ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Breaking Stereotypes: How Valuing Workers' Preferences Improves Task Allocation and Performance

Giulio Ecchia
Natalia Montinari
Raimondello Orsini
Sveva Vitellozzi

Quaderni - Working Paper DSE N°1202

Breaking Stereotypes: How Valuing Workers' Preferences Improves Task Allocation and Performance[1]

Giulio Ecchia[2], Natalia Montinari[3], Raimondello Orsini[4], Sveva Vitellozzi[5]

Abstract

Firm performance depends critically on the efficient allocation of tasks across employees. Yet, task assignment decisions are often shaped not only by productivity considerations but also by managerial biases and gender stereotypes—frequently resulting in women being disproportionately assigned low-promotability, female-stereotyped tasks. This paper investigates whether making workers' task preferences visible to managers can reduce gender-stereotypical assignments and improve overall outcomes. We conduct two complementary experiments. In the first, participants act as workers, completing real-effort tasks and reporting their task preferences. In the second, a separate group of participants from the same subject pool takes on the role of managers and assigns tasks to pairs of workers under varying information conditions. In the control condition—where managers lack access to workers' preferences—task assignments are more likely to reflect gender stereotypes. In contrast, when managers are informed of workers' preferences, stereotypical assignments decrease, and managerial earnings improve. We also find that preference-informed task allocation leads to higher managerial earnings, suggesting that reducing gender bias not only promotes fairness but also enhances organizational efficiency. Our findings highlight the potential of low-cost informational interventions to promote fairer and more effective task allocation practices.

JEL codes: J16, J71, M12, C91

Keywords: Gender stereotypes, task allocation, managerial decision-making, labor market discrimination, experimental economics

[11]We thank Josè Dominguez and Irene Maria Buso for excellence research assistance. We also thank the participants of the 17th workshop on Social Economy for useful comments. The authors are grateful to "Fondazione Cassa dei Risparmi di Forlì" for the financial support on the project.

[2] University of Bologna, Department of Economics, email: giulio.ecchia@unibo.it

[3] University of Bologna, Department of Economics, email: natalia.montinari2@unibo.it

[4] University of Bologna, Department of Economics, email: raimondello.orsini@unibo.it

[5] Corresponding Author: University of Bologna, Department of Economics, email: sveva.vitellozzi@unibo.it

Non-Technical Abstract

In many workplaces, tasks are rarely assigned at random. Managers and supervisors often decide who does what, based on a combination of observed skills, perceived strengths, and informal judgments. However, these judgments are not always neutral. A growing body of research shows that gender stereotypes -such as the idea that men are naturally better at analytical tasks while women excel in communication or support roles- can shape these decisions in subtle but impactful ways. As a result, equally qualified women and men can end up doing very different types of work, with very different career consequences.

This study explores whether giving decision-makers information about workers' actual preferences can reduce the influence of stereotypes and lead to more efficient and equitable task allocation. In many cases, managers do not know what their employees prefer to do—and in the absence of this information, they may fall back on assumptions shaped by social norms or group averages. But what happens if we make preferences visible?

To investigate this question, we conducted a laboratory experiment that mimics real-life task allocation. Participants were asked to perform two types of tasks—one involving math and logic (typically seen as "masculine"), and another involving word associations (often viewed as more "feminine"). After completing the tasks, participants were asked which one they preferred. Then, in the second phase, other participants were put in the role of "managers" and asked to assign one of the two tasks to each worker. Some managers were shown only the workers' names and past performance; others also saw the workers' stated preferences.

We found that in the absence of preference information, managers tended to assign tasks in ways that aligned with traditional gender stereotypes. Female workers were more likely to be assigned the verbal task, while male workers were more often chosen for the math task—even when their performance levels were similar. However, when managers were shown the workers' task preferences, they were significantly more likely to take those preferences into account and less likely to rely on gendered assumptions. This shift in allocation patterns improved overall performance and increased the managers' own earnings, suggesting that more accurate, individualized information leads to better outcomes for everyone.

Our findings offer several key insights. First, stereotypes can persist even when objective performance data is available; they serve as a cognitive shortcut in the absence of more specific information. Second, revealing workers' preferences is a simple but powerful intervention. It helps decision-makers tailor assignments more accurately, which improves productivity and reduces inequality. Third, taking preferences seriously does not come at a cost -it can actually enhance efficiency and organizational performance.

This research speaks directly to current debates about workplace fairness and inclusion. It shows that reducing gender bias does not necessarily require expensive training programs or radical organizational overhauls. Often, it is a matter of changing the information environment. By making individual preferences visible, we empower managers to move beyond stereotypes and make more meritocratic decisions. At the same time, we give workers a greater sense of agency and recognition -both of which are crucial for motivation and engagement.

In short, this study suggests that transparency matters. When we take the time to understand what people want to do -and not just what we assume they should do- we can build more inclusive, efficient, and fair workplaces. This is a lesson that applies not only to gender, but to all kinds of diversity in the labor market.

Breaking Stereotypes: How Valuing Workers' Preferences Improves Task Allocation and Performance

1. Introduction

Firm performance is intricately linked to the effective allocation of tasks among employees. On the one hand, the division of labour and allocation of tasks are crucial in ensuring that organizational goals are met efficiently (Delfgaauw et al., 2020). Assigning more important tasks to more talented employees has been shown to enhance firm performance. However, task allocation is often influenced by more than just performance considerations. Managers' biases, stereotypes, and personal preferences frequently shape decisions, leading to allocations that may deviate from efficiency. For instance, evidence suggests that women are more likely than men to be assigned low-promotability tasks—tasks that are essential to the organization but offer limited career advancement opportunities. This phenomenon may result both from demand-side factors, where managers disproportionately assign such tasks to women, and from supply-side factors, where women are more likely to accept such requests (Babcock et al., 2017a; Babcock et al., 2017b).

A mismatch between employees' task preferences and the tasks they are assigned can significantly undermine performance. While Kamei and Markussen (2023) do not study task assignment directly, they show that simply allowing team members to express and sort based on their preferences (without managerial intervention) significantly impacts team performance. Their findings suggest that even in the absence of formal allocation, individual preferences matter for the success of the group. These challenges—particularly the central role that task assignment plays in shaping organizational performance—are further exacerbated by gender stereotypes, which influence perceptions of competence and behavior. Men, for example, are often perceived to outperform women in scientific disciplines, under competitive environments, and in their willingness to compete. Conversely, under stereotypically female tasks, women's performance relative to men's may improve, and their willingness to compete may increase (Halladay and Landsman, 2022). Such stereotypes perpetuate inefficiencies in task assignments and can hinder optimal outcomes. Experimental evidence shows that even when people recognize the productivity potential of diverse teams, they underestimate how much others value that diversity, leading to persistent misallocation and gender-based sorting (Fischbacher, Kübler, and Stüber, 2022). Further, when contributions are difficult to observe—as is often the case in team settings—women are less likely to receive credit for group achievements, especially when collaborating with men, reinforcing gender gaps in recognition and advancement (Sarsons et al., 2021). In many organizational contexts and, more generally, in team settings, task assignments are often influenced by stereotypes rather than objective assessments of skill or interest. This not only risks undermining

individual motivation and performance but can also impair team dynamics and overall company outcomes. Introducing mechanisms that make employees' preferences visible to decision-makers can help disrupt this pattern, allowing for more meritocratic and efficient task allocation.

Given the significance of task allocation in shaping organizational performance, a pertinent question arises: does providing managers with information on workers' task preferences encourage gender-neutral task assignments and, by extension, improve overall performance? Addressing this question is vital for understanding how to mitigate the adverse effects of biases and stereotypes in the workplace while fostering a more equitable and effective organizational environment. This study seeks to contribute to the growing literature on task allocation and performance by investigating the impact of preference-informed managerial decisions on gender-neutral task assignments and worker performance.

We find that making workers' preferences visible to employers significantly reduces gender-stereotypical task assignments, particularly in situations where preferences provide informative guidance—such as when both workers are equally productive but express opposite preferences. In these cases, preference information helps employers deviate from gender stereotypical associations and assign tasks more equitably. However, when preferences are less informative—for instance, when both workers prefer the same task—employers tend to revert to stereotypical patterns, disregarding the counter-stereotypical choice regardless of the worker's gender. Overall, the disclosure of preferences not only reduces bias in task allocation but also leads to more efficient matches, translating into higher earnings for employers. These findings highlight the potential of low-cost informational interventions to improve both fairness and organizational performance. Our study contributes to two distinct but related strands of the literature. First, it builds on the growing body of evidence highlighting the positive effects of employees sorting into payment schemes and contracts that align with their preferences. Existing research has documented that alignment between preferences and workplace structures can enhance motivation, productivity, and firm performance (Dohmen & Falk, 2011). Recent experimental studies reinforce the importance of aligning work tasks with worker preferences. For example, while Kamei and Markussen (2023) do not study task assignment per se, they show that allowing individuals to express and sort based on their task preferences significantly affects effort provision and reduces free riding within teams. Their findings highlight that even in the absence of managerial intervention, preference alignment plays a key role in group outcomes. Similarly, Dasgupta et al. (2024) demonstrate that granting workers task autonomy increases effort, particularly when individuals have strong task preferences. While these studies emphasize the benefits of preference alignment through self-selection or autonomy, our work introduces a complementary perspective: we explore the implications of managerial allocation—where decision authority remains with the manager—but where preference information is made salient. This allows us to isolate the impact of information disclosure on both

efficiency and fairness, especially in the context of potential gender biases. By identifying a new channel—the role of managers being informed about workers' preferences—this study sheds light on how such information can improve managerial decision-making and reduce gender bias in task allocation. This provides actionable insights into how organizations can operationalize preference-based task assignments to achieve greater efficiency and equity.

Second, this study contributes to research on the impact of employers' gender stereotypes in the workplace. A large body of literature highlights how both descriptive and prescriptive gender stereotypes can distort managerial decisions related to task assignments, evaluations, promotions, and leadership opportunities. Descriptive stereotypes—generalizations about what women and men are like—create biased expectations about competence and performance, particularly when women are perceived to "lack fit" with stereotypically male tasks or roles (Heilman, 2012). These biases are especially likely to manifest in ambiguous settings, such as when performance is difficult to measure or when evaluative criteria are ill-defined. Prescriptive stereotypes—norms about how women and men should behave—further penalize women for engaging in agentic or leadership behaviors, even when they are effective, thus reinforcing inequitable patterns in organizational advancement (Heilman, 2001). While this literature has thoroughly documented how gender stereotypes operate, practical interventions to mitigate their effects remain underexplored. By demonstrating that making employees' preferences visible to managers can attenuate the influence of gender stereotypes on task allocation, our study offers a simple, implementable mechanism that promotes fairer and more efficient decision-making processes.

The remainder of the paper is structured as follows: Section 2 describes the experimental design; Section 3 outlines our pre-registered hypotheses; Section 4 presents the results; and Section 5 concludes.

2. Experimental Design

The experiment consisted of two stages, both involving participants from the University of Bologna's BLESS laboratory subject pool. In the first stage—the Workers Experiment—participants acted as "workers" and were invited to the laboratory to complete two distinct real-effort tasks: a verbal task (solving anagrams) and a math task (adding three-digit numbers). Full details of this stage are provided in Section 2.1.

In the second stage—the Employers Experiment—a new set of participants from the same subject pool took part in an online experiment as "employers." Each employer was asked to assign one of the two real-effort tasks (verbal or math) to each member of a pair of workers from the first stage. To support their allocation decisions, employers were shown information about each worker's socio-demographic

characteristics, a signal of their task performance, and—in the preference treatment—their self-reported task preference. Full details of this stage are provided in Section 2.2.

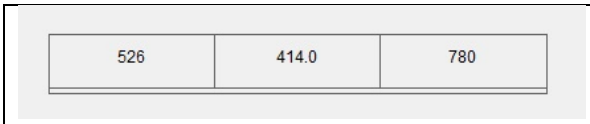
The English version of the experimental instructions used in both stages is available in Appendix B.

2.1 The Workers Experiment

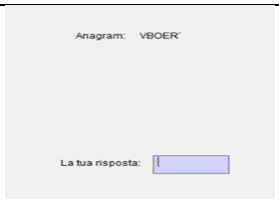
The Workers Experiment consisted of three sequential parts, followed by a final questionnaire. In Part 1, participants completed two real-effort tasks, presented in random order: a math task, involving the addition of three-digit numbers, and a verbal task, involving the formation of five-letter words from scrambled letters (anagrams). Examples of both tasks are shown in Figure 1.

The Addition task, commonly used in experimental economics (e.g., Niederle & Vesterlund, 2007; Sloof & van Praag, 2010; Eckartz et al., 2012; Lezzi et al., 2015), captures quantitative skills and is often associated with math-related gender stereotypes—where men are perceived to perform better than women, despite little to no actual difference in performance. Conversely, the Anagram task reflects verbal ability, typically linked to the humanities, and is stereotypically associated with higher performance by women. Yet, similar to math tasks, there is no consistent evidence of gender-based performance differences in verbal tasks (Coffman, 2014). These tasks were deliberately chosen to explore potential gender stereotypes around numerical and verbal competencies in the workplace.

Figure 1. Example of the real effort tasks used.



a) Math Task: additions



b) Verbal Task: anagrams

In both tasks, a new problem was generated each time a participant submitted a response or skipped a question by clicking the "submit" button without answering. Immediate feedback was provided on the correctness of each response. Each task lasted three minutes, yielding a total of six minutes for Part 1. Participants were compensated according to a piece-rate scheme: €0.30 per correct addition and €0.20 per correctly solved anagram. This phase served two purposes: it familiarized participants with the task structure and provided a baseline measure of individual ability.

After completing Part 1, participants were asked to indicate their preferred task, with the information that this choice would influence later stages of the experiment.

In Part 2, participants performed both tasks again, presented in random order and under the same incentive structure as in Part 1. This phase allowed us to assess learning effects and investigate whether participants improved when working on their preferred task.

In Part 3, participants completed only their self-selected preferred task for an additional three minutes, again under the same piece-rate scheme.

Following the task rounds, participants were asked to evaluate their own performance by estimating the number of correctly solved problems in each task. Correct self-assessments were rewarded with €1 each. They were also asked to rank themselves relative to their peers for both tasks.

Final earnings were determined by randomly selecting one of the three parts (Part 1, 2, or 3) for payment. If Part 1 or Part 2 was selected, a second random draw determined which of the two tasks would be used to calculate earnings. Participants received individualized feedback on their performance in each part and were informed about which part was selected for payment. In accordance with BLESS Lab procedures, all payments were processed securely via PayPal.

After the experimental tasks, participants completed a post-experimental questionnaire collecting demographic information, including age, gender, and field of study. This data was used to construct each participant's curriculum vitae (CV) for the second stage of the experiment—the Employers Experiment. A summary of the Workers Experiment timeline is provided in Table 1.

Table 1: Workers' Experiment- timeline

Part	Activity	Duration
Part 1	Addition and anagrams in random order	3 minutes for each task, 6 minutes in total
Elicitation of Preferred Task	Elicitation of participants' preferred task	-
Part 2	Addition and anagrams in random order	3 minutes for each task, 6 minutes in total
Part 3	Preferrred Task only	3 minutes
		-
Final questionnaire		
Feedback	Participants received detailed feedback on their performance and earnings across all parts. For each part, feedback included results for both tasks.	
Payment Structure	Participants received a €5 show-up fee, plus earnings from a randomly selected task within a randomly selected part, along with any additional compensation based on the outcome of the beliefs elicitation which was paid for sure.	

2.1.1 Experimental procedures

The Workers Experiment was conducted with 124 participants recruited from the BLESS subject pool at the University of Bologna using the ORSEE recruitment system (Greiner, 2015). The experiment was programmed and implemented using zTree (Fischbacher, 2007). Upon arrival at the laboratory, participants were randomly assigned to individual workstations. The session lasted approximately 45 minutes, including a public reading of the instructions to ensure common understanding. Participants received a €5 show-up fee, and average total earnings—including performance-based compensation—were approximately €11.

2.2 The Employers Experiment

The second stage of the experiment involved a new set of participants, recruited from the same BLESS subject pool, who took part in an online experiment in the role of *employers*. Their task was to assign two tasks—addition or anagram—to a pair of *workers* (i.e., participants from the first stage) based on information provided about each worker.

For each worker, employers received information on gender, age, field of study, and a performance signal for both tasks. In the preference treatment (between-subjects), employers also received information about each worker's stated task preference. The experiment followed a between-subjects design: employers were randomly assigned to either a baseline condition (no preference information) or the preference treatment.

Each employer completed 12 decision rounds, each involving a new pair of workers. Employers made 8 decisions involving mixed-sex pairs and 4 involving same-sex pairs, presented in randomized order. Further details on the construction of worker pairs are provided in Appendix A.1.

Given our focus on gender bias and preference alignment, we are particularly interested in employer behavior in mixed-sex pairs. These were organized into four experimental groups:

1. **Group 1: Counter-stereotypical preferences, equal ability**: Female workers prefer the addition task, while male workers prefer the verbal task. Both workers have equal productivity across tasks. Each employer saw two of three possible pairs (see Table A1-pnel 1 in the Appendix).
2. **Group 2: Female counter-stereotypical, male stereotypical, equal ability**: Female workers prefer the addition task; male workers prefer the math task as well. Productivity is equal across tasks. Each employer saw two of three possible pairs (see Table A1-pnel 2 in the Appendix).
3. **Group 3: Male counter-stereotypical, female stereotypical, equal ability**: Male workers prefer the verbal task; female workers prefer the verbal task as well. Productivity is equal across tasks. Each employer saw two of three possible pairs (see Table A1-pnel 3 in the Appendix).

4. **Group 4: Counter-stereotypical preferences with asymmetric ability**: Female workers prefer and outperform males in the addition task; male workers prefer the verbal task. Each employer evaluated both pairs (see Table A1-pnel 4 in the Appendix).

Same-sex worker pairs (Groups 5–8) served as comparative controls. Each group contained two worker pairs, and employers were randomly assigned one of them (see Tables A5–A8).

Participants were informed that one of their 12 decisions would be randomly selected for payment at the end of the experiment. Since the selected round was unknown in advance, employers were instructed to treat each decision independently and make choices that reflected their true preferences and strategic considerations throughout the session.

2.2.1 Experimental Treatments

The experiment featured two between-subjects treatments, differing only in the information provided to employers about the workers:

- **Baseline Treatment (BT):** Employers were shown workers' socio-demographic characteristics (gender, age, and field of study) and performance signals for both tasks.
- **Preference Treatment (PT):** In addition to the information provided in the baseline, employers were also shown workers' stated task preferences.

Figure 2 provides an example of how information was displayed to employers. The only difference between treatments is the inclusion of workers' self-reported preferences in the PT condition.

Figure 2. Pair a) from Group 1. The information about workers' preferences is only displayed in the Preference Treatment.

Male worker

Correct n. math task	6
Correct n. verbal task	7
Age	25
Field of study	Archeology
Preferred task	Verbal task

Female worker

Correct n. math task	6
Correct n. verbal task	7
Age	25
Field of study	Italian studies
Preferred task	Mathematical task

These treatments allow us to assess whether explicit knowledge of task preferences influences employer behavior, particularly in the presence or absence of informative performance signals.

In Groups 1–3, both workers are equally productive in both tasks. The performance signals, therefore, offer no guidance for task allocation, making these groups ideal for testing whether employers' default to gender-stereotypical assignments in the absence of performance-based justification. If the preference treatment reduces stereotypical allocations in these groups, it would suggest that employers rely on stereotypes only when lacking more salient, preference-based cues.

Group 4 introduces an informative performance signal: the female worker has a clear advantage in the addition task, while productivity is equal in the verbal task. This group allows us to test whether employers continue to follow stereotypes even when presented with counter-stereotypical performance evidence. If employers persist in assigning the addition task to the male worker, despite the female worker's stronger performance, it would indicate that gender biases override objective indicators. The preference treatment in this group allows for an additional test: whether explicit preferences, when aligned with productivity, are enough to correct for stereotype-driven misallocation.

Together, these treatment comparisons offer insight into how different types of information—demographic, performance-based, and preference-based—interact to influence task allocation decisions, and whether gender biases can be mitigated when employers are provided with richer contextual information.

2.2.2. Experimental procedures

The Employers Experiment was conducted with 385 participants, recruited from the BLESS subject pool using the ORSEE recruitment platform (Greiner, 2015). The experiment was programmed in Qualtrics and administered online.

Participants who registered for a session received a personalized link and were given 48 hours to complete the experiment. On average, participants completed the experiment in 13.37 minutes (*SD* = 4.62). Compensation consisted of a €3 show-up fee plus performance-based earnings, with an average total payment of €8.40.

3. Research questions and hypotheses

Based on our experimental design, we pre-registered and tested the following hypotheses[6]:

Hypothesis 1: Stereotypical Task Assignment

In the baseline treatment, where employers do not receive information about workers' task preferences, we hypothesize that employers will be more likely to assign tasks in line with gender stereotypes. Specifically, female candidates will be more frequently assigned to the anagram task, while male candidates will be more frequently assigned to the addition task.

[6] The experimental design and hypotheses were pre-registered using the platform as predicted. #158973 - Task Assignment: the role of Employers' Stereotypes and Employees' Preferences, https://aspredicted.org/f2r4-56v2.pdf

This hypothesis tests whether stereotypes drive task allocation decisions in the absence of explicit preference information. Prior research shows that even when there are no actual gender differences in task performance (Coffman, 2014), women are perceived to be less capable in math-related domains (Niederle & Vesterlund, 2007) and more capable in verbal ones. These perceptions reflect well-documented descriptive stereotypes (Heilman, 2012), which can lead to biased evaluations—especially in ambiguous decision contexts. Moreover, when workers are equally productive across tasks, reliance on gendered assumptions may serve as a cognitive shortcut, reinforcing occupational segregation and constraining access to promotable or skill-enhancing work (Babcock et al., 2017a; Fischbacher et al., 2022).

HP2: Between-Treatment Effect
In the preference treatment, where employers do receive information about workers' stated task preferences, we expect that counter-stereotypical preferences—such as men preferring anagrams and women preferring addition—will be less likely to be accommodated compared to stereotypical preferences. We expect that this effect will be stronger for female candidates, meaning that women's counter-stereotypical preferences (choosing the addition task) will be disregarded more often than men's counter-stereotypical preferences (choosing the anagram task).

This hypothesis examines whether explicit preference information mitigates or reinforces gender-biased task assignments. If stereotypes are merely a substitute for missing information, employers may respond positively to preference cues. However, if stereotypes are prescriptive (Heilman, 2001), women who deviate from traditional roles may face greater penalization than men who do the same—reflecting a gender asymmetry in norm violations. This is consistent with recent findings that women are often undervalued or under-credited in collaborative settings, especially when contributions are difficult to evaluate (Sarsons et al., 2021). If employers continue to disregard women's preferences even when informed, this would suggest that task segregation is driven not just by workers' self-selection, but by persistent and deep-rooted biases in managerial decision-making.

4. Data

In this section we present summary statistics about the workers' experiment (Section 4.1) and the employees' experiment (Section 4.2).

4.1 The Workers' Experiment

Table 2 reports summary statistics for the 124 participants in the Workers Experiment. The sample was well balanced in terms of gender, with 52.5% female participants, and had an average age of 23 years

(*SD* = 2.655). Performance in each round was measured by the number of correctly solved additions (mathematical task) and correctly completed anagrams (verbal task). From Part 1 to Part 2, participants significantly improved their performance in both tasks (Wilcoxon matched pairs signed-rank test: p = 0.000 for both tasks), suggesting a learning effect over time. On average, participants in the role of employees were more productive in the verbal task than in the mathematical task, both in Part 1 and Part 2 (Wilcoxon matched-pairs signed-rank test: p = 0.000 in both cases). After completing both tasks in Part 1, and before starting Part 2, 64.5% of participants (80 out of 124) expressed a preference for the verbal task.

Table 2. Summary statistics of workers' sample

	N	Mean	Median	SD
Part 1 – math: # correct answers	124	5.669	6	2.248
Part 2 – math: # correct answers	124	6.96	7	2.552
Part 3 – math: # correct answers (chosen)	44	7.204	7	2.741
Part 1 – verbal: # correct answers	124	10.5	10	5.902
Part 2 – verbal: # correct answers	124	14.605	14	7.982
Part 3 – verbal: # correct answers (chosen)	78	21.875	21	8.370
Choice preferred task (=1 verbal)	124	0.645	1	0.48
Earnings math task – rounds 1 and 2	124	3.788	3.75	1.268
Earnings verbal task – rounds 1 and 2	124	5.02	4.6	2.600
Earnings math task (chosen)	44	2.16	2.1	0.82
Earnings verbal task (chosen)	78	4.37	4.2	1.67
Total earnings from experimental tasks	124	12.4	11.5	4.879
Female (=1)	122	0.525	.5	0.549
Age of participant in the role of workers	122	23.59	23	2.655
Recipient of Scholarship (1=yes, 0=no)	122	1.311	1	0.465
Competitiveness (Self-reported, 0-10)	122	7.795	8	2.459
Confidence (Self-reported, 0-10)	122	7.648	8	2.367
Patience (Self-reported, 0-10)	122	7.787	9	2.716

Note. In the post-experimental questionnaire, we lost N=2 observations on participants who skipped some questions.

Table 3 presents average productivity by gender across both tasks. The results reveal no significant differences in performance between men and women, in either the verbal or the mathematical task. This absence of a gender performance gap is particularly important: it confirms that any observed gender-based differences in employer task assignment decisions cannot be attributed to underlying productivity differences. In this sense, the experimental environment provides an ideal setting to investigate the role of gender stereotypes in shaping task allocation.

Table 3: Mean difference tests in performance between men and women (Workers sample)

	Male	Mean	Female	Mean	Diff	St Err	p-value
Part 1 – math: # correct answers	61	5.787	58	5.569	0.218	0.414	0.599
Part 2 – math: # correct answers	61	7.312	58	6.707	0.605	0.471	0.202
Part 3 – math: # correct answers	23	2.738	20	2.431	0.306	0.705	0.664
Part 1 – verbal: # correct answers	61	11.213	58	10	1.213	1.087	0.267
Part 2 – verbal: # correct answers	61	15.639	58	13.431	2.208	1.468	0.135
Part 3 – verbal: # correct answers	38	15.59	38	12.449	3.142	2.303	0.175
Verbal task preferred (=1 verbal task)	61	.623	58	0.655	-0.032	0.089	0.718

Note. The last column report p-values from a set of Mann Whitney tests comparing the performance of men and women from the workers sample in each task and in each part.

4.2 The Employers' Experiment

Table 4 reports descriptive statistics for the sample of participants in the Employers Experiment (N = 385). The sample is evenly balanced by gender, with 51.26% female participants (mean of the binary female variable = 0.51). The average age of participants is 24.4 years (*SD* = 3.468). In addition to demographic characteristics, Table 4 also summarizes information on participants' personality traits and gender-stereotypical beliefs, which were collected to explore potential heterogeneity in employer behavior.

Table 4: Summary statistics of employers' sample

	N	**Mean**	**Median**	**SD**
Female (=1, 0 if male)	385	0.51	1	0.500
Age	385	24.418	24	3.468
Highschool (=1)	385	0.857	1	0.350
Bachelor (=1)	385	0.288	0	0.454
TIPI: 5 personality traits				
Extraversion	385	4.07	4	1.43
Agreeableness	385	4.929	5	1.07
Conscientiousness	385	5.36	5.5	1.181
Emotional stability	385	4.154	4	1.358
Openness	385	4.94	5	1.115
Risk loving (Self-reported, 1-10)	385	5.953	6	1.935
Competitiveness (Self-reported, 1-10)	385	6.291	7	2.297
Boys better in STEM	385	3.397	3	2.736
Girls better in arts	385	3.709	3	2.784
Boys better leaders	385	2.868	1	2.48
Girls better caregivers	385	3.662	3	2.916

Participants were also assessed using the Ten-Item Personality Inventory (TIPI), a brief psychological measure designed to evaluate the Big Five personality traits: Extraversion, Agreeableness, Conscientiousness, Emotional Stability (Neuroticism – reversed), and Openness to Experience (Gosling et al., 2003). The table highlights participants' explicit gender-stereotypical beliefs, with responses on a

scale from 0 to 5 capturing agreement with statements such as “Boys are better in STEM” (mean = 3.397) and “Girls are better caregivers” (mean = 3.662). Overall, the sample exhibits a mix of traditional and modern views on gender and a diversity of personality traits.

Table A.2 in Appendix reports mean difference tests between participants in the treatment and control group. Inspection of the table evidences no significant differences between the two groups, suggesting that the randomization worked properly.

5. Results

This section presents the core results of our experiment. In Section 5.1, we assess both Hypothesis 1 and Hypothesis 2, exploring how task preferences interact with gender stereotypes and whether counter-stereotypical preferences are treated differently depending on the worker’s gender. Finally, in Section 5.2, we analyze the effect of preference information on employer earnings, evaluating whether more equitable task assignments also lead to more efficient outcomes.

5.1 Task assignment

To analyze the frequency of stereotypical task assignments, as outlined in Hypothesis 1, we focus on the four mixed-gender groups described in Section 2.2. These groups allow us to examine how employers allocate tasks when faced with either ambiguous or informative productivity signals, and whether gender stereotypes influence these decisions. The data are analyzed by treatment to assess whether making workers' task preferences visible mitigates reliance on stereotypes.

We define a stereotypical task assignment as the employer assigning the female worker to the verbal task and the male worker to the addition task, regardless of actual productivity signals. This definition is applicable in two distinct contexts:

- Groups 1–3: In these groups, both workers are equally productive in both tasks, meaning the productivity signal is uninformative. Here, a high frequency of stereotypical assignments would suggest that employers rely on gender stereotypes in the absence of objective performance differences, potentially disregarding workers’ preferences.
- Group 4: In this group, the female worker has a higher productivity signal in the addition task, creating a clear performance-based justification for assigning her to the math task. Persisting with stereotypical allocations in this setting would indicate that employers’ biases override objective productivity considerations, even in the presence of counter-stereotypical performance evidence.

Preference treatment plays a critical role in interpreting these patterns. In Groups 1–3, if stated preferences reduce the frequency of stereotypical assignments, this would suggest that stereotypes serve as a default heuristic in the absence of alternative information. In Group 4, the interaction

between productivity signals and stated preferences allows for a stronger test: whether employers adjust their decisions when preferences and performance indicators are aligned, or whether biases persist despite both sources of counter-stereotypical information.

Table 5 reports the share of stereotypical task assignments across the four mixed-gender groups, disaggregated by treatment and calculated at the employer level. The last row presents results from a set of Mann-Whitney tests comparing the treatments. For a more detailed view, Appendix Table A.1 replicates these results, breaking them down at the pair level within each group.

Table 5. Fraction of stereotypical choice in each group

Treatment	Group 1	Group 2	Group 3	Group 4	Overall
Preferences	Both counter-stereotypical: female prefers math task, male prefers verbal task	Female workers counter-stereotypical: both prefers the math task	Male workers counter-stereotypical: both prefers the verbal task	Both counter-stereotypical: female prefers math task, male prefers verbal task	
Ability	No differences in signal for both tasks	No differences in signal for both tasks	No differences in signal for both tasks	Female workers more productive in math, no differences in verbal tasks	
Baseline (BT)	.481	.714	.685	.311	.548
Preferences (PT)	.235	.710	.677	.127	.436
Mann-Whitney test (BT vs PT)	0.000	0.928	0.762	0.000	0.000

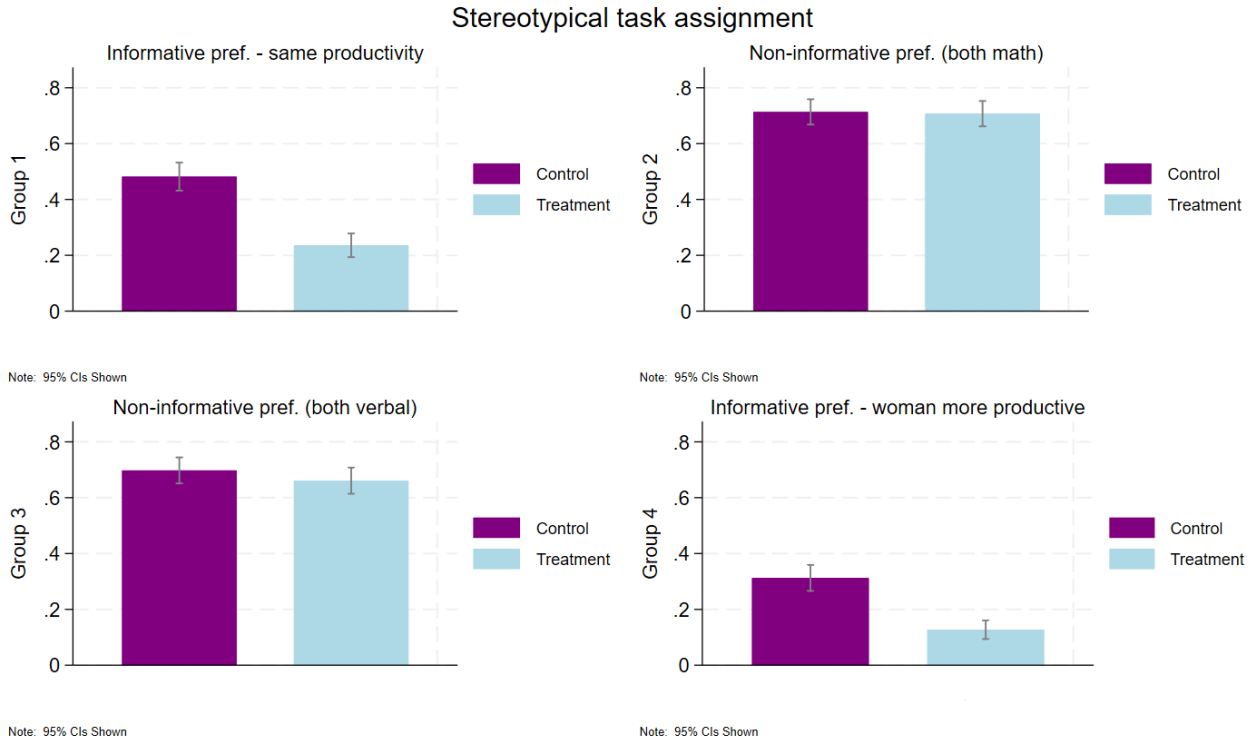
Note. Table 5 reports the share of stereotypical task assignments across the four mixed-gender groups, disaggregated by treatment and calculated at the employer level. For Groups 1–3, each employer evaluated a random selection of 2 out of 3 worker pairs; in Group 4, all employers evaluated 2 predefined pairs (see Table A.1 in the Appendix). The last row presents results from a set of Mann-Whitney tests comparing the treatments..

This analysis enables us to evaluate the extent to which gender stereotypes influence task allocation decisions, even when objective productivity signals are available. We then assess whether providing employers with information about workers' task preferences helps mitigate these biases.

When task preferences are made visible, the overall frequency of stereotypical task assignments decreases significantly—from 55% in the baseline treatment to 44% in the preference treatment (Mann-Whitney test, $p < 0.000$). The reduction is particularly pronounced in two key cases. First, in Group 1, where workers have identical productivity signals but express opposite, counter-stereotypical preferences, the share of stereotypical assignments drops from 48% to 24% when preferences are revealed (Mann-Whitney test, $p < 0.000$). Second, in Group 4, where the female worker is objectively more productive in the math task, stereotypical assignments decline from 31% to 13% under the preference treatment (Mann-Whitney test, $p < 0.000$).

These findings, also illustrated in Figure 3, suggest that explicit preference information can contrast gender biases in task assignment, particularly when it reinforces or clarifies existing productivity signals.

Figure 3: Fraction of stereotypical task assignment by treatment at the group level

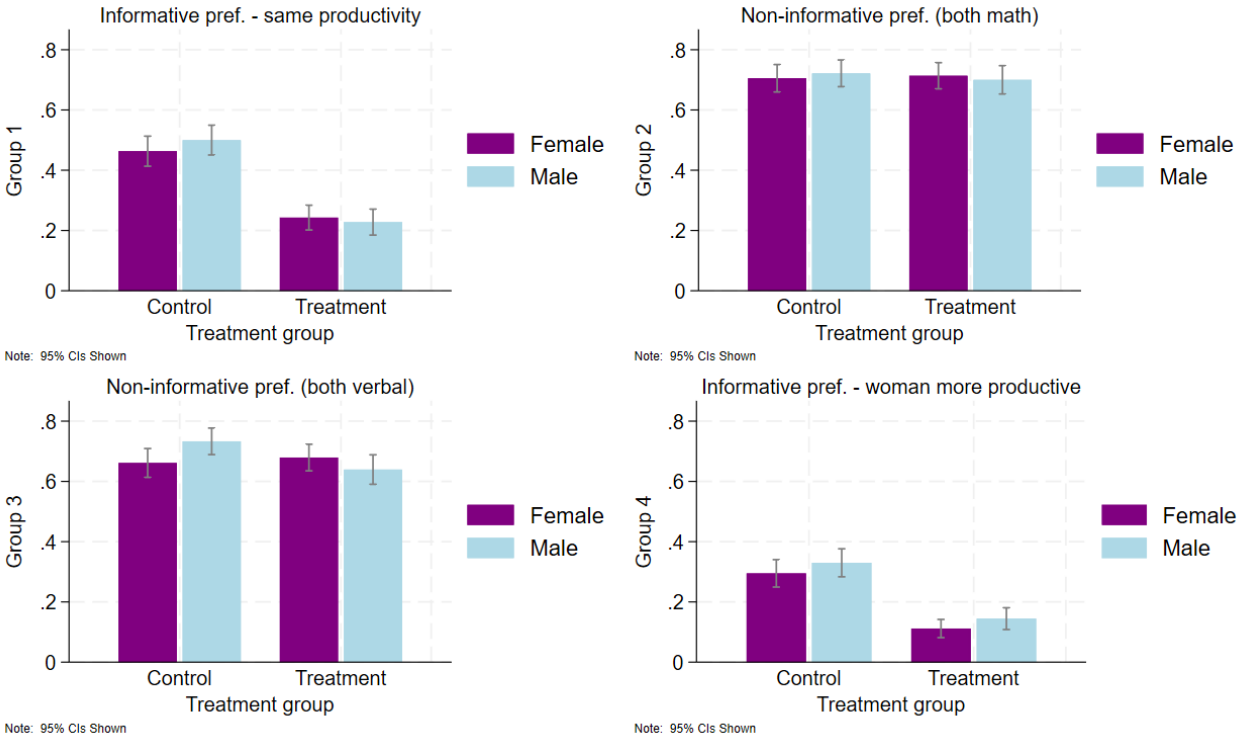


Note: Each panel corresponds to one of the four mixed-gender groups evaluated by 385 employers (192 in the control group, 193 in preference treatment). Purple bars indicate control treatment, while light blue bars represent the preference treatment. Groups 1 and 4 feature informative productivity signals, whereas Groups 2 and 3 contain non-informative signals. The figure reports the proportion of stereotypical task assignments, defined as assigning the male worker to the math task and the female to the verbal task. Error bars show 95% confidence intervals.

Next, we examine whether the likelihood of making a stereotypical task assignment differs by the gender of the employer. To do so, Figure 4 replicates the analysis from Figure 3, but disaggregates the sample by employer gender.

Visual inspection of the figure reveals no statistically significant differences between male and female employers in either the baseline or preference treatment. This suggests that the tendency to assign tasks along gender-stereotypical lines is not systematically driven by the employer's gender, in either the mathematical or verbal domain.

Figure 4: Fraction of stereotypical task assignment by gender of employer



Note. Each panel corresponds to one of the four mixed-gender groups evaluated by 385 employers (192 in the control group, 47.98% female, 193 in the preference treatment, 52.02% female). Within each treatment, light blue bars represent male employers, and purple bars represent female employers. Groups 1 and 4 feature informative productivity signals, whereas Groups 2 and 3 contain non-informative signals. The figure reports the proportion of stereotypical task assignments, defined as assigning the male worker to the math task and the female to the verbal task. Error bars show 95% confidence intervals.

Before turning to the regression analysis, we briefly revisit Hypothesis 2 and the structure of Groups 2 and 3. Hypothesis 2 posited that employers would be less likely to accommodate counter-stereotypical preferences when expressed by women compared to men. In both Groups 2 and 3, the two workers express the same task preference, meaning only one preference can be accommodated in each case. By comparing the frequency with which female counter-stereotypical preferences are accommodated in Group 2 and male counter-stereotypical preferences in Group 3, we can assess whether there is support for Hypothesis 2. In Group 2, the female worker has counter-stereotypical preferences (favoring math), while the male worker has stereotypical preferences (also favoring math). Employers assign the male worker to the math task in about 71% of cases, disregarding the woman's preference. In Group 3, the male worker has counter-stereotypical preferences (favoring verbal), while the female worker has stereotypical preferences (also verbal), and employers again assign the female worker to the verbal task in about 71% of cases, disregarding the man's preference.

These symmetrical patterns suggest that employers tend to fall back on gender-stereotypical task assignments when faced with conflicting preferences—regardless of which gender expresses the

counter-stereotypical preference. The preference treatment does not significantly alter these outcomes in either group. As such, we find no support for Hypothesis 2: counter-stereotypical preferences are equally unlikely to be accommodated for both men and women in settings where only one can be prioritized.

A further comparison between Group 1 and Group 3 allows us to explore how employers respond to women's preferences, depending on whether they are stereotypical or counter-stereotypical. In Group 3, women express stereotypical preferences for the verbal task and are matched to it in over 70% of cases, regardless of the treatment. In contrast, in Group 1, where women express counter-stereotypical preferences (favoring the math task), they are accommodated in less than half of the cases in the baseline condition. However, under the preference treatment, the rate of accommodation increases significantly. This comparison reinforces the idea that women's preferences are more likely to be respected when they conform to gender stereotypes, and that making preferences visible can mitigate—but not fully eliminate—this asymmetry.

To estimate the effect of our treatment on our outcomes of interest, we adopt a Linear Probability Model (LPM) and estimate the following specification:

$$Y_i = \alpha + \beta T_i + \gamma F_i + \delta X_i + \varepsilon_i$$

Where:

- Y_i represents the dependent variable, which varies across models:
  i. A dummy variable at the employer level taking value 1 if the employer assigns the worker to the task in a counter-stereotypical way (i.e., a female worker is assigned to the addition anagram, and a male worker is assigned to the anagram addition task) and 0 otherwise.
  ii. The employer's earnings from the assignment task.
- T_i is a dummy variable for the treatment condition.
- F_i is a dummy variable indicating whether the employer is female (=1), or male (=0).
- X_i is a vector of control variables, progressively included in different models:
  a. In **Model 1 and 4**, we include only the treatment dummy (T_i) and the employer's gender (F_i).
  b. In **Model 2 and 5**, we add a set of socio-demographic controls, including respondents' age, type of secondary school diploma, type of school (e.g., law, economics), and city of birth;
  c. In **Model 3 and 6**, we further control for personality traits, including risk aversion and competitiveness.
- ε_i is the error term.

Each employer made 12 task allocation decisions. We focus only on decisions involving mixed-gender pairs in Groups 1 and 4, where employers had the opportunity to act in a stereotypical or counter-stereotypical way. For this purpose, we construct a binary outcome variable at the employer level, equal to 1 if the employer consistently assigns the female worker to the verbal task and the male worker to the math task, and 0 otherwise. These focal scenarios occur in two distinct contexts:

1. Group 1: Male and female workers are equally productive but express opposite, counter-stereotypical preferences (e.g., the woman prefers math, the man prefers verbal tasks).
2. Group 4: The female worker has higher productivity in the math task, while both workers again express counter-stereotypical preferences.

To increase realism and avoid repetition, each group contains three unique worker pairs matched in productivity and preferences, but differing in background characteristics (e.g., age, degree). Each employer was randomly assigned to evaluate two of these three pairs per group, ensuring variation across rounds. Table 6 reports the results of the LPM estimations. Models 1–3 refer to Group 1, while Models 4–6 refer to Group 4. In all models, we use "female" as the binary gender variable for consistency with summary tables. We also estimate a pooled model across Groups 1 and 4 (see Table A.4. in the Appendix), including a dummy for group membership: results are unchanged.

Inspection of Table 6 reveals a significant effect of the preference treatment on employer decision-making. When workers' task preferences are disclosed, employers are significantly more likely to deviate from stereotypical assignments.

In the baseline specification (Model 1), the preference treatment reduces the likelihood of stereotypical task assignment by 24 percentage points ($p < 0.01$). When controlling for socio-demographic characteristics and personality traits (Model 3), the effect holds, with a reduction of 23.9 percentage points ($p < 0.01$). The effect is smaller in size when we look at Group 4, in which employers were exposed to a stronger preference treatment. In the baseline specification (Model 4), the preference treatment reduces the likelihood of stereotypical task assignment by 18 percentage points ($p < 0.01$), and the effect holds constant when we include control variables (Model 5 and Model 6).

The comparison between the baseline and preference treatments reveals a significant shift in employer decision-making when workers' preferences are disclosed. In the preference treatment, managers are more likely to assign tasks that align with workers' stated preferences rather than relying on gender-based assumptions. This shift reduces the likelihood of stereotypical task assignments, leading to a more equitable distribution of tasks and better alignment between worker capabilities and task requirements.

These results suggest that explicit preference information is a powerful and effective tool for mitigating gender-stereotypical behavior in task allocation. By making individual preferences salient, the preference treatment appears to redirect employers' attention away from heuristic-driven choices

based on gender stereotypes, promoting more personalized and merit-based assignments. This finding has important implications for organizational practices: even without structural changes or formal diversity policies, low-cost informational nudges—such as eliciting and sharing workers' task preferences—can meaningfully reduce bias in managerial decisions and support fairer, more efficient workplace dynamics. Tables A5 and A6 in the Appendix report, respectively, the results for the non-stereotypical assignment (Groups 2 and 3) and the estimation results, without control variables, for the choices across all four groups.

Table 6: Treatment effect on employers' stereotypical choice

Model	(1)	(2)	(3)	(4)	(5)	(6)
Group considered	1st group			4th group -		
Dependent Variable	Stereotypical task assignment			Stereotypical task assignment		
Estimation method	Linear Probability Model			Linear Probability Model		
Preferences	-0.245***	-0.239***	-0.239***	-0.184***	-0.174***	-0.172***
Treatment	(0.0359)	(0.0403)	(0.0404)	(0.0269)	(0.0295)	(0.0296)
Female Employer	-0.0325	-0.0331	-0.00604	-0.151	-0.217	-0.193
	(0.105)	(0.111)	(0.115)	(0.118)	(0.135)	(0.135)
Constant	0.302***	0.211	-0.0220	0.130	0.518	0.361
	(0.103)	(0.715)	(0.772)	(0.101)	(0.638)	(0.719)
Socio-demographics	No	Yes	Yes	No	Yes	Yes
Personality traits	No	No	Yes	No	No	Yes
Employers	385	379	379	385	357	357
Observations	770	758	758	770	714	714

Note. Clustered standard errors at the individual level in parentheses (N=385). *** p<0.01, ** p<0.05, * p<0.1. Socio-demographic controls include: age, secondary school, type of school (e.g., law, economics), city of birth. Personality traits include: risk aversion, competitiveness. In models (2) and (3) we loose 6 and 28 observations due to dropped observations for multicollinearity when adding the variable "city of birth" as control variable.

5.2 Task assignment and employers' earnings

In the final part of our analysis, we examine whether the preference treatment not only reduced stereotypical task assignments but also improved employers' outcomes in terms of earnings. Since employers' payoffs were directly tied to the performance of the workers in their assigned tasks, this serves as a meaningful proxy for task allocation effectiveness.

Table 7 reports the results from a set of regressions similar to the ones reported in Table 6 in which the dependent variable is the employer's earnings from each allocation decision. Models 1 and 2 refer to Group 1, where both workers are equally productive but express counter-stereotypical preferences. Models 3 and 4 refer to Group 4, where the female worker has a higher productivity signal in the math task. In all models, we test whether employers in the preference treatment—who were informed about workers' stated task preferences—earned more than those in the baseline condition.

The results show that employers in preference treatment earned significantly more, suggesting that making workers' preferences visible facilitated better task matching, which in turn led to higher-performing pairings. These findings highlight that reducing bias is not only normatively desirable but also instrumentally valuable for organizations.

To confirm the robustness of these findings, we also estimate a pooled model using data from all four mixed-gender groups (see Appendix Table A.8). The treatment effect remains positive and statistically significant, reinforcing the conclusion that preference-based task assignment improves both fairness and efficiency in managerial decision-making. Table A.7 conducts the same analysis on the other groups: no impacts are found on the earnings.

Table 7: *Treatment effects on employers' earnings*

Model	(1)	(2)	(3)	(4)	(5)	(6)
Group considered	Group 1			Group 4		
Dependent Variable	Earnings from task allocation			Earnings from task allocation		
Estimation method	Linear Regression			Linear Regression		
Preferences Treatment	0.279***	0.324***	0.338***	0.0866***	0.0837***	0.0854***
	(0.0828)	(0.0863)	(0.0875)	(0.0323)	(0.0323)	(0.0323)
Female	0.0710	-0.0267	-0.0443	-0.0127	-0.0175	-0.00192
	(0.0808)	(0.119)	(0.123)	(0.0314)	(0.0340)	(0.0348)
Constant	4.870***	7.236***	7.218***	4.852***	5.113***	5.245***
	(0.317)	(0.711)	(0.753)	(0.120)	(0.318)	(0.336)
Socio-demographics	No	Yes	Yes	No	Yes	Yes
Personality traits	No	No	Yes	No	No	Yes
# Employers	385	385	385	385	385	385
Observations	770	770	770	770	770	770
R-squared	0.709	0.778	0.779	0.250	0.301	0.304

Note. Clustered standard errors at the individual level in parentheses (N=385). *** p<0.01, ** p<0.05, * p<0.1. Socio-demographic controls include: age, secondary school, type of school (e.g., law, economics), city of birth. Personality traits include: risk aversion, competitiveness.

Conclusions

This study investigates whether making workers' task preferences visible to managers can reduce gender-stereotypical task assignments and improve organizational outcomes. Through a two-stage experiment involving real-effort tasks and allocation decisions, we find that even minimal informational interventions—such as disclosing employees' preferences—can meaningfully alter managerial behavior.

Our results strongly support Hypothesis 1: in the absence of preference information, employers systematically assign tasks along gender-stereotypical lines, despite no underlying performance differences between men and women. However, when preferences are made visible, stereotypical assignments decrease significantly, especially in cases where workers express counter-stereotypical

preferences or where the productivity signal is informative. This shift results in not only fairer task allocation but also improved managerial payoffs, suggesting that reducing bias enhances both equity and efficiency.

In contrast, our evidence does not support Hypothesis 2. We find no indication that women's counter-stereotypical preferences are less likely to be accommodated than men's. Instead, employers tend to ignore counter-stereotypical preferences regardless of the worker's gender when only one preference can be fulfilled. This suggests that, in contexts of conflict, employers fall back on gender stereotypes rather than privileging one group's preferences over another. However, additional comparisons indicate that women's preferences are more likely to be honored when they conform to traditional expectations—highlighting that stereotype-consistent choices still receive preferential treatment.

Taken together, our findings suggest that simple, low-cost interventions like preference elicitation can help reduce the role of gender stereotypes in managerial decision-making. While such mechanisms do not fully eliminate bias, they represent an important step toward more meritocratic and inclusive workplace practices.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGTP 4 to improve language and readability, with caution. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

References

Babcock, L., Recalde, M. P., Vesterlund, L., & Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review*, *107*(3), 714-747.

Babcock, L., Recalde, M. P., & Vesterlund, L. (2017). Gender differences in the allocation of low-promotability tasks: The role of backlash. *American Economic Review*, *107*(5), 131-135.

Dasgupta, U., Mani, S., & Souther, A. (2024). "Let It Be": Motivating Workers Through Task Autonomy. Available at SSRN 4813049.

Dohmen, T., & Falk, A. (2011). Performance Pay and Multi-dimensional Sorting – Productivity, Preferences and Gender. *Labour Economics*, 18(4), 586–597

Eckartz, K., Kirchkamp, O., & Schunk, D. (2012). *How do incentives affect creativity?* (No. 4049). CESifo Working Paper.

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

Fischbacher, U., Kübler, D., & Stüber, R. (2024). Betting on diversity—Occupational segregation and gender stereotypes. *Management Science*, *70*(8), 5502-5516.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504-528. https://doi.org/10.1016/S0092-6566(03)00046-1

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

Halladay, B., & Landsman, R. (2022). Perception matters: The role of task gender stereotype on confidence and tournament selection. *Journal of Economic Behavior & Organization*, *199*, 35-43.

Heilman, M. E. (2001). Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder. *Journal of Social Issues*, 57(4), 657–674

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in organizational Behavior*, *32*, 113-135.

Kamei, K., & Markussen, T. (2023). Free riding and workplace democracy—heterogeneous task preferences and sorting. *Management Science*, *69*(7), 3884-3904.

Lezzi, E., Fleming, P., & Zizzo, D. J. (2015). Does it matter which effort task you use? A comparison of four effort tasks when agents compete for a prize. *A Comparison of Four Effort Tasks When Agents Compete for a Prize (April 15, 2015)*.

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.

Sarsons, H., Gërxhani, K., Reuben, E., & Schram, A. (2021). Gender differences in recognition for group work. *Journal of Political economy*, *129*(1), 101-147.

Sloof, R., & Van Praag, C. M. (2010). The effect of noise in a performance measure on work motivation: A real effort laboratory experiment. *Labour Economics*, *17*(5), 751-765.

Appendix A.

A.1 Composition of the pairs evaluated by the employers.

Using data from the worker experiment, we created eight groups, detailed below. Each group consists of pairs of workers to whom employers must assign tasks based on available information, including gender, age, field of study, productivity signals for each task, and, in the revealed preferences treatment, their task preferences,

Groups 1-4 feature mixed-gender pairs, with three pairs in Groups 1-3 and two pairs in Group 4. Groups 5-8 consist of same-gender pairs, with Groups 5-6 including two male pairs each and Groups 7-8 including two female pairs each. Groups 1-3 consist of worker pairs with equal productivity in both tasks but different preferences. In Group 1, workers exhibit counter-stereotypical preferences, with the woman preferring the mathematical task and the man preferring the verbal task. In Group 2, both workers share a preference for the mathematical task, while in Group 3, the situation is reversed, with both workers preferring the verbal task. In Group 4, the woman is more productive than the man in the mathematical task. Groups 5-8 feature same-gender pairs with identical task preferences. These groups are less central to our analysis and were primarily included to prevent employers from inferring the study's objective.

Each employer participated in 12 decision rounds, each involving a randomly selected pair of workers from the groups below. For each pair, the employer was required to assign a task to each worker. Specifically, from Groups 1-3, each employer viewed two out of three pairs, selected randomly. In Group 4, employers always viewed both pairs, while from Groups 5-8, each employer was randomly assigned one of the two pairs. Within each group, the order of presentation was fully randomized to eliminate potential ordering effects.

Participants were informed that, at the end of the twelve decision rounds, one decision would be randomly selected, and their payments would be determined based on the choice they made in that specific round. Given that the selected decision was unknown in advance, they were encouraged to carefully evaluate each choice, ensuring that their decisions accurately reflected their true preferences and strategic considerations throughout the task.

Table A.1. panel 1.– GROUP 1

Group 1: Counter-stereotypical preferences for both workers, same productivity in both tasks										
	Signal of performance in math task		**Signal of performance in verbal task**		**Age**		**Education**		**Preferences (only displayed in the treatment)**	
Pair	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**
a	6	6	7	7	25	25	Literature	Archaeology	Math	Verbal
b	8	8	10	10	21	24	Economics	Low carbon technologies	Math	Verbal
c	5	5	6	6	20	19	Chemistry	Management	Math	Verbal

Figure A1. Reproduces the decision about pair a) from Group 1. The information about the preferred task was only displayed in the Treatment.

Figure A.1. Decision screen presented to the employers featuring pair a) from **Group 1.**

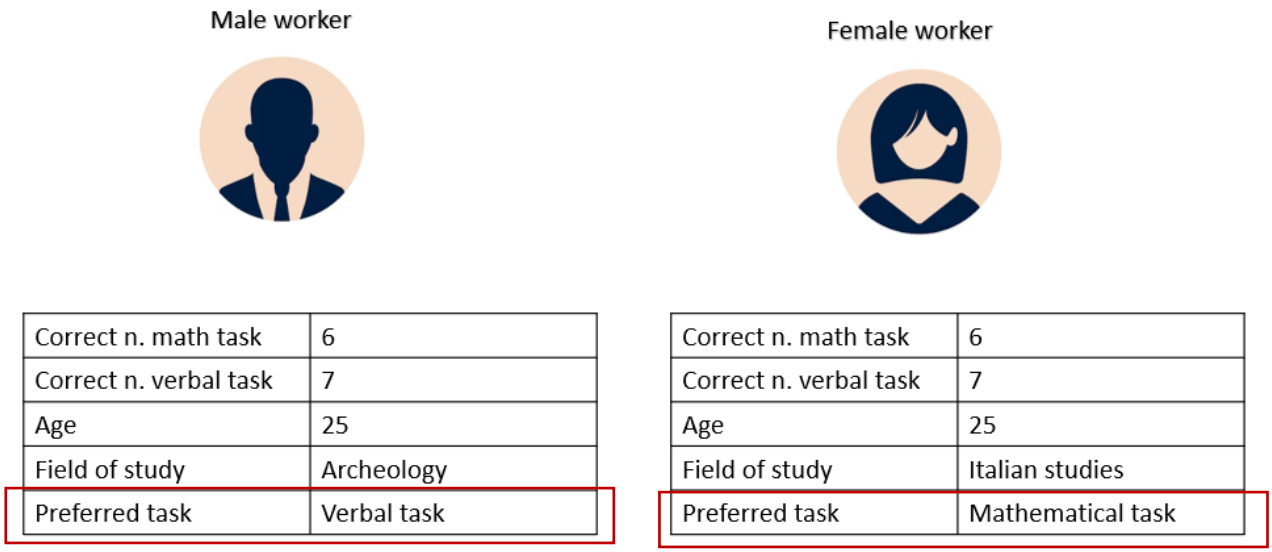


Table A.1 panel 2. – GROUP 2

Group 2: Counter-stereotypical preferences for female worker and stereotypical for male worker, same productivity in both tasks										
	Signal of performance in math task		**Signal of performance in verbal task**		**Age**		**Education**		**Preferences (only displayed in the treatment)**	
Pair	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**
a	6	6	7	7	25	24	Italian Literature	Applied Cognitive Psychology	Math	Math
b	8	8	9	9	20	21	Education	International cooperation	Math	Math
c	6	6	7	7	25	21	Global cultures	Management	Math	Math

Below, Figure A2. Reproduces the decision about pair a) from group 2. The information about the preferred task was only displayed in the Treatment.

Figure A.2. Decision screen presented to the employers featuring pair a) from **Group 2**.

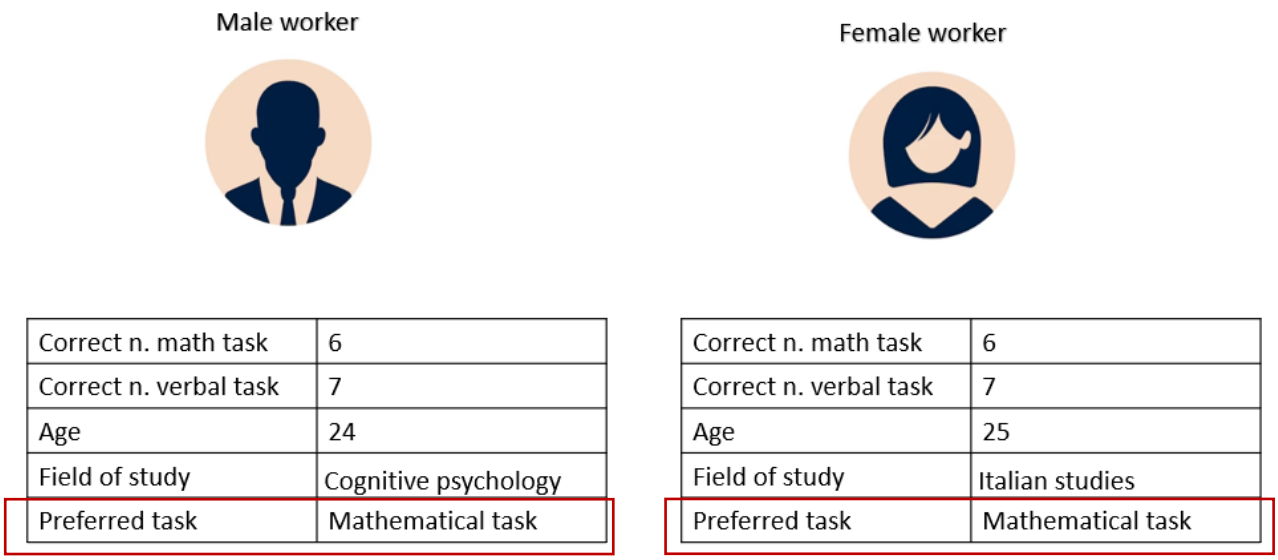


Table A.1. panel 3 GROUP 3

Group 3: Counter-stereotypical preferences for male worker and stereotypical for female worker, same productivity in both tasks										
	Signal of performance in math task		Signal of performance in verbal task		Age		Education		Preferences (only displayed in the treatment)	
Pair	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**
a	6	6	7	7	23	25	Law	Archaeology	Verbal	Verbal
b	7	7	9	9	24	24	Biology	Economics and politics	Verbal	Verbal
c	7	7	9	9	23	24	Global change ecology	Economics	Verbal	Verbal

Below, Figure A3. Reproduces the decision about pair a) from **Group 3**. The information about the preferred task was only displayed in the Treatment.

Figure A.3. Decision screen presented to the employers featuring pair a) from Group 3.

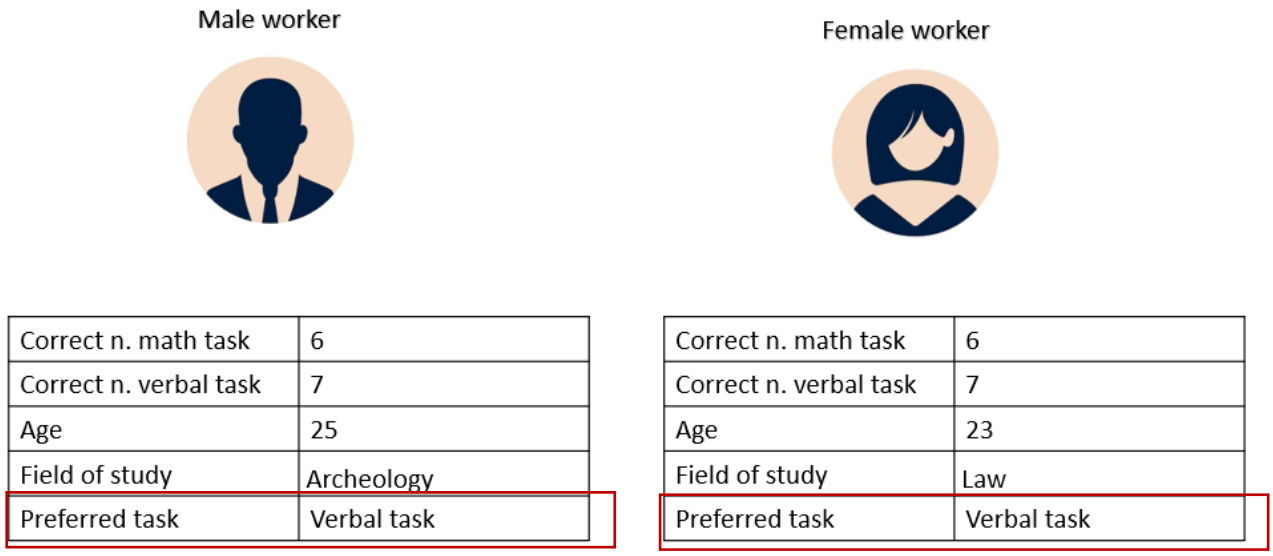


Table A.1. panel 4– GROUP 4

Group 4: Counter-stereotypical preferences for both workers, female worker more productive in the mathematical task, same productivity in the verbal task										
	Signal of performance in math task		**Signal of performance in verbal task**		**Age**		**Education**		**Preferences (only displayed in the treatment)**	
Pair	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**	**Man**
a	7	6	8	8	23	20	Education	Computer Engineering	Math	Verbal
b	8	7	9	9	23	24	Economics and finance	Literature	Math	Verbal

Below, Figure A4. Reproduces the decision about pair a) from group 4. The information about the preferred task was only displayed in the Treatment.

Figure A.4. Decision screen presented to the employers featuring pair a) from Group 4.

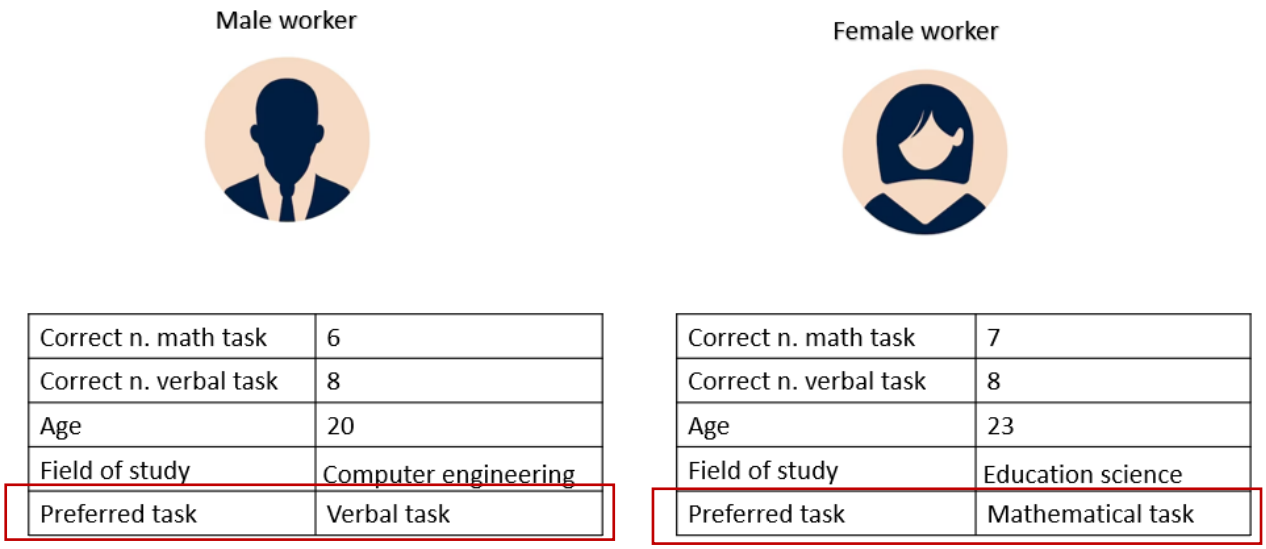


Table A.1. panel 5. – GROUP 5

Group 5: Stereotypical preferences, pairs with male workers only.										
	Signal of performance in math task		**Signal of performance in verbal task**		**Age**		**Education**		**Preferences (only displayed in the treatment)**	
Pair	**Man 1**	**Man 2**	**Man 1**	**Man 2**	**Man 1**	**Man 2**	**Man 1**	**Man 2**	**Man 1**	**Man 2**
a	6	6	7	7	24	21	Psychology	Management	Math	Math
b	7	8	6	7	25	20	Agricultural science	Literature	Math	Math

Table A.1. panel 6. – GROUP 6

Group 6: Counterstereotypical preferences, pairs with male workers only.										
	Signal of performance in math task		Signal of performance in verbal task		Age		Education		Preferences (only displayed in the treatment)	
Pair	**Man 1**	**Man 2**	**Man 1**	**Man 2**	**Man 1**	**Man 2**	**Man 1**	**Man 2**	**Man 1**	**Man 2**
a	7	7	10	10	24	24	Economics	Statistics	Verbal	Verbal
b	8	8	9	9	24	22	Literature	Economics	Verbal	Verbal

Table A.1. panel 7. – GROUP 7

Group 7: Counterstereotypical preferences, pairs with female workers only.										
	Signal of performance in math task		Signal of performance in verbal task		Age		Education		Preferences (only displayed in the treatment)	
Pair	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**
a	6	6	7	7	25	25	Global cultures	Literature	Math	Math
b	6	6	7	7	20	25	Management	Literature	Math	Math

Table A.1. panel 8. – GROUP 8

Group 8: Stereotypical preferences, pairs with female workers only.										
	Signal of performance in math task		Signal of performance in verbal task		Age		Education		Preferences (only displayed in the treatment)	
Pair	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**	**Woman 1**	**Woman 2**
a	6	6	7	7	24	23	Medicine	Law	Verbal	Verbal
b	7	7	9	9	23	25	Engineering	Law	Verbal	Verbal

Additional tables

Table A.2 Mean difference tests between participants in treatment and control groups, employers experiment

	Control	Mean	Treat	Mean	Diff	St Err	p value
Female (=1)	192	.494	193	.533	-.038	.051	.761
Age	192	24.567	193	24.27	.298	.353	.400
Highschool (=1)	192	.88	193	.834	.046	.036	.198
Bachelor (=1)	192	.276	193	.3	-.025	.046	.597
Personality trait: extroversion	192	4.023	193	4.139	-.1164	.145	.798
Personality trait: agreeableness	192	4.921	193	4.958	-.0366	.109	.335
Personality trait: conscientiousness	192	5.348	193	5.360	-.011	.121	.091
Personality trait: emotional stability	192	4.171	193	4.199	-.0276	.139	.197
Personality trait: open-minded	192	4.989	193	4.924	.064	.112	.572
Risk-lover (self-reported, 0-10)	192	5.979	193	5.928	.052	.198	.793
Competitive (self-reported, 0-10)	192	6.172	193	6.41	-.238	.234	.311
Boys better at STEM	192	3.558	193	3.239	.319	.279	.254
Girl better at arts	192	3.828	193	3.591	.238	.284	.404
Boys better leaders	192	2.885	193	2.849	.036	.253	.888
Girls better caregivers	192	3.729	193	3.596	.134	.297	.654

Table A.3: Fraction of stereotypical choice in each group at the pair level within each group

Treatment	**Group 1**			**Group 2**			**Group 3**			**Group 4**	
	1a	**1b**	**1c**	**2a**	**2b**	**2c**	**3a**	**3b**	**3c**	**4a**	**4b**
Preferences	Both counter-stereotypical: female prefers math task, male prefers verbal task			Female workers counter-stereotypical: both prefer math task			Male workers counter-stereotypical: both prefer verbal task			Both counter-stereotypical: female prefers math, male prefers verbal	
Ability	No differences in signal for both tasks			No differences in signal for both tasks			No differences in signal for both tasks			Female workers more productive in math, no differences in verbal tasks	
Pair	1a	1b	1c	2a	2b	2c	3a	3b	3c	4a	4b
Baseline	.359	.139	.139	.375	.429	. 265	.353	.429	.265	.281	.031
N	128	122	134	137	121	126	130	121	126	192	192
Preferences	.119	.169	.169	.384	.423	.257	.316	.423	.257	.103	.023
N	126	127	133	130	124	132	139	124	132	193	193
Mann-Whitney test	0.000	0.307	0.000	0.739	0.778	0.790	0.194	0.778	0.790	0.000	0.493

Note. Table A.3 reports the share of stereotypical task assignments across the four mixed-gender groups, disaggregated by pair and treatment. Data are calculated at the employer level. For Groups 1–3, each employer evaluated a random selection of 2 out of 3 worker pairs; in Group 4, all employers evaluated 2 predefined pairs. The last row presents results from a set of Mann-Whitney tests comparing the treatments.

Table A.4: Treatment effect on employers' stereotypical choice, pooled regression

Model	(1)	(2)	(3)
Group considered	1st and 4th group pooled together		
Dependent Variable	Stereotypical task assignment		
Estimation method	Linear Probability Model		
Preferences	-0.206***	-0.209***	-0.207***
Treatment	(0.0238)	(0.0256)	(0.0259)
Female Employer	-0.0127	-0.0128	-0.00933
	(0.0238)	(0.0256)	(0.0265)
	0.0114	0.00912	0.00289
Constant	0.372***	0.313**	0.285*
	(0.0442)	(0.142)	(0.148)
Socio-demographics	No	Yes	Yes
Personality traits	No	No	Yes
Employers	385	385	385
Observations	1,037	1,037	1,037

Note. Clustered standard errors at the individual level in parentheses (N=385). *** p<0.01, ** p<0.05, * p<0.1. Socio-demographic controls include: age, secondary school, type of school (e.g., law, economics), city of birth. Personality traits include: risk aversion, competitiveness. We control for group membership including in the analysis a dummy variable that equals 1 for group 1, and 0 for group 4.

Table A.5: effect of treatment on employers' non-stereotypical choice

Model	(1)	(2)	(3)	(4)	(5)	(6)
Group considered	2nd group Both prefer math task			3rd group Both prefer verbal task		
Dependent Variable	Stereotipical task assignment (Assign the math task to the male)					
Estimation Method	Linear Probability Model					
Preferences Treatment	-0.0110	0.0352	0.0310	-0.105	-0.138	-0.145
	(0.106)	(0.116)	(0.117)	(0.104)	(0.111)	(0.111)
Female Employer	-0.0106	-0.0213	-0.0584	-0.0499	-0.0262	-0.0538
	(0.106)	(0.116)	(0.120)	(0.104)	(0.116)	(0.117)
Constant	0.718***	0.915	1.288	0.519***	0.577	0.732
	(0.112)	(0.745)	(0.849)	(0.110)	(0.695)	(0.758)
Socio-demographics	No	Yes	Yes	No	Yes	Yes
Personality traits	No	No	Yes	No	Yes	Yes
Employers (N)	385	367	367	385	366	366
Observations	770	734	734	770	739	739

Note. Clustered standard errors at the individual level in parentheses (N=385). *** p<0.01, ** p<0.05, * p<0.1. Socio-demographic controls include: age, secondary school, type of school (e.g., law, economics), city of birth. Personality traits include: risk aversion, competitiveness. In models (2) and (3) we loose 6 and 28 observations due to dropped observations for multicollinearity when adding the variable "city of birth" as control variable.

Table A.6: effect of treatment on employers' choices

	(1)	(2)	(3)	(4)
	1st Group- Stereotypical task assignment	2nd Group- prefer math task	3rd Group – both prefer verbal task	4th Group – Stereotypical task assignment
Preferences Treatment	-0.684***	-0.0110	-0.105	-0. 740***
	(0.106)	(0.106)	(0.104)	(0. 127)
Female	-0.0325	-0.0106	-0.0499	-0. 151
	(0.105)	(0.106)	(0.104)	(0. 117)
Constant	0.302***	0.718***	0.519***	0. 130
	(0.103)	(0.112)	(0.110)	(0. 101)
Socio-demographics	No	No	No	No
Personality traits	No	No	No	No
Employers	385	385	385	385
Observations	770	770	770	770

Note. Clustered standard errors at the individual level in parentheses (N=385). *** p<0.01, ** p<0.05, * p<0.1. The table presents results for all four groups of male–female worker pairs. Columns 1 and 4 show results for pairs in which the woman prefers the mathematical task and the man prefers the verbal task. Column 2 shows results for pairs in which both prefer the mathematical task, while Column 3 reports results for pairs in which both prefer the verbal task. No control variables are included in the analysis.

Table A.7: effect of treatment on employers' earnings from non-stereotypical choice

Model	(1)	(2)	(3)	(4)	(5)	(6)
Group considered	2nd group Both prefer math task			3rd group Both prefer verbal task		
Dependent variable	Employers' earnings from the non-stereotypical task assignment					
Estimation method	Linear regression					
Preferences Treatment	0.00406	0.0309	0.0337	0.0384	0.0743	0.0745
	(0.0243)	(0.0276)	(0.0278)	(0.0437)	(0.0484)	(0.0489)
Female	-0.00530	-0.0128	-0.0143	0.0435	0.0241	0.0405
	(0.0234)	(0.0339)	(0.0353)	(0.0434)	(0.0480)	(0.0490)
Constant	4.961***	4.803***	4.814***	4.854***	4.630***	4.735***
	(0.100)	(0.273)	(0.283)	(0.183)	(0.279)	(0.289)
Socio-demographics	No	Yes	Yes	No	Yes	Yes
Personality traits	No	No	Yes	No	No	Yes
Employers	385	385	385	385	385	385
Observations	770	770	770	770	770	770
R-squared	0.735	0.806	0.806	0.915	0.922	0.922

Note. Clustered standard errors at the individual level in parentheses (N=385). *** p<0.01, ** p<0.05, * p<0.1. Socio-demographic controls include: age, secondary school, type of school (e.g., law, economics), city of birth. Personality traits include: risk aversion, competitiveness.

Table A.8: effect of treatment on employers' earnings – all groups pooled together

Model	(1)	(2)	(3)
Dependent variable	Earnings from task allocations - pooled		
Estimamation method	Linear Regression		
Preferences Treatment	0.424***	0.337*	0.345*
	(0.163)	(0.178)	(0.177)
Female	0.0409	-0.158	-0.187
	(0.158)	(0.180)	(0.185)
Constant	14.83***	14.66***	14.56***
	(0.343)	(1.226)	(1.318)
Socio-demographics	No	Yes	Yes
Personality traits	No	No	Yes
Employers	385	385	385
Observations	1,155	1,155	1,155
R-squared	0.027	0.038	0.039

Note. Clustered standard errors at the individual level in parentheses (N=385). *** p<0.01, ** p<0.05, * p<0.1. The table reports results from earnings pooled across the four task allocations. Socio-demographic controls include: age, secondary school, type of school (e.g., law, economics), city of birth. Personality traits include: risk aversion, competitiveness.

APPENDIX B

In what follows we reproduce the Instructions used for the workers experiment and the employers experiment. The experiment was conducted in Italian, the Italian version of the instructions are available from the authors upon request.

B.1 Workers' Experiment

Instructions

You are participating in a study on decision-making processes funded by the University of Bologna. During this study, you can earn a sum of money based on the rules described on the following pages. Payments will be made in cash and in a way that preserves confidentiality.
The entire study will last approximately 40 minutes.

The study consists of three parts and a short final questionnaire. At the end of the study, one of the three parts will be randomly selected by the computer and used to determine your earnings. Therefore, your final earnings from the study will be composed of the sum of your earnings for the selected part plus a participation fee of €5.

The methodology for determining your earnings will differ across the three parts. You will receive instructions for each part, one at a time. Before each part begins, we will describe in detail how your earnings will be determined for that specific part.

Please do not speak with others during the study. The use of mobile phones is strictly prohibited, and using a phone during the study will result in immediate disqualification from payments.

If you have any questions during the study, please raise your hand. An assistant will come to your station to respond privately.

INSTRUCTIONS FOR PART 1

In Part 1, you will perform two different tasks, and for each task, your earnings will depend on your performance.
We will now explain the details of each task and how your earnings will be determined.

Task A: Additions

Your goal in this task is to correctly solve as many addition problems as possible. You will have 3 minutes to solve additions of three 3-digit numbers. The numbers to be added are selected randomly. You will see a screen similar to the one shown below.

You can use the sheets of paper and pen provided at your station. When you are ready, you can enter your answer in the designated box and click the red button. After clicking the red button, the computer will immediately tell you whether your answer is correct or not. Your responses are anonymous.

Earnings for Task A
You will earn €0.50 for each correct addition solved within the 3-minute time limit. Your earnings will not decrease if you provide an incorrect answer.

Task B: Anagrams

Your goal in this task is to correctly solve as many anagrams as possible. You will have 3 minutes to find a 5-letter word using a set of 5 randomly ordered letters. You will see a screen similar to the one shown below.

For example, in this case, the correct answer is: VERBO.
Each anagram has a single solution and is in Italian.

You can use the sheets of paper and pen provided at your station. When you are ready, you can enter your answer in the designated box and click the red button. After clicking the red button, the computer will immediately tell you whether your answer is correct or not. Your responses are anonymous.

Earnings for Task B
You will earn €0.50 for each correct anagram solved within the 3-minute time limit. Your earnings will not decrease if you provide an incorrect answer.

What happens now?

Each participant will perform both Task A and Task B, but the order in which they are presented will be random.

Earnings for Part 1
If Part 1 of the study is selected for the final payment, there will be a new random selection between the two tasks you completed. The earnings for Part 1 will correspond to the earnings obtained in the randomly selected task.

If you have any questions about Part 1, please raise your hand, and an assistant will come to your station to respond privately.

INSTRUCTIONS FOR PART 2 AND PART 3

In Part 2, you will perform the same tasks as in Part 1, each for 3 minutes, and you will receive the same payments for correct answers. As in Part 1, the two tasks will be presented in random order.

Earnings for Part 2
If Part 2 of the study is selected for the final payment, there will be a new random selection between the two tasks you completed. The earnings for Part 2 will correspond to the earnings obtained in the randomly selected task.

IMPORTANT
Before starting Part 2, we will ask you to indicate which of the two tasks you performed in Part 1 is your preferred task. This choice will determine the content of Part 3 of the study.

In Part 3, you will perform only one task—the task you indicated as your preferred task in Part 2—for an additional 3 minutes.

Earnings for Part 3
If Part 3 of the study is selected for the final payment, your earnings will depend on your performance in the task you performed in Part 3.

If you have any questions about Part 2, please raise your hand, and an assistant will come to your station to respond privately.

B.2 Employers' Experiment

The text displayed only in the treatment group is displayed in parenthesis.

Welcome!
You are about to participate in a study that consists of a questionnaire. The study will take approximately 20 minutes to complete. You will receive €3 for your participation.

Additionally, by answering certain questions, you may earn an additional amount ranging from €0 to €5. Further details will be provided before these questions.

Once the data collection is complete, you will receive information about the correctness of your answers and your final earnings. Payments will be made via PayPal within seven days of completing the study.

At the end of the questionnaire, we will ask you to provide some personal information (email and name). This information will be used to notify you of your earnings and to process payments.
Your personal data will be used solely for payment purposes, while all the answers you provide will remain anonymous. No one, not even the researchers handling the data collected in this questionnaire, will be able to identify the author of the responses.

By clicking the arrow below, you will begin the study. You can only participate once, and the questionnaire cannot be paused or resumed. At 1:00 PM on Thursday, January 25, 2024, the data collection will end, and no further responses will be accepted. Only those who complete the questionnaire in its entirety will receive payment.

You will be presented with a decision-making scenario, and by putting yourself in that situation, you will need to answer 12 questions.

The decision-making scenario is as follows: You are the manager of a company and must allocate two different tasks to two employees under your supervision.

The two alternative tasks you need to assign to the workers are:
A) Solving as many additions as possible of three 3-digit numbers, or
B) Solving as many anagrams as possible of words composed of 5 letters (2 vowels and 3 consonants).

Below, you can see examples of the two tasks performed by the workers. Try to solve them:

1. The additions are similar to these:
   536 + 414 + 780
   235 + 689 + 327

2. The anagrams are similar to this:
   VBOER
   ZAZPI

You will need to make 12 decisions. Specifically, you will be presented with 12 pairs of employees in sequence, and each time you must decide who to assign the first task to and who to assign the second task to. Your earnings will depend on the performance of the two workers in the tasks assigned to them.

Before selecting the workers, you will have the opportunity to review their CVs using the information each worker provided when participating in the study.

The workers you can assign to the two alternative tasks all belong to a group of 60 students who participated in an in-person study at the BLESS laboratory in Bologna in December 2023. The study consisted of two equal parts. In the first part, the students/workers had 3 minutes to solve as many addition problems as possible and 3 minutes to solve as many anagrams as possible. The two tasks were presented in random order, and at the end, each participant was asked which task they preferred. In the second part, they again had 3 minutes + 3 minutes to complete each of the two tasks, presented again in random order. Each student/worker thus completed each task twice and was then paid based on the number of correct answers provided in one of the four tasks performed, randomly selected by the computer.

To make your decision, you will be presented with a similar scenario:

Male worker

Correct n. math task	X
Correct n. verbal task	Y
Age	25
Field of study	XXX

Female worker

Correct n. math task	X
Correct n. verbal task	Y
Age	25
Field of study	XXX

The following scenario was presented in the Treatment (it presents a final row with information about the worker's preferred task.

Male worker

Correct n. math task	X
Correct n. verbal task	Y
Age	25
Field of study	XXX
Preferred task	XXXXX

Female worker

Correct n. math task	X
Correct n. verbal task	Y
Age	25
Field of study	XXX
Preferred task	XXXXX

For each worker, in addition to information about their age and the degree program they are enrolled in, you will also have information about their productivity in the two tasks from part 1. You will need to decide which worker to assign to each task.

[The text in parenthesis was displayed only in the Treatment group: Additionally, you will be provided with information about each worker's preferred task. This preference was collected at the end of part 1, before the start of part 2.**]**

Your earnings depend on the performance of the two workers in the tasks assigned to them. To calculate your earnings, we will use the workers' performance in part 2 of the study (on which you have no information). Remember that in part 2, the workers performed both tasks for 3 minutes in random order.

Specifically, you will receive 30 cents for each addition correctly solved by the worker you assigned the addition task to, and 20 cents for each anagram correctly solved by the worker you assigned the anagram task to.

On average, in the second part of the study, workers correctly completed 10 addition tasks and 14 anagram tasks.

At the end of your twelve decisions, one will be randomly selected, and you will receive payments based on the choice you made in that selected decision. Since you do not know which decision will be chosen, make sure to carefully consider each one as you go through the task.

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