10*(2), 171-178.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). "Measuring individual differences in implicit cognition: the implicit association test." *Journal of personality and social psychology*, *74*(6), 1464.

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

Heinz, M., Normann, H. T., & Rau, H. A. (2016). "How competitiveness may cause a gender wage gap: Experimental evidence." *European Economic Review*, 90, 336-349.

Kamas, L., & Preston, A. (2012). "The importance of being confident; gender, career choice, and willingness to compete." *Journal of Economic Behavior & Organization*, 83(1), 82-97.

Kim, S., Klager, C., & Schneider, B. (2019). "The effects of alignment of educational expectations and occupational aspirations on labor market outcomes: Evidence from NLSY79." *The Journal of Higher Education*, 90(6), 992-1015.

Mueller, A. I., & Spinnewijn, J. (2023). "Expectations data, labor market, and job search." *Handbook of Economic Expectations*, 677-713.

Niederle, M., & Vesterlund, L. (2007). "Do women shy away from competition? do men compete too much?" *The Quarterly Journal of Economics*, 122 (3), 1067-1101.

Peterle, E., & Rau, H. A. (2017). "Gender differences in competitive positions: Experimental evidence on job promotion," CEGE Number 303 – February 2017.

Reuben, E., Rey-Biel, P., Sapienza, P., & Zingales, L. (2012). "The emergence of male leadership in competitive environments." *Journal of Economic Behavior & Organization*, 83(1), 111-117.

Reuben, E., Sapienza, P., & Zingales, L. (2014). "How stereotypes impair women's careers in science." PNAS *Proceedings of the National Academy of Sciences of the United States of America*, 111(12), 4403–4408. https://doi.org/10.1073/pnas.1314788111

Salvi del Pero, A. and A. Bytchkova (2013), "A Bird's Eye View of Gender Differences in Education in OECD Countries," OECD Social, Employment and Migration Working Papers, No. 149, OECD Publishing, Paris, https://doi.org/10.1787/5k40k706tmtb-en.

Sipe, S., C.D. Johnson, and D.K. Fisher (2009), "University Students' Perceptions of Gender Discrimination in the Workplace: Reality Versus Fiction," *Journal of Education for Business* 84(6), 339-349.

von Bieberstein, F., Jaussi, S., & Vogel, C. (2020)." Challenge-seeking and the gender wage gap: A lab-in the-field experiment with cleaning personnel." *Journal of Economic Behavior & Organization*, 175, 251-277.

Entering a gender-neutral workplace?

College students' expectations and the impact of information provision.

Francesca Barigozzi, José J. Domínguez, and Natalia Montinari

Appendix

Table A1a. Pools of candidates in the Demand-side game.

Group	#	Year of birth	Gender	Field	Signal	#Correct ST*	#Correct CT**
1	1	1998	Female	Social Sciences	8	16	15
	2	1994	Male	Social Sciences	8	15	10
	3	1996	Female	Sciences	7	14	12
	4	1980	Male	Sciences	7	16	10
	5	1995	Male	Sciences	6	11	8
	6	1993	Female	Sciences	6	9	8
2	1	1997	Male	Social Sciences	8	16	4
	2	1992	Female	Sciences	7	17	7
	3	1993	Female	Social Sciences	6	12	6
	4	1997	Male	Languages	6	14	3
	5	1996	Male	Engineering	5	9	11
	6	1999	Female	Social Sciences	5	9	10
3	1	1992	Male	Sciences	6	13	2
	2	1997	Female	Sciences	5	10	18
	3	1995	Male	Sciences	4	11	9
	4	1998	Female	Social Sciences	4	8	9
	5	1997	Male	Sciences	4	13	13
	6	1996	Male	Engineering	4	9	11
4	1	1995	Male	Social Sciences	10	18	9
	2	1995	Female	Social Sciences	10	20	8
	3	1998	Male	Sciences	10	19	7
	4	1997	Female	Social Sciences	9	19	10
	5	1993	Male	Engineering	9	19	14
	6	1994	Male	Engineering	9	16	11

Note: * Performance in the Simple Task, ** Performance in the Challenging Task.

Table A1b. Candidates' gender differences in signal and performance in the Demand-side game.

Gender	Signal	Correct ST	Correct CT
Female candidates (n=10)	6.70	13.40	10.30
Male candidates (n=14)	6.86	14.21	8.71
Total (n=24)	6.79	13.88	9.38
Mann-Whitney test	Z=-0.148	Z=-0.471	Z=0.441

of gender differences	p=0.882	p=0.637	p=0.658

Note: Candidates' signal and performance in the demand-side game. The 24 candidates were extracted from the 56 subjects analyzed by Domínguez and Montinari (2021). Neither the subsample of 24 candidates nor the overall sample of 56 subjects present significant gender differences in performances in ST and CT; see also Domínguez and Montinari, 2021).

Table A2. Sample characteristics.

	Economics	Engineering	Law	Total
Proportion of men	0,476	0,661	0,302	0,500
Average Age	22,447	23,983	24,406	23,583
Average enrolment year in College	2,552	3,247	4,825	3,448

Table A3. Proportion of respondents assigning a higher probability to men (Giovanni) than to women (Angela) for each question pre and post-COVID19.

	A1		B1		B2		B3		B5		B5		B6	
	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men
Pre-COVID19	0.115	0.191	0.46	0.356	0.398	0.278	0.362	0.304	0.398	0.278	0.46	0.295	0.407	0.347
Post-COVID19	0.209	0.170	0.604	0.292	0.488	0.39	0.488	0.39	0.488	0.39	0.488	0.365	0.488	0.463
Mann-Whitney p	0.131	0.772	0.108	0.460	0.310	0.183	0.153	0.315	0.310	0.183	0.753	0.407	0.360	0.191

Table A4. Average probability assigned to *Angela* and *Giovanni* in each question.

	Economics				Engineering				Law			
	Women		Men		Women		Men		Women		Men	
Question	GIO	ANG	GIO	ANG	GIO	ANG	GIO	ANG	GIO	ANG	GIO	ANG
Ability	62.73	71.35	61.74	59.00	63.17	66.32	59.30	61.93	69.62	73.47	62.50	76.00
Permanent contract	59.15	52.96	55.64	54.14	71.32	63.22	70.56	67.49	37.00	33.75	39.62	41.31
Medium wage in 1 year	58.64	52.38	55.60	52.82	57.02	52.46	64.89	61.26	53.85	50.35	50.19	49.23
High-stake wage in 1 year	37.24	32.11	36.50	34.86	42.29	36.39	48.90	45.04	31.25	28.27	27.50	24.00
Medium wage in 5 years	67.42	62.78	58.68	57.02	63.37	54.66	64.96	61.64	59.82	55.53	62.50	62.31
High-stake wage in 5 years	55.38	50.27	40.56	38.34	46.44	35.49	51.35	48.36	42.95	38.87	41.92	43.04
Self-employed in 5 years	69.82	64.49	53.18	50.16	63.07	57.71	60.50	55.08	60.33	54.20	60.42	59.81

Note: The table summarizes the average probability assigned to men (Giovanni, GIO) and women (Angela, ANG) in each question, conditional on gender and field of study.

Table A5. Proportion of subjects attributing G<A, G=A, or G>A in each question by gender.

	Men			Women			Total		
Question	G<A	G=A	G>A	G<A	G=A	G>A	G<A	G=A	G>A
Ability	0.41	0.40	0.19	0.40	0.46	0.14	0.41	0.43	0.16
Permanent contract	0.09	0.57	0.34	0.09	0.41	0.50	0.09	0.49	0.42
Medium wage in 1 year	0.06	0.63	0.31	0.02	0.56	0.31	0.04	0.60	0.36
High-stake wage in 1 year	0.06	0.62	0.33	0.02	0.58	0.40	0.04	0.60	0.36
Medium wage in 5 years	0.06	0.63	0.31	0.01	0.56	0.43	0.04	0.59	0.37
High-stake wage in 5 years	0.05	0.64	0.31	0.01	0.52	0.47	0.03	0.58	0.39
Self-employed in 5 years	0.04	0.58	0.38	0.04	0.53	0.43	0.05	0.55	0.40

Note: This table reports the proportion of subjects assigning higher, equal or lower probability to men (G) than to women (A) in each question. Where G<A means Higher probability to Angela, G>A means Higher probability to Giovanni, and G=A means equal probability to Giovanni and Angela.

Table A6. Treatment effect: Hiring in the demand-side game.

Dependent variable:	Pr (Hired=1)					
Estimation technique:	Probit regression					
Population:	All	All	Unaware of outcomes gap	Aware of outcomes gap	Unaware of ability gap	Aware of ability gap
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Female candidate	-0.068***	-0.080***	-0.090***	-0.063	-0.114***	-0.018
	(0.019)	(0.028)	(0.034)	(0.047)	(0.031)	(0.051)
Feedback	-0.000	-0.005	0.003	-0.011	-0.017	0.017
	(0.001)	(0.012)	(0.014)	(0.020)	(0.014)	(0.021)
Feedback x Female candidate		0.014	-0.007	0.028	0.051	-0.053
		(0.033)	(0.039)	(0.056)	(0.039)	(0.058)
Salience	-0.001	-0.007	-0.008	-0.004	-0.026**	0.025
	(0.001)	(0.011)	(0.013)	(0.019)	(0.013)	(0.019)
Salience x Female candidate		0.016	0.023	0.006	0.074**	-0.077
		(0.031)	(0.038)	(0.053)	(0.037)	(0.055)
Signal (rank)	-0.090***	-0.090***	-0.082***	-0.103***	-0.091***	-0.090***
	(0.011)	(0.011)	(0.015)	(0.018)	(0.015)	(0.017)
Age	-0.024***	-0.024***	-0.018***	-0.032***	-0.024***	-0.024***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
STEM	0.446***	0.446***	0.418***	0.490***	0.444***	0.451***
	(0.025)	(0.025)	(0.032)	(0.039)	(0.031)	(0.041)
Male evaluator	-0.001	-0.001	-0.002**	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dscore>0	0.000	0.000	-0.000	0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Aware of outcomes gap (0-1)	0.001	0.001			0.000	0.001
	(0.001)	(0.001)			(0.001)	(0.001)
Aware of ability gap (0-1)	0.001	0.001	0.000	0.001		
	(0.001)	(0.001)	(0.001)	(0.001)		
Additional Controls	✓	✓	✓	✓	✓	✓
Observations	7,488	7,488	4,392	3,096	4,440	3,048

Note: The table shows the marginal effects of different Probit estimations of candidates' probability of being assigned to the Challenging Task. The overall number of observations corresponds to 7488 (312 evaluators x 6 candidates x 4 rounds). Additional controls include the order of the tasks, candidates' position in the decision screen, number of women in the pool, and number of women among the top-1 performers according to the signal and period. Standard errors, clustered at the employer level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7. Analysis of efficiency in the Demand-side game.

Treatment	Earnings (Total)	Earnings ST	Earnings CT	Deviation (Total)	Deviation ST	Deviation CT
Baseline	13.47	3.03	10.44	0.238	0.091	0.270
Salience	13.20	2.88	10.31	0.249	0.093	0.283
Feedback	12.91	2.96	9.95	0.238	0.085	0.270
P-value (Kruskal-Wallis)	0.733	0.174	0.663	0.558	0.461	0.595

Note: The table shows the average earning obtained by employers in each treatment and the percentage deviation from the maximum attainable earning. It considers the total measure and the earnings deviation associated with each task.

Table 4: specifications of the models.

Table 4 analyzes the probability of candidates being assigned to the Challenging Task (CT) and presents the marginal effects of different specifications of a probit model. The models regress a dummy variable that equals 1 if the candidate is assigned to CT and 0 otherwise (*Task*) on different sets of covariates:

- Candidates' characteristics:
  - *Female candidate* is a dummy variable that equals 1 if the candidate is female and 0 if the candidate is male.
  - *Signal (rank)* is a continuous measure of the relative ranking of the candidate in her pool according to the signal of performance.
  - *Age* is a continuous measure of candidates' age.
  - *STEM* is a dummy variable that equals 1 if the candidate is from *STEM* fields and 0 otherwise.

- Employers' characteristics:
  - *Supply first* is a dummy variable that equals 1 if the supply-side game was played first and 0 otherwise. Playing the supply-side game first may introduce a reinforcement of prior beliefs about gender differences in performance and affect the decisions of the employer.
  - *Male evaluator* is a dummy variable that equals 1 if the employer is male and 0 if the employer is female.
  - *Aware of outcomes gap* is a dummy variable that specifies whether the subject is aware of inequalities (1) or not (0).
  - *Aware of ability gap* is a dummy variable that specifies whether the subject is aware of ability (1) or not (0).
  - *Dscore>0* is a dummy variable that equals 1 if the employer score positively in the Implicit Association Test (IAT). A positive score predicts an implicit correlation between gender (male) and performance in math tasks.

- Decision environment:
  - *WorkerNumber* is a dummy variable that specifies the fixed position of the candidates in the decision screen.
  - *Period* is a dummy variable that specifies the period of the decision.

- *NWomen* is a continuous variable that specifies the number of women in each pool.
- *Topw* is a continuous variable specifying the number of women in the top-1 of each pool.

Table 5: specifications of the models.

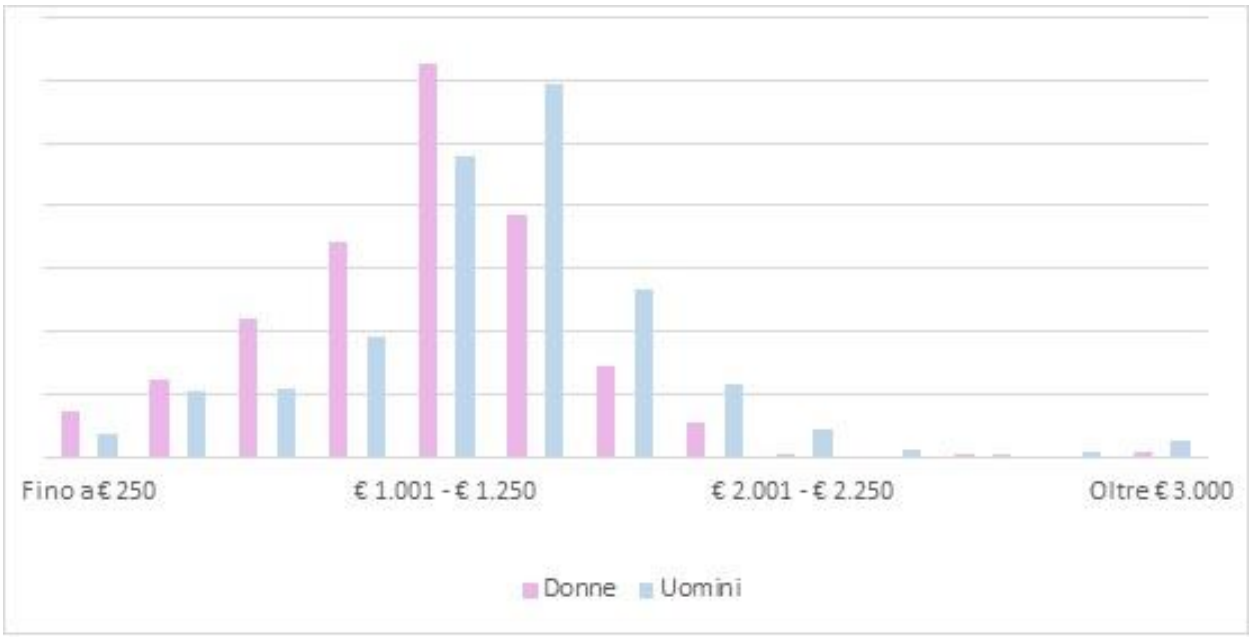
Table 5 analyzes the subjects' probability of choosing the tournament. The models in Table 5 regress a dummy variable that equals 1 if a subject chooses a tournament payment scheme in Stage 3 of the Supply-side game and 0 otherwise (*Choice*) on different sets of covariates:

- *Female* is a dummy variable that equals 1 if the subject is female and 0 if the candidate is male.
- *Performance (piece-rate)* is a continuous variable that measures subjects' performance in Stage 1 of the Supply-side game, where subjects were rewarded on a piece-rate basis.
- *Performance (tournament)* is a continuous variable that measures subjects' performance in Stage 2 of the Supply-side game, where subjects were rewarded on a tournament basis.
- *Engineering* is a dummy variable that equals 1 if the subject is from the School of Engineering and 0 otherwise.
- *Economics* is a dummy variable that equals 1 if the subject is from the School of Economics and Management and 0 otherwise.
- *Age* is a continuous measure of subjects' age.
- *Risk-lover* is a continuous variable that measures the degree of subjects' risk-loving. The higher the value of the variable, the higher their risk-loving attitude. In the post-experimental questionnaire, subjects were asked the following question: Do you consider yourself a person who prefers to take risks or avoid them? Please, indicate a number considering that 0 means "absolutely don't willing to take risks" and 10 means "absolutely willing to take risks".
- *Underconfident* is a dummy variable that equals 1 if the subjects' expected ranking about their performance in the addition task was below their actual ranking and 0 otherwise.
- *Dscore>0* is a dummy variable that equals 1 if the employer score positively in the Implicit Association Test (IAT). A positive score predicts an implicit correlation between gender (male) and performance in math tasks.

- *Supply first* is a dummy variable that equals 1 if the supply-side game was played first and 0 otherwise. Playing the supply-side game first may introduce a reinforcement of prior beliefs about gender differences in performance and affect the decisions of the employer.
- *Aware of outcomes gap* is a dummy variable that specifies whether the subject is aware of inequalities (1) or not (0).
- *Aware of ability gap* is a dummy variable that specifies whether the subject is aware of ability (1) or not (0).

Feedback provision

An example of feedback provision about question A (graduation grades) and question B1 (permanent contract) for students of Economics and Management.



According to the data of "Almalaurea" on graduates in Economics and Management at the University of Bologna (academic years 2010, 2011, and 2012), the average grade of male students like Giovanni and Hassan is 104/100 while the average duration of the completion of the undergraduate studies is of 2.7 years.

The average grade of female students like Angela is 107/100, obtained in 2.6 years on average.

Female students tend to graduate with better results than male students in a period slightly lower.

Of those students interviewed after one year of graduation, 17.5% of males employed have a long-term contract versus 17.1% of females.

INSTRUCTIONS

(This is an example of Feedback treatment)

Introduction

You are taking part in a decision-making study financed by the University of Bologna. During this study, you can earn an amount of money according to the rules that will be described in the following pages. The payment will be paid in cash and confidentially.

The present study will last around 90 minutes and is composed of 3 parts. You will be paid for one part randomly selected by the computer. So, your final earnings will be composed of the randomly selected part's earnings plus a 5 Euro show-up fee. The rules we will follow to determine your earnings are different in each part, and you will receive the instructions for each part sequentially. These instructions will describe in detail how your earnings will be determined in that part.

Communicating with other participants during the study is forbidden. The use of electronic devices will determine the exclusion for this study. If you have questions during the study, please raise your hand. An assistant will arrive at your station to answer privately.

Instructions Part 1

In this part, you will have to answer a questionnaire about inequalities in the labor market. We ask you to think carefully about your answers and to tell what you believe. If this part is selected for payment, you will receive a fixed payment of 10 Euros irrespectively of the correctness of your answers. You will receive information about the correctness of your answer at the end of this part.

What is happening now?

If there are questions, please raise your hand, and an assistant will come to your desk to answer. After that, we will start with this par

Instructions Part 2

In this part, we will assign you the firm's manager role. The firm is composed of one manager (you) and six workers. Your goal is to hire 3 of the six workers and assign them to two different tasks that differ in their difficulty and profitability for the firm. The first is called "Adding Task" and consists of summing up as many three three-digit numbers as possible in 6 minutes. The second task, "Problem Task," consists of solving as many mathematical problems as possible in 10 minutes. In this session, all participants play the role of managers. Workers were recruited in previous sessions, performed both tasks (the "Adding Task" and the "Problem Task"), and were paid according to their performance.

Your job is to:

1) select three workers out of 6 for your team;
2) assign two workers to the "Adding Task" and one worker to the "Problem task".

Your earnings as manager will be determined by the performance that the two workers have obtained in the "Adding task" plus the performance of the worker assigned to the "Problem task". Specifically, you will earn 10 cents per each addition correctly solved by the two workers assigned to the "Adding task" plus 1 Euro per each problem correctly solved by the worker assigned to the "Problem task".

Example:

- Imagine that you selected Worker 1, Worker 2, and Worker 3 for your firm and assign the first two to the "Adding Task" and the last one to the "Problem task",
- Imagine that Worker 1, and Worker 2 correctly solved 10 (C_{W1}) and 15 additions in the "Adding Task" (C_{W2}), respectively, and that Worker 3 correctly solved 6 (C_{W3}) problems in the "Problem task, your earnings will be computed as follows:

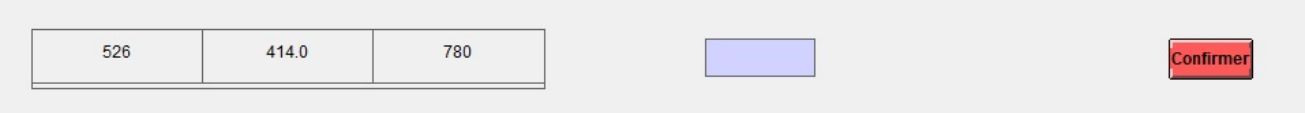
Earnings= 0.10 Euro x (C_{W1} + C_{W2}) + 1 Euros x (C_{W3})

Earnings= 0.10 Euro x (10 + 15) + 1 Euro x 6 = 2.5 + 6 Euro = 8.50 Euro.

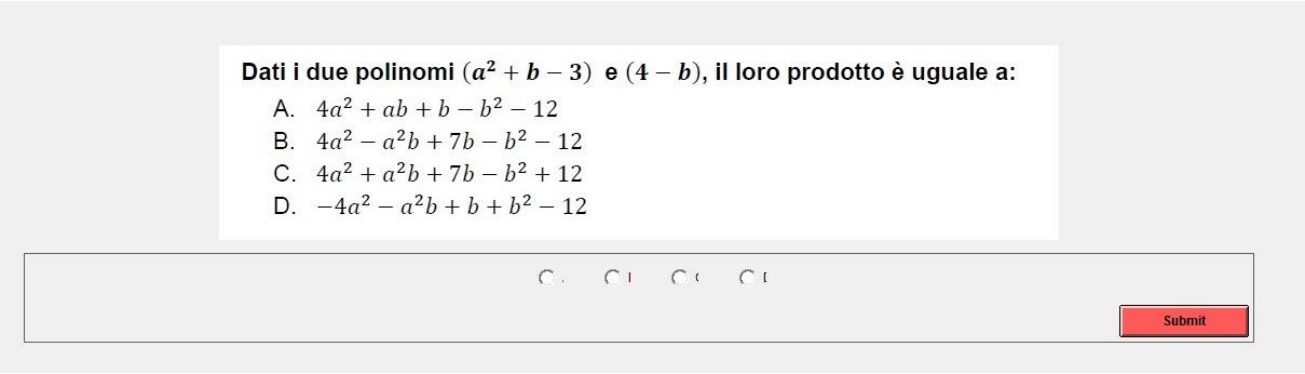
This task (i.e. the hiring decision and the task assignment decision) will be repeated four times. Every time you will face a new set of 6 workers.
If this part is selected for the payment, the computer will make another random draw and select one of the four repetitions, which will be used to determine your payment in part 2.

Workers' information
Before selecting workers, you will have the chance to look at their CVs by using the information relative to each worker that has already participated in the study. Apart from the information shown in the CV, you will receive information about the productivity of the workers in the first 3 minutes of Part 1 (number of correct calculations). We will call this information SIGNAL. You won't be provided with information relative to the performance in Part 2.

Before this part starts
In order to familiarize yourself with the tasks for which you have to make a decision, you will participate in a non-rewarded stage of the "Adding Task" (Screen 1) and the "Problem Task" (Screen 2).
In Screen 1, your task is to correctly solve the higher number of additions, as the workers did. As illustrated below, you will have 1 minute to solve additions of three three-digit numbers. The numbers to sum will be selected randomly. You will see a scheme like the one represented below.



Your task in Screen 2 will be to solve as many mathematical problems as possible correctly. You will have 2 minutes. You will see a scheme like the one presented below and have to select one of the possible answers.



Calculators or electronic devices are forbidden. Using the sheets of paper and the pencil on your desk is possible. When you are ready, you can insert your answer by choosing one of the answers and clicking the red button. Immediately, the computer will say if the answer is correct or not. Your answers are anonymous.
What is happening now?
If there are questions, please raise your hand, and an assistant will come to your desk to answer. Before starting this part, we will ask you to answer some questions to verify if you understood the rules correctly. After that, we will start with this part

Instructions Part 3. Part 3 is composed of 3 stages. If Part 3 is selected for the payment, the computer will make another random draw and select one of the three stages, which will be used to calculate your earnings in Part 3.

Instructions Part 3-Stage 1

Stage 1- Task

Your task in stage 1 is to solve correctly as many addition exercises as possible. To be more precise, you will have 3 minutes of time in order to solve as many additions of five randomly selected two-digit numbers as possible by entering the sum of the five numbers.

You are not allowed to use calculators, but you can write down the numbers and use the provided scribbling paper for your calculations. You enter an answer by clicking the "Confirm" button with the mouse. When you enter an answer, you are immediately told on the screen whether the latter is correct. Your answers are anonymous.

Earnings for Stage 1

In Stage 1, you get 50 cents per addition exercise you solve correctly in the 3 minutes. Your payment does not decrease if you provide an incorrect answer. We refer to this payment as the *Piece Rate* payment.

If Stage 1 is randomly selected for the final payment, you will obtain the money earned in Stage 1 under the *Piece Rate* payment.

Instructions Part 3 -Stage 2

Stage 2 - Task

As in stage 1, you will have 3 minutes of time to correctly solve as many addition exercises as possible. However, your payment in this stage depends on your performance relative to the performance of a group of participants. Groups are formed randomly at the beginning of this Stage, and each participant stays in the same group until the end of Part 3.

Allocation in groups

Each group consists of 6 participants, 3 of whom are men and 3 are women. Each group member receives an identification code. All members keep their identification code for the entire part 3. The three women randomly receive the identification code F1, F2, and F3. The 3 men randomly receive the identification code M1, M2, and M3. You will not discover the other participants' identities in your group during or after the study, so all decisions remain anonymous.

In Stage 2, your earnings depend on the number of additions you correctly solve compared to the five other people in your group. The two winners of the tournament are the two winners who correctly solve the largest number of addition exercises.

Earnings for Stage 2

Two winners (out of the 6 participants in the group) will receive 1.50 Euro per correct addition exercise, while the other participants do not receive any payment. In the case of a tie, the ranking among the members with equal performances is determined randomly. We refer to this as the *Tournament* payment. During the task and at the end of Stage 2, you will only be informed about the number of your correct answers, but you will not be informed of whether you were among the winners of the tournament until the study has been completed.

If Stage 2 is randomly selected for the final payment, then your earnings depend on whether you were among the two winners in the tournament or not.

Instructions Part 3- Stage 3

Stage 3- Task and Earnings

As in Stages 1 and 2, you will have 3 minutes of time to solve as many additions as possible correctly. However, you must now choose your preferred payment method for your performance in stage 3. You can choose the Piece-rate payment (as in Stage 1) or the Tournament payment (as in Stage 2).

To summarize:

- *Earnings in Stage 3 if you choose the Piece Rate payment scheme:*

you receive 50 cents per addition exercise you correctly solve.

- *Earnings in Stage 3 if you choose the Tournament payment scheme:*

if you choose the Tournament payment, your performance in Stage 3 will be evaluated in comparison to the performance of the other 5 group members in Stage 2 (Tournament).
As a reminder: That is the stage that you have just completed. If you enter more correct answers than four of your group members did in stage 2, you will receive €1.50 per correct answer (i.e., three times the Piece-rate payment). In other words, only one member of your group can have a Stage 2 performance higher than your Stage 3 performance; otherwise, you receive no payment for this stage. In case of a tie, the ranking among the members with equal performances is again determined randomly. The group composition (three men and three women) is in Stage 2. If you choose the Tournament payment, you will not be informed about the tournament's outcome until the study end.

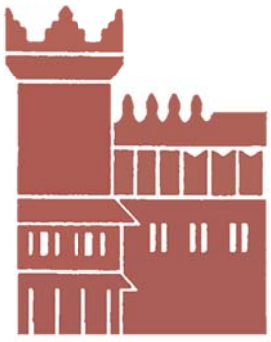
On the next screen, you will be asked whether you want to choose the Piece-rate payment or the Tournament payment for your performance in Stage 3. Afterward, you will have 3 minutes in order to calculate the sums of the two-digit numbers.

If Stage 3 is randomly selected for the final payment, then your earnings depend on the choice of the payment and your performance in this task.

What is happening now?
If there are questions, please raise your hand, and an assistant will come to your desk to answer.
Before starting this part, we will ask you to answer some questions to verify if you understood the rules correctly.
After the control questions, you will have one minute in order to familiarize yourselves with the task. During this time, you can solve addition exercises, which do not count for your earnings. Once the practice time is terminated, you will be informed when Stage 1 is going to start.

Instructions Part 4 [displayed on the screen only]
Part 4 is composed of 2 stages:
1) an association test, and
2) a general questionnaire.

If this part is selected for payment, you will receive a fixed payment of 10 Euros for both stages (the questionnaire and the association test).
What is happening now?
If there are questions, please raise your hand, and an assistant will come to your desk to answer.
After that, we will start with this part.



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it