Journal of Behavioral and Experimental Economics 109 (2024) 102142

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

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

Women lean back when representing others in competitions

Maliheh Paryavi *

Harvard University, 79 JFK Street, Cambridge, MA 02138, USA

ARTICLE INFO

Keywords:
Gender gap
Representation
Competition
Confidence
Leadership

ABSTRACT

This study examined the role that being a 'representative' plays in competition behavior of women and men through a laboratory experiment where self-representing and other-representing individuals had to decide whether to enter a mix-gender tournament that involved performing in a male-stereotyped task. While self-representing men and women exhibited very similar performance and competitiveness behaviors, women 'representatives' *leaned back*. Controlling for differences in performance, female 'representatives' were less likely to enter into competitions than male 'representatives' and self-representing women. The leaning back of female 'representatives' from competition entry as compared to self-representing women is largely attributed to these women experiencing lower levels of confidence and being less likely to enter a competition when they guessed their performance to be less than the best. Furthermore, compared to male 'representatives', female 'representatives' were significantly less confident in their abilities, explaining the gender gap in competition entry amongst other-representing individuals.

1. Introduction

Women's willingness to enter leadership roles and advance their candidacy has been a central focus of the literature on women's leadership thus far. While women are often encouraged to "lean in" (Sandberg, 2013), compared to men, women have been found to be less willing to self-select into leadership roles (Alan, Ertac, Kubilay & Loranth, 2020; Born, Ranehill & Sandberg, 2022; Chen & Houser, 2019) and opt into leadership positions even when they know that they are the top performers (Erkal, Gangadharan & Xiao, 2022). Women are also less likely than men to compete (Markowsky & Beblo, 2022; Niederle, 2017; Niederle & Vesterlund, 2007), put themselves forward for opportunities (Coffman, Collis & Kulkarni, 2020), promote themselves (Coffman, Flikkema & Shurchkov, 2021; Exley & Kessler, 2019), and contribute their ideas (Bordalo, Coffman, Gennaioli & Shleifer, 2019; Chen & Houser, 2017; Coffman, 2014). In most of these circumstances, stereotyped beliefs are shown to significantly contribute to the observed gender gaps.

The work environment and responsibilities involved in high profile jobs, which include higher levels of competitive pressure, may also be contributing to the gender gaps in leadership[1] (Marianne, 2011; Markowsky & Beblo, 2022; Sandberg, 2013). Specifically, the responsibility of representing the economic interests of others is a key feature of many top-level leadership positions, whether it is representing employees and investors of a company and competing for market share or representing voters and stakeholders in political office and competing for resources. Since these positions continue to be dominated by men, for individuals to maximize gains for themselves and their constituents and advance in these positions, they would be required to perform and compete in stereotypically male domains. How these conditions around the leadership responsibility of representing and competing on behalf of others impact men and women differently remains unclear and serve as the motivation for this research.

The competition literature suggests that women are at a major disadvantage in competitive environments that involve performing in stereotypically male tasks as they are significantly more likely to opt out of competitions (Niederle & Vesterlund, 2007; Saccardo, Pietrasz & Gneezy, 2018). At the same time, the negotiation literature has found that while women opt out of negotiations when only representing their own interests, they are particularly "energized" in negotiations when representing the interests of others (Bowles, Babcock & McGinn, 2005). Combining these results could suggest that while women are more likely to opt out of competition environments that involve stereotypically male tasks when only representing their own interests, they could become

* Correspondence to. Email: maliheh_paryavi@post.harvard.edu
E-mail address: maliheh_paryavi@post.harvard.edu.

[1] As of September 1st, 2021, only 25.8 percent of all national parliamentarians were women (IPU Parline, 2021). In the US, women hold only 24% of senate seats, make up 8.2% of Fortune 500 CEOs, and serve on 26% of corporate boards (Catalyst, 2020, 2021; Rutgers University, 2021).

https://doi.org/10.1016/j.socec.2023.102142
Received 29 July 2022; Received in revised form 20 November 2023; Accepted 21 November 2023
Available online 1 December 2023

"energized" in competitions when representing the interests of others. In fact, there is a prevalent hypothesis in the competition literature that "just as women have been found to opt out of negotiations, and yet be willing to negotiate on behalf of others, it may be that they are more eager to compete when doing so benefits others" (Niederle & Vesterlund, 2011). However, the answer is not so clear as there are key differences between competition and negotiation experiments.

First, while representing another individual may reduce the effects of a constraint responsible for the gender gaps in negotiation experiments, that constraint may not be present or binding in competition experiments. In the negotiation literature, perceptions and the possibility of backlash has been identified as a significant deterrent for women in negotiations (Amanatullah & Morris, 2010; Bowles, 2012), which is expected to be greatest when women are assertive for their own benefit, contradicting communal prescriptions for female behavior (Rudman, 1998; Rudman & Glick, 1999). Hence, in negotiation settings, which are conducted face-to-face in a social environment, negotiating on behalf of others and someone else's benefit may give women a boost and energize them to negotiate assertively without fearing backlash from their male counterparts for taking an "agentic" role (i.e. a stereotypically male behavior) (Bowles et al., 2005). However, participants in competition experiments experience no possibility of backlash given that they remain anonymous, have no direct contact with their multiple competitors, and receive no feedback on others' performance or choices. They also do not experience a possibility of social approval given that no one will know how well they performed. This suggests that women would not be expected to experience a boost in their performance or competition entry from reduced threat of backlash by representing another person in a competition experiment environment.

Secondly, as compared to negotiation settings where a range of positive payoffs are possible and there is room for maneuvering and negotiating, a competition environment like the one in Niederle and Vesterlund (2007) presents a binary decision for the participant which includes the possibility of a zero-payoff. This difference in decision structures could interact with participants' other-regarding preferences. In fact, a fundamental feature of the Niederle and Vesterlund (2007) experiment was that individual choices imposed no externality on the payoffs of others, therefore reducing the effect of other-regarding preferences (Niederle & Vesterlund, 2011). However, while individual incentives would remain constant, if individuals are in charge of representing the interests of another person, the size of the pie would be doubled, and their individual choices would impose an externality on the payoff of another person. Depending on their other-regarding preferences, individuals may respond to either the ability to help another person, the larger pie, or both, which could shape their competition behavior. Numerous studies reveal that men are more likely to be social-surplus maximizers and women are more inequality averse and altruistic (Croson & Gneezy, 2009; Kamas & Preston, 2010; Niederle & Vesterlund, 2011). Furthermore, Kamas and Preston (2010) find that even in a no-externality environment similar to the Niederle and Vesterlund (2007) experiment, individuals who were social-surplus maximizers were more likely to enter into the tournament as compared to inequity averse individuals. Therefore, if individual choices imposed externalities for the payoff of another individual, given women's tendencies to be altruistic and more concerned with equalizing earnings, they may prefer that their constituent earn at least some payoff rather than none. As a result, as compared to negotiation environments where a range of positive payoffs can be negotiated, women would be less likely to enter a win or lose tournament that could have a zero-payoff outcome when representing another individual than if they were only representing themselves.

Furthermore, while the role of gender differences in confidence is not evaluated in the negotiation setting of Bowles et al. (2005), throughout the competition literature, gender differences in beliefs about relative ability (i.e. confidence) and attitudes towards performing in a competition are identified as the main reasons for the gender gap in competitiveness (Niederle & Vesterlund, 2011). These factors could also interact with individuals' other regarding preferences when serving as representatives in a competition environment. In addition, risk and feedback preferences may also become a determining factor when individuals are responsible for another individual even though they have been found to play a negligible role in traditional competition experiments. Therefore, it is not clear whether the findings from the Bowles et al. (2005) negotiation study could be directly transferable and added to the findings from the competition experiments of Niederle and Vesterlund (2011).

As a result, this paper connects the dots between these two research studies, while informing the literature on women's leadership. It examines the impact of the 'representative' role, as in the Bowles et al. (2005) negotiation research, on competition behavior of men and women in a stereotypically male domain, as in the Niederle and Vesterlund (2007) competition research. Would women also be energized to "lean in" and become competitive when leading and representing others in a stereotypically male domain? Or does the leadership responsibility of representing another individual discourage them from competing? By surveying mutual-fund managers (yet, another cohort in which women are underrepresented), Beckmann and Menkhoff (2008) found that female financial experts were more averse to competition and less likely to choose a strategy that could allow them to outperform the market. While mutual-fund managers are clearly responsible for the financial interests of others in a male-dominated field, the role that being a 'representative' plays in their competition behavior by gender is unclear.

To understand if equally qualified and ambitious women become more aggressive competitors when representing others in a male domain, I conducted laboratory experiments in a mix-gender competitive environment, as in Niederle and Vesterlund (2007). The experiment design allowed for careful examination of individual behaviors as well as the contributing factors—such as confidence, risk and feedback aversion, and attitudes towards competition. In the treatment condition, female and male 'representatives' were asked to decide whether they choose to enter into a tournament that involved performing in a male-stereotyped task. Similar to Bowles et al. (2005), these 'representatives' were responsible for the interest of someone else, their 'constituents', in addition to their own. Whereas, in the control group, individuals faced the exact same tasks and decisions, but were only responsible for their own interests.

A significant majority of study participants were Harvard University students, with no significant differences in their levels of confidence, therefore, unlike previous competition experiments, men and women exhibited very similar performance and competition behaviors when only representing themselves. This baseline comparison provides a unique opportunity to examine how the competitive behavior of similarly ambitious and qualified men and women—such as men and women in leadership positions—is influenced by the responsibility of representing another person's interests in a male domain.

The results suggest that not only were women not eager to enter competitions when charged with representing others, but that they leaned back. Controlling for differences in performance, female 'representatives' were less likely to enter into tournaments than male 'representatives' and self-representing women. In addition, female 'representatives' did not significantly increase their performance, while male 'representatives' were particularly energized and experienced a boost in their performance levels as compared to the self-representing men. The leaning back of female 'representatives' from competition entry as compared to self-representing women is largely attributed to these women experiencing lower levels of confidence and being less likely to enter a competition when they guessed their performance to be less than the best. Furthermore, compared to male 'representatives', female 'representatives' were significantly less confident in their abilities, which explained the gender gap in competition entry amongst other-representing individuals.

While subjects did not self-select into the representation role in this

study, these insights illuminate how the responsibility that is a common feature of most high-profile leadership positions (i.e. representing others' interests in stereotypically male-dominated domains) impacts men and women differently.

The following sections present the conceptual framework (Section 2), experimental design (Section 3), results (Section 4), explanations of the results (Section 5), and concluding remarks (Section 6).

2. Conceptual framework

Studies have shown that women are significantly more likely to opt out of competitions and competitive negotiations when only representing their own interests (Bowles et al., 2005; Niederle & Vesterlund, 2007, 2011; Saccardo et al., 2018). However, in one negotiation study, women were found to become particularly energized when charged with representing the interest of another (Bowles et al., 2005). As discussed in Section 1, there are key differences between negotiation and competition experiment settings, which do not permit us to directly apply the findings of the negotiation experiment to the question at hand: how does representing another individual in a competitive environment involving a stereotypically male task (as in Niederle and Vesterlund (2007)) impact individuals' competition behavior?

This paper aims to answer this question and connect the dots between the Bowles et al. (2005) negotiation research and the Niederle and Vesterlund (2007) competition research. Throughout the competition literature, gender differences in beliefs about relative ability (i.e. confidence) and attitudes towards performing in a competition are identified as the main reasons for the gender gap in competitiveness (Niederle & Vesterlund, 2011), which could also interact with individuals' other regarding preferences when serving as representatives. In addition, risk and feedback aversion may also play a role. These factors will be discussed in the next section to examine how representing another individual in a competition environment is expected to shape competition entry decisions.

2.1. Differences in competition entry under representation

2.1.1. Role of differences in confidence

Women have been shown to be less confident than men regarding their relative abilities (Barber & Odean, 2001; Beyer, 1990; Beyer & Bowden, 1997; Exley & Kessler, 2019; Pulford & Colman, 1997; Soll & Klayman, 2004), especially in stereotypically male domains (Coffman et al., 2020; Grosse & Riener, 2010; Kamas & Preston, 2010; Lundeberg, Fox & Punćcohaŕ, 1994). In fact, Niederle and Vesterlund (2007) and others (Balafoutas & Sutter, 2010; Dargnies, 2011; Healy & Pate, 2011; Niederle & Vesterlund, 2010; Shurchkov, 2012; Sutter & Rützler, 2010) directly measure the impact of beliefs about relative ability in winner-take-all tournament entry. They find that women are less confident than men and these beliefs help explain some of the gender gap in tournament entry that they observed. Cason, Masters and Sheremeta (2010), Wozniak, Harbaugh and Mayr (2009), Ertac and Szentes (2010) also assess the role of beliefs about relative ability by manipulating beliefs directly and providing feedback about the performance of others before the decision to enter a tournament was made. They all find that the gender gap in tournament entry is reduced or eliminated when information about the performance of others is provided.

Moreover, while stereotypes seem to strongly predict individual's beliefs about their own ability (Bordalo et al., 2019; Coffman, 2014), men's confidence seems to be derailed less by tasks or circumstances that may shift women's confidence. Wozniak et al. (2009) found no gender difference in confidence in a stereotypically female word task (with men being just as confident as women in this female-stereotyped task), and that men were significantly more confident than women in the math task – the difference being due to women being less confident in math (a stereotypically male-typed task). In fact, Lundeberg et al. (1994) argue that gender differences in beliefs about relative ability exist mostly in masculine domains. Fox and Lawless (2011) also hold that the lower confidence of eligible female electoral candidates can be linked to women's perceptions of the political environment as being sexist and masculinized. In the laboratory as well, the gender gap in confidence and entry is larger in tasks that are stereotypically male (e.g. math tasks vs. word tasks; see Kamas and Preston (2010), Grosse and Riener (2010)) when subjects believe that women are better at the verbal tasks and men better in the math task. Additionally, Shurchkov (2012) varies the time limit allowed for completing tasks to create high-pressure and low-pressure environments. She found that women were more likely to enter a tournament and outperform men in a low-pressure verbal task, possibly suggesting that they were more confident in their abilities in low-pressure stereotypically-female environments.

Given gender differences in other regarding preferences and the higher stakes involved when representing another individuals' interests, women may become more likely to shy away from competition unless they are confident that they will win and increase their payoff and those of their constituent. Women have been shown to raise their expectations of themselves in traditional male domains and when the stakes are high. For example, when assessing their readiness to pursue elective office, women have been found to set an extremely high bar for themselves (Fox & Lawless, 2011; Lawless, 2012). Thebaud (2010) also found that when compared to similarly situated men, women hold themselves "to a stricter standard of competence" and are significantly less likely to view themselves as able to be an entrepreneur. Moreover, in studying gender differences in willingness to guess on tests, Baldiga (2014) finds that women are significantly less likely to guess and answered fewer questions than men when facing a penalty for wrong answers, whereas they answered all questions when there was no imposed penalty.

Additionally, stereotype threat may be implicitly activated for women in high-pressure stereotypically masculine domains (e.g. representing someone else while performing a math task). While stereotype literature (Spencer & Steele, 1999; Steele & Aronson, 1995) focuses on performance evaluation in non-competitive environments, stereotype threat may be contributing to gender differences in beliefs about relative ability to perform well in a competition involving a stereotypically male math task. If stereotype threat is implicitly activated, women may expect to not perform as well as men, which could also shift their beliefs about their relative performance.

While the exact mechanism behind shifts in women's confidence is unclear, given the additional pressure and responsibility of representing the interests of someone else, women are expected to experience lower confidence levels when serving as 'representatives' in a stereotypically male domain while men's confidence is expected to remain constant across treatments. Furthermore, all else equal, women's lower level of confidence as 'representatives' is projected to deter them from entering competition, leading to a significant increase in the gender and treatment gap in competition entry.

2.1.2. Role of differences in attitudes toward performing in a competition

The competition literature contends that men and women hold significantly different preferences for entering competitions that require them to subsequently perform in a task. Women are considered to be more likely to select out of situations that require them to perform in a competition (Niederle & Vesterlund, 2007), while men may even be drawn to them. It has been theorized, "while the prospect of engaging in a future competition may cause women to anticipate a psychic cost and deter them from tournaments, men may anticipate a psychic benefit and instead be drawn to them" (Niederle & Vesterlund, 2007, pg. 1070). This could be a result of girls and boys experiencing different upbringings, where typically girls are encouraged to show empathy, be fair and considerate of others, and boys are encouraged to be assertive (Ruble, Martin & Berenbaum, 2006). Evolutionary psychology also lends support to this theory by using reproductive strategies of each gender to argue that men have "evolved" to enjoy competition (Campbell, 2002;

Daly & Wilson, 1983). Buser, Dreber and Mollerstrom (2017) also show that willingness to compete is more sensitive to cortisol changes for women than for men.

Following from their analysis, Niederle and Vesterlund (2007) argue that gender differences in attitudes toward performing in competition explain a significant portion of the gender differences in tournament entry. Despite the established differences in attitudes towards performing in a competition, the theory is unclear whether women would be encouraged to be assertive or to lean back from entering a competition when representing the interest of another individual. Given the added requirement of representing another individual and the larger stakes, female 'representatives' may become even less drawn to the pressure of performing in a competitive setting and as a result, shy away from entering competitions that require them to perform. In contrast, the added pressure may excite men even more and increase their likelihood of competition entry. At the same time, considering the fact that women are found to be more altruistic than men (Croson & Gneezy, 2009), they could be encouraged to take on the additional pressure for the sake of increasing winnings for their constituent, and choose to perform in a competitive setting. This is also supported by the literature on gender and interdependent self-construal. While men are thought to consider representations of others as separate from themselves, women are thought to construct interdependent self-construal, where they consider others as part of themselves (Cross & Madson, 1997). This may serve to shift women's attitudes towards competition, energizing them to compete if they felt a personal responsibility to represent the interests of another person (Bowles et al., 2005). It is even likely for both forces to be at play with some women being encouraged to compete by the representation role, and others being discouraged by the added pressure of the representation role. While the exact mechanisms behind shifts in attitudes towards performing in a competition are not under examination here, the experiment under consideration allows for specifically identifying the impact of differences in attitudes towards competition when representing another individual.

2.1.3. Role of differences in risk attitudes and feedback responses

When considering gender differences in risk attitudes, the economics literature generally finds that women are more risk averse than men both in the lab with low stakes and the field where higher stakes are involved (Croson & Gneezy, 2009; Eckel & Grossman, 2008). In terms of whether gender differences in risk attitudes affect competitiveness, a series of studies show that it plays a limited role in explaining gender differences in tournament entry. Niederle and Vesterlund (2007) and Healy and Pate (2011) provided participants with a choice that mimics the risk of the tournament-entry decision (i.e. asked participants to choose to enter past piece-rate performance into a tournament or receive a piece-rate compensation), and found risk aversion to be an insignificant factor. Cason et al. (2010), Wozniak et al. (2009), Sutter and Rützler (2010) control for risk aversion by directly asking participants to make a series of choices over incentivized lotteries, and also find that their measures do not explain the gender differences in entry. Similarly, studies in the psychology literature suggest that women tend to internalize negative feedback more than men (Dweck, 2000; Roberts & Nolen-Hoeksema, 1989), and thus, would be more averse to environments where they would receive feedback on their relative performance. However, in their analysis Niederle and Vesterlund (2007) find that feedback aversion also plays a negligible role in explaining gender differences in the choice to enter the tournament.

Given that the literature does not support strong *within* gender variations, it is not likely for the representation role to noticeably shift women's risk and feedback preferences. In other words, it is not expected for other-representing women to have significantly different risk and feedback preferences from self-representing women and, for these preferences to be shaping competition entry decisions.

In summary, based on the outlined conceptual framework, the effects of risk and feedback aversion on competition entry when representing another individual in a stereotypically male domain are expected to be minimal. In addition, differences in attitudes towards performing in a competition could energize male 'representative' competition entry, while the impact on women could be mixed. However, women 'representatives' are expected to hold significantly less confident beliefs about their relative ability, which would discourage them from entering into competitions. In contrast, men are expected to remain relatively more confident regardless of the treatment, with confidence not being a large determining factor for male 'representative's' choice to enter competition. In addition, as likely social-surplus maximizers, men may be more likely than women to enter competitions when representing another individual in their effort to maximize payoffs for both participants. Whereas women's tendency towards altruism and equalizing earnings could discourage them from entering a tournament that could have a zero-payoff outcome for the individual they are representing. Combined with gender differences in beliefs about relative ability, this lends support to the expectation that women may expect themselves to be top performer in order to enter competitions when representing another individual. Therefore, given the strong expected treatment and gender gap in confidence coupled with women's other-regarding preferences, controlling for differences in performance (discussed below), it is expected for female 'representatives' to experience lower rates of competition entry as compared to self-representing women (*Hypothesis I.A*). In addition, a significant representation effect on men's likelihood to enter competition is not expected (*Hypothesis I.B)*.

The next section discusses possible changes in competitive performance, another measure of competitiveness, as a result of the representation role.

2.2. Differences in performance under representation

Niederle and Vesterlund (2007) found no gender differences in performance in mix-gender competitive and non-competitive environments with performance in the male-stereotyped math task not being very responsive to the small differences in the incentive schemes. However, in the case where individuals need to represent another individual, their individual incentives remain constant, but the size of the pie is doubled as it involvs the payoffs of another individual. Therefore, being charged with the responsibility of representing someone else increases the financial stakes and places added pressure on the 'representatives'.

Laboratory and field studies seem to suggest that the gender performance gap increases with competitive pressure, with men's performance improving and women's performance not improving or diminishing in response to competition (Gneezy, Niederle & Rustichini, 2003; Niederle & Vesterlund, 2011). For example, Dilmaghani (2020) found that the gender gap in performance increases as the time constraint in chess games is tightened, with women underperforming as compared to men in fast chess games. Shurchkov (2012) compared gender differences in performance in math (i.e. stereotypically male) and verbal (i.e. stereotypically female) tasks. She found that in a mix-gender high-pressure treatment where subjects are allowed less time to solve each problem, the gender difference in the math task performance increased, with men significantly outperforming women in the tournament, and there was no gender difference in the verbal task. In addition, when representing others in a male domain, not only is the task at hand stereotypically male, but also perhaps, even the role of the individual as a 'representative'. This suggests the possibility of stereotype threat being activated for women 'representatives' in this high-pressure environment that involves a male-stereotyped task and role, and thus, hindering women's performance (Kiefer & Sekaquaptewa, 2007; Nosek, Banaji & Greenwald, 2002; Spencer & Steele, 1999; Steele & Aronson, 1995).

Furthermore, Paserman (2007) found that when evaluating point-by-point data from Grand Slam tennis tournaments, women were more likely to make unforced errors when the stakes are higher, at

crucial junctures of the match. In examining responses to a competitive fellowship program that awards graduate students who make quick advance to candidacy, J. Price (2008) found that this competition on speed had no effect on the time to candidacy for female students, while leading to substantial and significant decrease in time to candidacy for eligible male students. Moreover, Attali, Neeman and Schlosser (2011) document a greater gender gap in performance on the GRE in a real and competitive setting than in an experimental and low-stakes environment. They attribute this to men responding more strongly to decrease in incentives in the experimental GRE and performing better in a high stakes environment.

Following these insights, given that individuals are required to represent another individual in a mix-gender competitive environment involving a male-stereotyped task and role, with higher stakes involved, it is reasonable to expect shifts in the performance of men and women. Specifically, either no change or a decrease in female 'representative' performance as compared to self-representing women is expected (*Hypothesis II.A)*, and a boost in male 'representative' performance as compared to self-representing men (*Hypothesis II.B*).

The following section presents the experiment where I tested the impact of being a 'representative' on tournament entry and performance and identified the specific reasons behind any observed gender and treatment differences.

3. Experimental design

I examined the effect of representing others on competitive behavior and performance of men and women using a laboratory environment with a between-subject experiment design. Individuals were randomly assigned to either a control or treatment condition (Table 1).[2] The 'self-representation' control condition aimed to establish a baseline for competitive performance and behavior when individuals only represent their own interests. The control condition was then compared to the 'other-representation' treatment condition, where individuals faced the same tasks and decisions as in the control condition, but they also represented the interest of someone else. The individual treatment conditions and experimental procedures are explained in turn.

3.1. Self-Representation control condition

The self-representation control (SRC) condition followed the Niederle and Vesterlund (2007) mix-gender competition design and was computerized using Z-tree software program (Fischbacher, 2007).[3] Study participants were seated in groups of four, next to one another, consisting of two men and two women in each row of cubicles in the lab. Each participant had their own computer terminal that was separated by partition walls. While individuals could see each other and who is in their competition group, they could not see each other's computer screens, performance or decision information. Individuals were asked to participate in a real math task, which involved adding up a row of five two-digit numbers. Previous studies that measure explicit gender attitudes (Price, 2012) and utilize implicit association tests (Nosek et al., 2002) show that math is generally perceived to be a stereotypically male domain. Therefore, the math task fits the study's objective to understand how women and men behave in stereotypically male domains.

The numbers were randomly generated, and participants were asked to fill in the sum of these numbers in the blank box and press "Submit" for a new problem to appear (Fig. 1). Subjects were given five minutes to complete the task under four different compensation schemes. Each of the five-minute trials was referred to as a "Task" and performance was defined as "number of correctly solved problems" on a task. Participants received details on each task individually, right before they were asked to perform the specific task. They were also informed of their own performance during each task (number of problems solved correctly and incorrectly) but did not receive any information on the performance of other participants until the end of the study.

Task 1 allowed participants to gain practice with the task without needing to make any decisions regarding the compensation scheme and competition. Subjects performed the math task under a noncompetitive piece-rate, earning $0.50 per correctly solved problem. Task 2 allowed individuals to gain experience with the task in a competitive environment, but again, without needing to make any decisions about competition entry. All subjects were asked to perform the math task in a tournament amongst their four-person group. The person with the highest score received $2 per correct problem and the others received nothing. In case of a tie, the computer program chose the winner randomly. The outcome of this tournament was announced at the end of the experiment. As in the Niederle and Vesterlund (2007) design, gender was not explicitly mentioned in the experiment, but the participants were able to see their competitors who were seated in cubicles next to them.

After participating in the piece-rate and tournament-compensation schemes, individuals were asked to choose which payment scheme (piece-rate or tournament) they preferred for the third Task. Therefore, Task 3 was the competition entry environment. If participants selected the tournament-compensation, their performance in this task was compared to the performance of the three other group members from the previous tournament (Task 2). Therefore, their choice would not influence their team members' payoffs, which eliminated concerns about other-regarding preferences. Once the decision for Task 3 was made, they proceeded to perform the math task for five minutes.

To control for beliefs about relative ability, risk and feedback aversion, while eliminating factors such as competitive attitudes that shape preferences for performing in a competitive environment, participants were asked to perform an additional task. Task 4 mimicked the competition-entry choice in Task 3 but did not require individuals to perform in a tournament.[4] Participants were asked to choose between a

Table 1
Experiment treatment and control groups.

Condition	Number of Female Participants	Number of Male Participants	Number of Constituents
Self-Representation Control	40	40	None
Other-Representation Treatment	40	40	40 Men & 40 Women

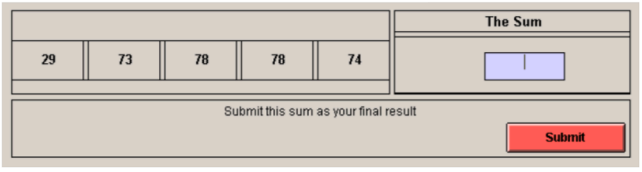

Fig. 1. Computer screen of sample math task.

[2] The sample size in this experiment replicated the Niederle and Vesterlund (2007) design with 40 women and 40 women in each experimental condition.

[3] Thanks to Muriel Niederle and Lise Vesterlund for sharing their Z-tree code.

[4] Therefore, as in Niederle and Vesterlund (2007) Task 4 is used "to distinguish between the role played by gender differences in preferences for performing in a competition, and the more general explanations such as gender differences in overconfidence and risk and feedback aversion." This experimental design does not have an independent measure of risk aversion.

piece-rate and tournament-compensation scheme for an already completed task: their Task 1 performance. After the Task 4 decision was made, to assess beliefs about relative ability, participants were asked to rank their performance in Task 1, Task 2, and Task 3 relative to that of other group members. Individuals could earn an extra $3 as those who correctly ranked their performance on each task received $1 for each correct ranking.[5] Hence, controlling for performance, beliefs about relative ability and the decision in Task 4, any gender gap in Task 3 can be attributed to differing attitudes towards having to compete against others. At the end of the experiment, one of the four tasks was randomly chosen for payment.

The self-representation control (SRC) condition was conducted in twelve sessions at the Harvard Decision Science Laboratory in Cambridge, MA with a total of 80 laboratory subjects participating (40 men and 40 women). This totaled twenty competition groups. Participants were provided with four sheets of blank white paper and a pen, as calculators were not allowed. Standard recruiting procedures were used and signing up for the study was restricted to students only. This was in response to pilot sessions that did not include the student restriction on the subject pool, which yielded sessions with mostly female Harvard students and significantly older male seniors competing against each other. The student requirement is standard practice in many laboratory experiments. For this experiment, given the location of the study, it yielded participants who were majority Harvard University students. This led to a homogenous sample of generally ambitious and qualified individuals to tackle this math task.

The experiment lasted approximately forty-five minutes. All participants remained anonymous throughout the study and were only identified by their randomly assigned code numbers. The instructions were read out loud and can be found in Appendix A. The math task was described to the participants and for each task, their decision, and the procedure to determine their earnings was explained. Participants were informed of their earnings at the end of the study. Subsequently, they were asked to answer a demographic questionnaire and were paid in cash (their study earnings plus a $5 show-up fee and $7 participation fee) at the end of their session. Average earning per participant was $22.49 in the self-representing control condition.

3.2. Other-representation treatment condition

To test the impact of representing others on competitive performance and behavior, the other-representation treatment (ORT) replicated the design of the self-representation control condition with only one difference: the participants who performed Task 1 through Task 4 were also required to represent the interests of another study participant in addition to their own (i.e. their payoffs also determined the payoffs of another individual). Therefore, the only difference between the self- and other-representation conditions in terms of monetary payoff was that representatives in other-representation treatment worked for a higher payoff for another randomly assigned study participant who remained anonymous but was physically present in the room. This replicated the Bowles et al. (2005) design, where individuals represented the interests of someone else in a negotiation setting, the individual economic incentives were held constant across these two conditions but the amount of their earnings was added to their constituent's earnings. Hence, while individual incentives remained constant across the two conditions, the size of the pie was increased in the other-representation treatment. Representatives may thus respond to both features of serving as a representative that are included in this treatment: the ability to help someone else and the larger pie. Thus, being in the representative role includes higher stakes and more pressure. The theory laid out in the previous sections helps in understanding which of the two features men and women are expected to respond to.

The other-representation treatment consisted of twenty sessions that were conducted using the same subject pool, recruiting procedures, and requirements as the self-representation control condition. Each session included eight individuals (total of 160 subjects) who were randomly assigned into two groups of four. Four of the individuals (two men and two women) were the 'representatives' who participated in the study and were paid based on their own decisions and performance (as in the control). The remaining four individuals (two men and two women) were the 'constituents'. They did not participate in the study. Instead, 'constituents' performed an alternative coding task for which they were not monetarily compensated and the earnings of a randomly assigned 'representative' determined their study earnings. Previous research in negotiation environments has shown that negotiation behavior of the representative was impacted by the gender of their constituent (Pruitt, Carnevale, Forcey & Van Slyck, 1986).[6] To make sure that 'representatives' in the other-representation treatment do not know the gender of their 'constituent', an equal mix of female and male 'constituents' was recruited for each session and they were randomly assigned to each 'representative.'

Hence, the other-representative treatment consisted of twenty competition groups (as in the control) with eighty 'representatives' and eighty 'constituents' (40 men and 40 women of each). The terms 'representatives' and 'constituents' were not used in the experiment. Instead, 'representatives' were identified as "yellow-labeled participants" and 'constituents' were identified as "green-labeled participants." Individuals were randomly assigned to either the green- or yellow-labeled cubicles by selecting an index card from a cup when they arrived at the lab and signed in for the study. The cup consisted of randomly ordered and folded white index cards with "green" or "yellow" written inside. At the start of the experiment, individuals with cards that read "green" were asked to sit in the row of cubicles labeled with green sticky notes and those with cards that read "yellow" were asked to sit in the row of cubicles labeled with yellow sticky notes.

All participants were paid a $5 show-up fee, a $7 participation fee and additional earnings from the study, which was based on either the individual's own performance (in the case of 'representatives') or a randomly selected individual's performance (in the case of 'constituents'). The average earning per participant was $23.30. As in the self-representation control condition, the experiment was computerized using Z-tree software program, lasted approximately forty-five minutes, all participants remained anonymous throughout the study and the instructions were read out loud and can be found in Appendix A. Constituents' followed along as the math task was described to the 'representatives' and for each task, their decision and the procedure to determine the representative's earnings and the earnings of their 'constituent' were explained. At the end of their session, all participants were informed of their earnings, asked to answer a demographic questionnaire, and were paid in cash. The next section discusses the experimental results.

4. Results

First, I present individual competitive behaviors in the control condition (i.e. Self-Representation Condition) where the Niederle and Vesterlund (2007) experiment was replicated. I found that female and male participants behaved very similarly, with the likelihood of choosing the tournament compensation scheme (instead of the piece-rate) in Task 3 being 65 percent for women and 67.5 percent for men. Therefore, there is no significant gender difference in the choice to enter a tournament in the self-representation condition (a Fisher's exact test yields p-value = 0.81). This contrasts with Niederle and Vesterlund's (2007) results, which found a significant difference between men and women's choice

[5] In the event of a tie, participants were compensated for any guessed ranking that could be correct.

[6] In Bowles et al. (2005)

of compensation in Task 3 with 73 percent of men and only 35 percent of women selecting the tournament (a Fisher's exact test yields p-value = 0.002). These differences could be due to differences in subject pools. This study was restricted to a student subject pool with 84 percent of subjects being Harvard University students, which unlike Niederle and Vesterlund (2007), yielded no significant gender differences in confidence in this control condition (see Section 5). Price (2010) did not replicate Niederle and Vesterlund's results either. They also used a student-only subject pool and Niederle and Vesterlund (2011) explain that the Price (2010) findings were due to lack of gender differences in confidence levels, perhaps as a result of having a more homogenous student sample.[7]

In addition, as shown in Fig. 2, women and men had very similar performances in both the piece rate round (Task 1) and the tournament round (Task 2). In the piece rate (Task 1), women and men correctly solved an average of 11.8 and 12.3 problems, respectively, with this difference being insignificant using a two-sided t-test (p-value = 0.65). Similarly, in the Task 2 tournament, women solved an average of 13.2 problems correctly and men solved 13.6 problems. This difference is also statistically insignificant (p-value = 0.65). Men did demonstrate a larger variance in their piece-rate and tournament performances than women, but only the differences in piece-rate score variances are significant (Variance ratio test, p-value = 0.08). There is also a strong correlation between performances in the piece-rate and tournament rounds (Spearman rank correlations of 0.67 for women and 0.84 for men); however, both men and women performed significantly better in the tournament than in the piece-rate round (one-sided $p < 0.01$ for each gender separately). This suggests that both men and women became equally competitive in a competitive environment in the self-representation condition. Though, as Niederle and Vesterlund suggest, this increase in performance could also be due to learning and the different incentives across the two tasks.

Consequently, there remained no gender differences in the probability of winning the Task 2 tournament given the similar performances by men and women. In fact, of the twenty Task-2 tournaments, women won 50 percent and men won 50 percent. To evaluate the probability of winning the tournament, I created random four-person groups from the observed performance distributions, following Niederle and Vesterlund (2007). Conditioning only on gender, the probability of winning the tournament was 27.5 percent for a man and 22.5 percent for a woman and this difference is not significant using a Fisher's exact test (p-value = 0.80). Conditioning on performance, in a sample of those who scored 13 or greater,[8] the probability of winning the tournament is 45.83 percent for a man and 36.36 percent for a woman (p-value = 0.56). The probability of winning the tournament increases to 52.38 percent for a man and 50 percent for a woman among those who scored 14 or greater (p-value = 1.00). Therefore, there were no significant gender differences in performance in the Task 2 tournament or in the probability of winning the tournament in the self-representation control condition.

Hence, in the self-representation treatment, men and women behaved very similarly; both in terms of performance in non-competitive (Task 1) and competitive (Task 2) environments and in terms of their choice to enter a competitive environment where they needed to subsequently perform (Task 3). In addition, men and women behaved similarly in their choice to enter a tournament when they did not need to perform in a competition (Task 4). In fact, 55 percent of men chose to submit their Task 1 piece-rate performance to a tournament (Task 4) and 47.5 percent of women made the same decision, with no significant gender differences (p-value = 0.51). Thus, men and women in the self-representation control condition exhibited very similar behaviors across all four tasks and environments.

This result places this study in a unique position to examine how the competitive behavior of similarly ambitious and qualified men and women, such as men and women in leadership positions, is impacted by the responsibility of representing another person's interests. Accordingly, I present the results of the other-representation treatment next, where individuals faced the same tasks and decisions as in the control condition, but they were also responsible for the interest of another study participant who was randomly assigned to them. Overall, I did not find any evidence in support of the Bowles et al. (2005) competitive negotiation environment findings being transferred to this competitive environment. Bowles et al. (2005) find that women were "particularly energized" when they were responsible for representing the interests of another person, negotiating significantly higher compensations as compared to women who only represented themselves. In contrast, in this competitive environment, I found that not only were female 'representatives' not "particularly energized", but they *leaned back* from entering into competitions when compared to self-representing women. In fact, I find that the male 'representatives' were the ones who were "particularly energized" in terms of performance when representing the interests of someone else in a competition.

The following sections expand on these results by first presenting a comparison among women across treatment conditions in Section 4.1, then a comparison among men across treatment conditions in Section 4.2, and finally, combining the entire subject pool and presenting the key experimental results in Section 4.3.

4.1. *Self-representing women vs. other-representing women*

In negotiation contexts, women have been found to negotiate less than men when negotiating on their own behalf, and when they represented someone else, women were "energized", became more competitive and negotiated more strongly. However, in this competitive environment, I started with a different reference point since women showed the same competition behavior as men when representing themselves (both in terms of tournament entry and performance). Moreover, when women had to represent someone else, they became less likely to enter a tournament (*Hypothesis I.A*), despite experiencing similar performance across the two conditions (*Hypothesis II.A*).

Compared to the self-representation control condition where 65 percent of women chose to compete in Task 3, the likelihood of competing significantly declined by 0.28 percentage points to 38 percent for female 'representatives' (two-tailed Fisher's exact test, p-value = 0.025). This was also significantly lower than the 58 percent of male 'representatives' who selected to compete in the other-representation condition (one-tail Fisher's Exact Test, p-value = 0.058). Hence, not only were female 'representatives' not "particularly energized" in this competitive environment, but they *leaned back* and shied away from entering into competitions (Fig. 3).

This is despite the fact that women exhibited similar performance patterns across treatment conditions, with no significant differences in their average performance in the first two tasks (Fig. 4).[9] As shown in Fig. 4, women's level of performance in the first two tasks was very similar regardless of whether or not they were just representing their

[7] Moreover, Price (2010) found no significant gender differences in competition entry when limiting the scope to only undergraduate students across Niederle and Vesterlund's study and his subject pool.

[8] In expected value terms, if individuals solved 12 correct answers, then risk-neutral rational decision makers would be indifferent between choosing the piece rate (EV = 12 x \$0.50 = \$6) or the tournament (EV = 12 x \$2 x 0.25 = \$6) compensation scheme.

[9] In the piece rate, the average number of problems solved by women is 11.8 in the control condition and 12.5 when representing others' interests. This difference is not significant using a two-sided t-test ($p = 0.55$) or a one-sided t-test ($p = 0.27$). In addition, in the tournament, the average number of problems solved by women is 13.2 in the control condition and 13.8 when representing others' interests. This difference is not significant using a two-sided t-test ($p = 0.54$) or a one-sided t-test ($p = 0.27$).

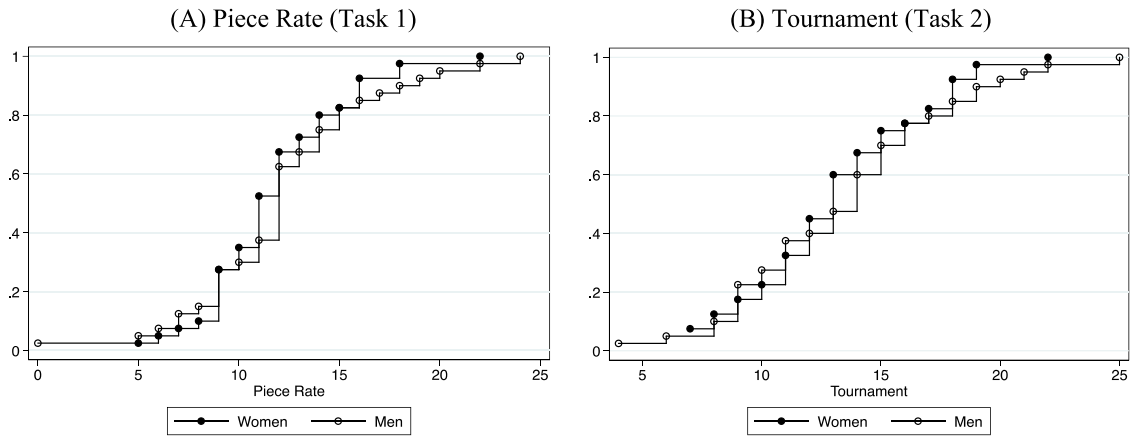

Fig. 2. CDF of number of correctly solved problems, self-representation control condition.

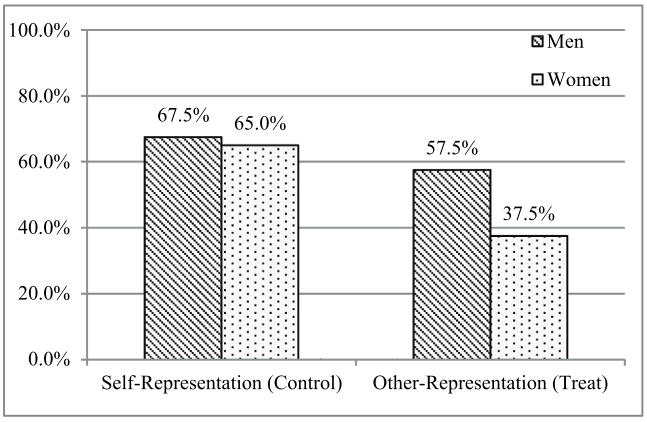

Fig. 3. Percentage of individuals choosing to enter tournament in Task 3.

own interests or the interest of another individual in addition to their own. Therefore, when representing another individual, there was no significant increase in women's performance.

Moreover, in the other-representation condition, women who competed in Task 3 performed significantly better than those who did not, both in the piece-rate (p-value = 0.04) and tournament (p-value = 0.06), and this significance is similar to the women in the self-representation condition (Table 2). Although the average female performance of those who competed was higher in the other-representation treatment than the self-representation condition, this difference was not significant for either task (p-value = 0.22 and p-value = 0.30 for piece rate and tournament, respectively).

In addition, the results from Task 3 performance were very similar to those of Task 2 across treatments. The participants who entered the tournament in the other-representation treatment did not have a significantly different increase in performance in the choice task (Task 3) relative to the former (Task 2) tournament than those who entered the tournament in the control condition (p-value = 0.48 overall, p-value = 0.61 for women and p-value = 0.21 for men). Comparing women who entered the tournament in the other-representation treatment and self-representation control condition, there were no significant increases in performance in the choice task (Task 3) relative to the former (Task 2) tournament (p-value = 0.61). In addition, women experienced the same difference in Task 3 performance across treatments between those who competed (p-value = 0.39) and those who did not (p-value = 0.57).

Table 2
Average female performance in piece rate (task 1) and tournament (task 2) by tournament entry decision (task 3).

Piece Rate (Task 1) Performance	Compete	Don't compete	p-value
Self-Representation	12.58	10.43	0.06
Other-Representation	14.87	11.04	0.04
p-value	0.22	0.57	
Tournament (Task 2) Performance	**Compete**	**Don't Compete**	**p-value**
Self-Representation	14.08	11.43	0.03
Other-Representation	16.07	12.48	0.06
p-value	0.30	0.34	

Note: p-values represent two-sample t-test of means.

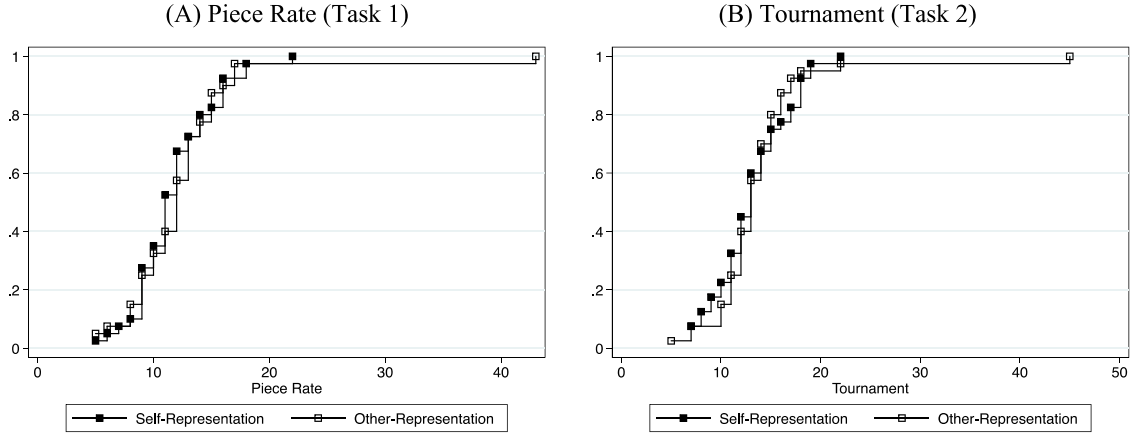

Fig. 4. CDF of Number of Correctly Solved Problems by Women Across Treatments.

However, in the self-representation control condition, women who chose the tournament performed significantly better than those who chose the piece-rate (one-tail *t*-test, *p*-value = 0.01), and the same was true for women in the other-representation treatment (one-tail *t*-test, *p*-value = 0.02).

Additionally, based on their performance on Task 2 and Task 3, women in the other-representative treatment entered the tournament less than those in the control condition (a one-sided Fisher' exact test yields *p*-value = 0.099 and 0.039 for Tasks 2 and 3, respectively). In fact, Fig. 5 shows that the proportion of women in the self-representation condition who selected tournament for Task 3 is higher for every performance quartile of Task 2 and Task 3 than women in the other-representation treatment.[10] Hence, despite similarities in performance across the two conditions, when compared to their self-representing counterparts, female 'representatives' leaned back from entering the competition.

4.2. Self-representing men vs. other-representing men

In this competitive environment, the male 'representatives' were the ones who were "particularly energized" in terms of experiencing enhanced performance (*Hypothesis II.B*). In fact, the average number of problems solved in the Task 1 piece-rate by male 'representatives' was 14.4, which was significantly higher than the 12.3 average male performance in the control condition (one-sided *t*-test, *p*-value = 0.045), and the 12.5 average female 'representative' performance (one-sided *t*-test, *p*-value = 0.08). In addition, as shown in Fig. 6, in the Task 2 tournament, the average number of problems solved by male 'representatives' was 15.3, which was marginally higher than the 13.6 male average of correctly solved problems in the control condition (one-sided *t*-test, *p*-value = 0.099); and while higher than the 13.8 average female 'representative' performance, the difference was not significant (one-sided *t*-test, *p*-value = 0.16).

Additionally, men were more competitive and aggressive in their performance following the choice task (Task 3). While men who chose the piece-rate did not experience a significant difference in Task 3 performance across treatments (*p*-value = 0.79), men who chose the tournament experienced a significantly higher Task 3 performance when representing others (Fig. 7). In Task 3, the men who chose to compete in the other-representation treatment solved an average 19 correct answers on the task, compared to an average performance of 13.8 by the men who selected the tournament in the self-representation treatment (two-sided *t*-test yields *p*-value = 0.007).

Given this enhanced performance by male 'representatives', it could be expected that they also experienced higher rates of tournament entry as compared to their self-representative male counterparts (*Hypothesis I.B*). This was not the case, however. In fact, there was no significant increase in the share of men who entered the tournament in the other-representative treatment as compared to men in the self-representative control group. In the self-representation condition 68 percent of the men entered the tournament in Task 3, which was not significantly different from the 58 percent of men who chose to compete in the other-representation treatment (two-tailed Fisher's exact test, *p*-value = 0.49).

4.3. Self-representing control vs. other-representing treatment

Overall, the most significant difference across the self-representation control condition and the other-representation treatment was that despite having similar performance scores, female 'representatives' leaned back and shied away from entering into a tournament as compared to women who only represented themselves. Table 3, Column (1) presents the regression analysis that confirms this result. The table presents Probit regression results with marginal effects reported in percentage points, where the dependent variable is the Task-3 choice of compensation scheme (1 = tournament and 0 = piece rate). The "Piece-rate" variable refers to Task-1 performance, "Tournament" variable refers to Task-2 performance, and "Represent" variable refers to the treatment condition (1 = Other-Representation Treatment and 0 = Self-Representation Control condition). The results in Column (1) show that controlling for past performance, women were much less likely to select a competitive-compensation scheme when charged with representing the interest of another individual.

In fact, for an average female study participant, there was a significant representation effect: being a representative significantly reduced her likelihood of entering the tournament by 37 percentage points. When comparing with an average male 'representative' in the other-representation condition (Table 3, Column (4)), female 'representatives' were 26 percentage points less likely to enter a tournament, though this effect is marginally significant. In contrast, women's decisions in the self-representation condition did not differ significantly from men (Column (3)), and there was no significant representation effect on men's decisions as compared to the men in the control (Column (2)). All in all, controlling for past performance, female 'representatives' leaned back and were significantly less likely to compete in Task 3 as compared to everyone else (Column (5)). The following section discusses the possible explanations for this finding.

5. Why female 'Representatives' leaned back from competition entry

The results suggest that differences in confidence levels play a significant role in explaining why female 'representatives' leaned back from tournament entry. To evaluate participants' beliefs on their relative tournament performance they were asked at the end of the experiment to guess how their performance in Task 2 ranked relative to the other members of their group. Participants received $1 if their guess was correct, and in the event of a tie they were compensated for any guess that could be correct. In the other-representation condition, the guesses of women and men differed significantly from one another (Fisher's exact test of independence *p*-value = 0.005).

Men were more confident about their relative performance than women in the other-representation treatment (Fig. 8). While 55 percent of the men in both the self- and other-representation conditions believed that they were the best in their group of four, only 25 percent of the women in the other-representation treatment held this belief (Fisher's exact test *p*-value = 0.01).[11] This was also less than the 40 percent of the women in the self-representation condition who believed that their performance ranked the highest, though this difference is not significant

[10] Among women whose expected earnings are lower in the tournament, those in the self-representation condition are more likely to enter a tournament than those in the other-representation condition (a one-sided Fisher' exact test yields $p = .072$ and .085 for Tasks 2 and 3, respectively). Thus, from a profit-maximizing perspective, low-performing women enter the tournament too often in the self-representation condition, and high-performing women enter it too rarely in the other-representation condition.

[11] In other-representation treatment, 25 percent of women believed they were a top performer. Of those women, 40% were the top performer in their group and correctly guessed their ranking. Of the 55 percent of men who believed that they were a top performer, 54 percent were the top performer and correctly guessing their ranking. The differences between the proportion of men and women who correctly ranked themselves as a top performer are not statistically significant (Fisher's exact test *p*-value = 0.13).

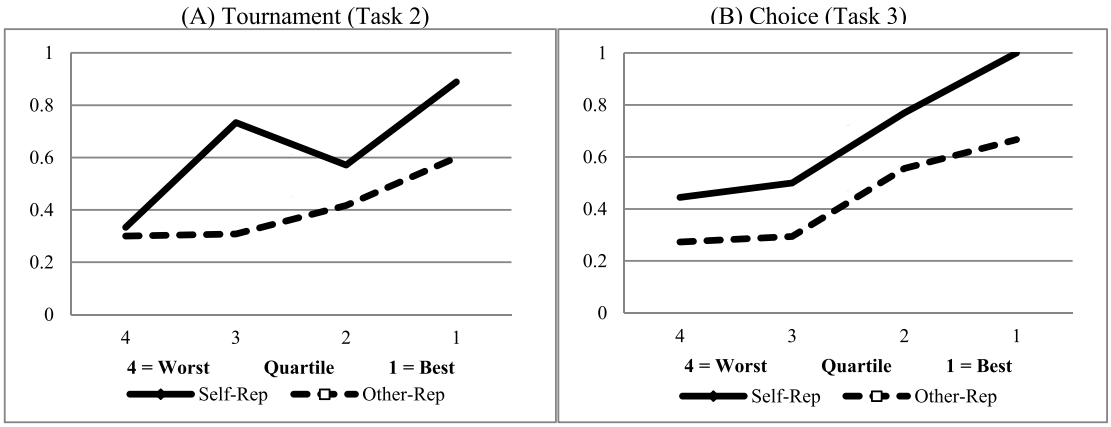

Fig. 5. Proportion of women selecting tournament for task 3 conditional on task 2 performance quartile (A) and task 3 performance quartile.

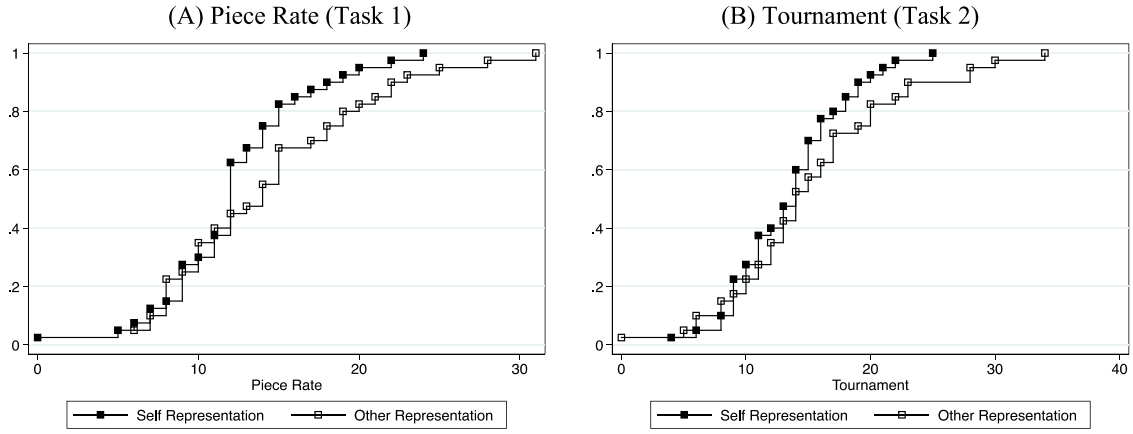

Fig. 6. CDF of number of correctly solved problems by men across treatments.

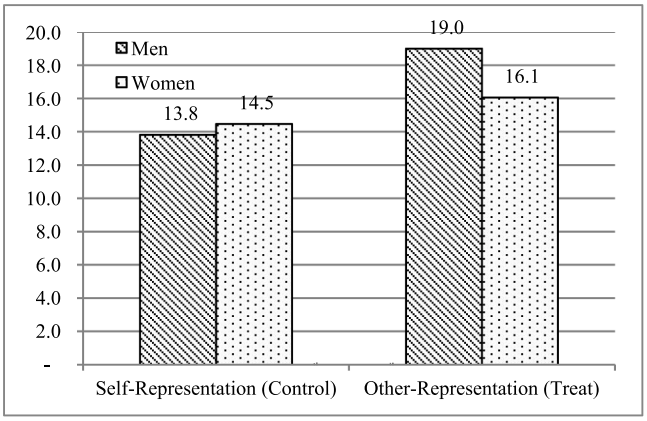

Fig. 7. Number of correctly solved problems by those who chose to compete (task 3).

(Fisher's exact test p-value = 0.23).[12]

To further examine why women in the other-representation treatment entered the tournament less frequently than the women in the self-representation condition, we turn to the proportion of women who entered the tournament across treatments for each guessed rank (Fig. 9). Looking at the women who believed that they were the top performer in their group, there was absolutely no treatment gap in tournament entry (Fisher's exact test p-value = 1.00). However, these women only accounted for 25 percent of other-representing women and 40 percent of self-representing women (Fisher's exact test p-value = 0.23). In contrast, amongst the women who believed that they were the second best in their group there was a significant treatment gap in tournament entry of about 36 percentage points (Fisher's exact test p-value = 0.04).[13] As compared to women in the self-representation condition, it appears that women who were responsible for the interest of someone else needed to

[12] Relative to their actual rank, both men and women were overconfident in the self-representation condition. Overall, the guesses of women and men in the self-representation condition did not differ significantly from one another and a Fisher's exact test of independence of the distributions for men and women delivered p-value = 0.445. In addition, the guesses of women across these two treatments did not differ significantly from one another; a Fisher's exact test of independence of the distributions delivered p-value = 0.383.

[13] There were no significant differences across treatments among men. Seventy-three percent of men who ranked themselves as first or second in the self-representation condition competed and 70 percent competed in the other-representation condition (two-tailed t-test, p-value= 0.75).

[14] None of the women who ranked themselves as fourth chose to the tournament compensation scheme for Task 3.

Table 3
Probit regression of tournament choice in task 3, marginal effects.

	(1) Women	(2) Men	(3) Self-Representatives	(4) Representatives	(5) All Participants
Piece-Rate	0.006	−0.006	0.012	−0.023	0.002
	(0.025)	(0.021)	(0.024)	(0.028)	(0.015)
Tournament	0.061**	0.040*	0.026	0.079***	0.035**
	(0.027)	(0.021)	(0.022)	(0.030)	(0.016)
Represent	−0.365***	−0.170			−0.180
	(0.120)	(0.114)			(0.117)
Female			−0.052	−0.264*	−0.072
			(0.112)	(0.145)	(0.109)
Represent*Female					−0.396***
					(0.116)
Observations	79	76	80	75	155
PsuedoR2	0.310	0.168	0.111	0.401	0.208

The table presents Probit regression results with marginal effects reported in percentage points. The dependent variable is Task-3 choice of compensation scheme (1 = tournament and 0 = piece rate). Piece-rate refers to Task-1 performance and Tournament refers to Task-2 performance. Represent refers to the treatment condition (1 = Other-Representation Treatment and 0 = Self-Representation Control condition). Demographic variables such as age, race, and household income level as well as the month the experiment was conducted are also controlled for. Column (1) presents marginal effects evaluated at means for a woman in the self-representation treatment. Column (2) presents marginal effects evaluated at means for a man in the self-representation treatment. Similarly, columns (3) and (4) presents marginal effects evaluated at means for a man. Column (5) presents marginal effects evaluated at means for a man in the self-representation treatment. Ranking guesses of 4, were excluded from the analysis. *** Significance at the 1 percent level. ** Significance at the 5 percent level. * Significance at the 10 percent level.

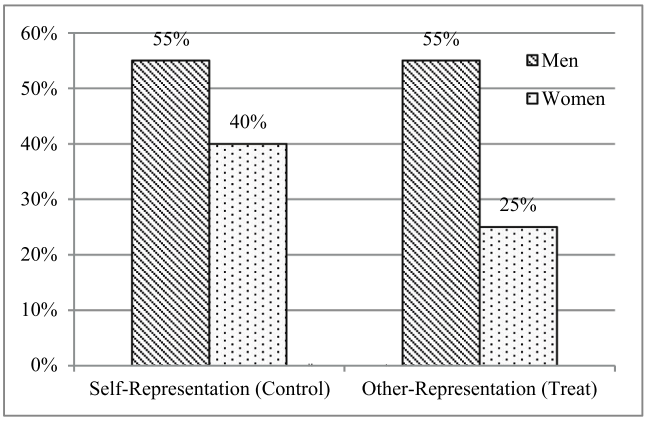

Fig. 8. Percent of individuals believing they are the top performer in the task 2 tournament.

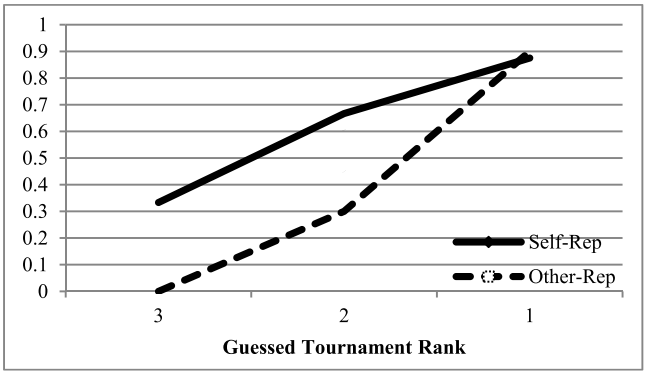

Fig. 9. Proportion of women selecting the tournament for task 3 conditional on guessed rank.[14]

be extremely confident about their relative performance in order to select the tournament in Task 3. In fact, when female 'representatives' believed that they were the top performer, 90 percent chose the tournament, but when they believed that they were the second-best performer, that number declined to 37.5 percent. This suggests that women's other-regarding preferences may be at play, with women being concerned about their 'constituent' ending up with earnings of zero if they are not the top performer. Yet, when they believed that they ranked first, they proceeded with the goal of maximizing the size of the pie for themselves and their 'constituent' and therefore, chose to compete.

This is confirmed by the regression analysis in Table 4. When the guessed tournament ranking of one and two are controlled for, the marginal effects of both binary variables are statistically significant, while the representation effect is nearly eliminated and is no longer significant. As also shown in Fig. 9, the representation effect was mostly driven by differences amongst women who guessed their performance to be less than the best. When women believed that they were the top performer in both the self-representing and other-representing conditions, they are just as likely to enter competitions. However, when women believed that they performed less than the best in their group, women in the other-representing condition were significantly less likely to join the tournament in Task 3. Given that when women ranked themselves as the top performer, they were just as likely to choose the tournament compensation scheme across treatments, this suggests that treatment differences in women's confidence levels explain the representation effect.

Moreover, differences in confidence levels appear to be driving the gender gap in tournament entry in the other-representing condition. When the guessed tournament rankings are controlled for (Column (4)), the gender effect is nearly eliminated and is no longer significant (compared to Column (3)). Contrary to the comparison with women in the control condition, guessed ranking of two is not statistically significant. This suggests that what is driving the gender difference in tournament entry is the significantly lower share of women who believe that they are the top performer as compared to the majority of men who hold the same belief in the other-representative condition.

To rule out other factors such as risk and feedback aversion, I also examined differences in individual tournament entry decisions in Task 4. Recall that in Task 4 participants were asked to select one of two compensation schemes for their Task 1 piece-rate performance, either a piece-rate or a tournament. If they chose the piece-rate, they earned $0.50 for each correctly solved problem in Task 1. If they selected the tournament compensation, their Task 1 performance was submitted into a competition against the Task 1 performances of the other participants in their group (independent of the other group members' choice of compensation scheme). If the individual's performance was higher than the other three players, they would win the tournament and earn $2 for

Table 4
Probit regression of tournament choice in task 3, marginal effects.

	(1) Women	(2) Women	(3) Representatives	(4) Representatives
Piece-Rate	0.006	0.032	−0.023	−0.009
	(0.025)	(0.027)	(0.028)	(0.040)
Tournament	0.061**	−0.004	0.079***	0.064
	(0.027)	(0.014)	(0.030)	(0.047)
Represent	−0.365***	−0.075		
	(0.120)	(0.075)		
Female			−0.264*	−0.073
			(0.145)	(0.163)
Guessed 1 as Tournament Rank		0.904***		0.446**
		(0.084)		(0.220)
Guessed 2 as Tournament Rank		0.681***		−0.157
		(0.134)		(0.226)
Observations	79	79	75	75
PsuedoR2	0.310	0.526	0.401	0.533

The table presents Probit regression results with marginal effects reported in percentage points. The dependent variable is Task-3 choice of compensation scheme (1 = tournament and 0 = piece rate). Piece-rate refers to Task-1 performance and Tournament refers to Task-2 performance. Represent refers to the treatment condition (1 = Other-Representation Treatment and 0 = Self-Representation Control condition). Guessed 1 Tournament Rank equals 1 if individual ranked their performance as the best in their group, and 0 otherwise. Guessed 2 Tournament Rank equals 1 if individual ranked their performance as the second best in their group, and 0 otherwise. Demographic variables such as age, race, and household income level as well as the month the experiment was conducted are also controlled for. Column 1 presents marginal effects evaluated at means for a woman in the self-representation treatment. Column 2 presents marginal effects evaluated at means for a woman in the self-representation treatment who believed that she ranked third in Task 2 performance. Column 3 presents marginal effects evaluated at means for a man in the other-representation treatment. Column 4 presents marginal effects evaluated at means for a man in the other-representation treatment who believed that he ranked third in Task 2 performance. Ranking guesses of 4, were excluded from the analysis. *** Significance at the 1 percent level. ** Significance at the 5 percent level. * Significance at the 10 percent level.

every correctly solved problem. However, if they lost the tournament, they earned nothing in Task 4.

In the self-representation control condition, 47.5 percent of women and 55 percent of men submitted their Task 1 piece-rate performance to the tournament in Task 4, yielding no significant gender differences (Fisher's exact test, p-value = 0.655). However, in the other-representation condition, significantly fewer women (22.5 percent) submitted their Task 1 piece-rate performance to the tournament as compared to men (45 percent), (Fisher's one-tail exact test, p-value = 0.029). Across the treatments, there were no significant differences in men's Task 4 decisions (Fisher's exact test, p-value = 0.503). Conversely, women in the other-representation treatment were significantly less likely to submit their Task 1 piece-rate performance to the tournament than women in the self-representation treatment (Fisher's one-tail exact test, p-value = 0.017). Thus, women in the other- and self-representation treatments differed in their compensation scheme choices even when the decision did not have the added risk of having to perform in a competition.

Performance on Task 1 appears to be a significant factor in individual decisions to choose the tournament compensation scheme in Task 4. In the self-representation condition, both men and women who submitted to the tournament solved significantly more problems (13.1 for women and 13.8 for men) than those who did not submit to the tournament (10.7 for women and 10.3 for men; p-value = 0.03 for women, p-value = 0.047 for men). These differences were even more pronounced in the other-representation treatment. Women who competed, solved 16.7 problems, and women who did not, solved 11.2 problems correctly (p-value = 0.01). In addition, men who competed, solved 18.9 problems, and men who did not compete, solved 10.6 problems (p-value = 0.00).

As shown in Fig. 10 (Panel A), across all performance quartiles, women in the self-representation treatment were more likely than those in the other-representation treatment to submit their Task 1 piece-rate performance to a tournament in Task 4. The regression analysis in Column (1) of Table 5 shows that when controlling for performance scores, women in the other-representation treatment were significantly less likely to submit their Task 1 piece rate into a tournament in Task 4 than women in the self-representing condition. However, when performance is controlled for, there are no significant gender differences in the Task 4 choice of 'representatives'.

In addition to performance in Task 1, individual beliefs about their relative ability appear to have played a significant role in Task 4 choices. As shown in Fig. 10 (Panel B), when female 'representatives' believed that they were the top performer, they were even slightly more likely to submit their piece-rate to the tournament than the women in the self-representation condition. However, only 20 percent of female 'representatives' believed that they were the top performer in Task 1, as compared to 37.5 percent of women in the self-representing condition, and this difference is marginally significant (one-tailed Fisher's exact test, p-value = 0.07). The regression analysis in Column (2) of Table 5 confirms this result. Including controls for guessed piece-rate rank eliminates the significance of the representation effect on women in Task 4, with individuals who guessed their rank to be one being significantly more likely to submit their Task 1 performance into a tournament. Given that when women ranked themselves as the top performer, they were just as likely to choose the tournament compensation scheme across treatments, this further confirms that treatment differences in women's confidence levels explains the representation effect in Task 4. This also suggests that conditioning on beliefs, factors such as risk and feedback aversion had an insignificant effect on the Task 4 decision, where there was no expectation of performing in a competition.[15]

Furthermore, to rule out the impact of differences in attitudes towards performing in a competition, the Task 4 decision is incorporated into the Task 3 tournament entry analysis. Recall that the only difference

[15] Following Niederle and Vesterlund (2007), to further confirm that risk aversion did not play significant a role, I compared women across treatments with fourteen or more correct answers in the Task-2 tournament that had a 35 percent or higher chance of winning the tournament. The decision to enter the tournament can be evaluated as a gamble of receiving, per correct answer, either $2 with a probability of 35 percent (or more) or receiving $0.50 for sure. For participants who had fifteen correct answers, that means a gamble of a 35 percent chance of $30 (i.e., an expected value of $10.50) versus a sure gain of $7.50. Of the participants who solved fourteen problems or more, a greater percentage of women in the other-representation treatment (nine out of seventeen, 53%) did not take this or a better gamble than the women in the self-representation condition (four out of sixteen, 25%). This difference is significant with a one-sided Fisher's exact test (p-value = 0.099). Similarly, for participants who have thirteen or fewer correct answers, the chance of winning the tournament is five percent or less. Thus, entering the tournament means receiving $2 per correct answer with a probability of five percent (or less) versus receiving $0.50 for sure. For participants who solve twelve correct answers, this was a choice between a five percent chance of winning $24 (i.e., an expected value of $1.20) compared to receiving $6 for sure. Of the participants, who solved twelve problems or less, ten out of eighteen (55.6%) of the women in the self-representation condition and four out of sixteen (25%) of the women in the other-representation condition took this or a worse gamble. This difference is significant with a two-sided Fisher's exact test (p-value = 0.09; p-value = 0.07 one-sided test). To explain these choices, women would have to be exceptionally risk-averse in the other-representation treatment and exceptionally risk seeking in the self-representation condition. However, such a significant difference in risk preferences amongst women is highly unusual, suggesting that other factors besides risk aversion may be playing a role.

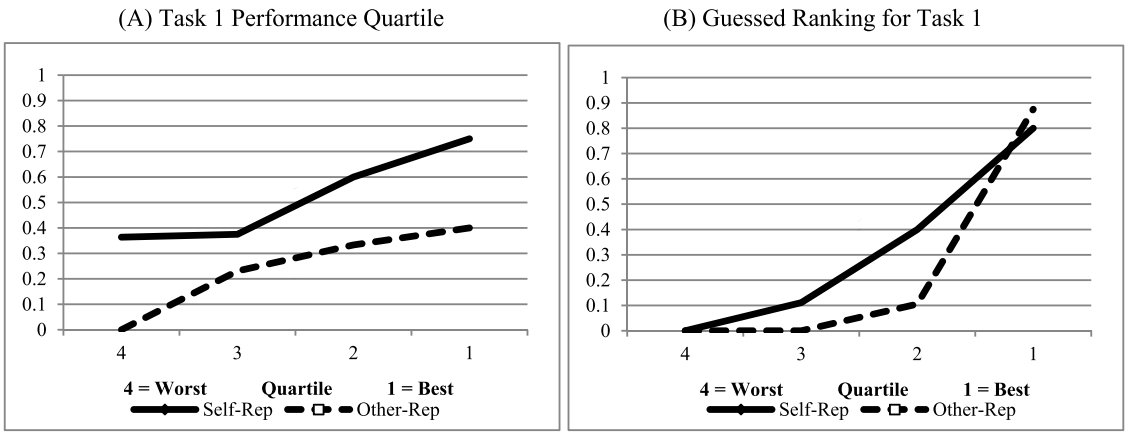

Fig. 10. Proportion of women selecting tournament for task 4 conditional on task 1 performance quartile (A) and guessed task 1 rank.

Table 5
Probit regression of task 4 choice, marginal effects.

	(1) Women	(2) Women	(3) Representatives	(4) Representatives
Piece-Rate	0.0680***	0.005	0.0770***	0.000
	(0.022)	(0.008)	(0.018)	(0.000)
Represent	−0.270**	−0.050		
	(0.119)	(0.048)		
Female			−0.078	0.000
			(0.134)	(0.000)
Guessed 1 as Piece-Rate Rank		0.763***		0.980***
		(0.121)		(0.031)
Guessed 2 as Piece-rate Rank		0.187		0.125*
		(0.119)		(0.075)
Observations	77	77	76	76
PsuedoR2	0.246	0.382	0.428	0.757

The table presents Probit regression results with marginal effects reported in percentage points. The dependent variable is Task-4 choice of compensation scheme (1 = tournament and 0 = piece rate). Piece-rate refers to Task-1 performance. Represent refers to the treatment condition (1 = Other-Representation Treatment and 0 = Self-Representation Control condition). Guessed 1 for Task1 equals 1 if individual ranked their performance as the best in their group in Task 1, and 0 otherwise. Guessed 2 for Task1 equals 1 if individual ranked their performance as the second best in their group in Task 1, and 0 otherwise. Demographic variables such as age, race, and household income level as well as the month the experiment was conducted are also controlled for. Column 1 presents marginal effects evaluated at means for a woman in the self-representation treatment. Column 2 presents marginal effects evaluated at means for a woman in the self-representation treatment who believed that she ranked third in Task 1 performance. Column 3 presents marginal effects evaluated at means for a man in the other-representation treatment. Column 4 presents marginal effects evaluated at means for a man in the other-representation treatment who believed that he ranked third in Task 1 performance. Ranking guesses of 4, were excluded from the analysis. *** Significance at the 1 percent level. ** Significance at the 5 percent level. * Significance at the 10 percent level.

between the Task 3 and Task 4 decisions was that individuals faced the prospect of performing in Task 3, while in Task 4, they only decided if they wanted to submit their previous performance to a tournament. Therefore, any residual effects after controlling for the Task 4 decision (i.e. Task 4 decision, "submitted the Piece Rate" equals 1 if piece-rate performance was submitted into a tournament, and 0 otherwise) can be attributed to differences in attitudes towards preforming in a competition. When the Task 4 decision is added to the Task 3 tournament-entry decision analysis (Table 6), the representation effect remains small in magnitude and statistically insignificant (Column 3) as well as the gender effect (Column 6). This suggests that attitudes towards performing in a competition does not play a significant role in explaining why women lean back from entering competitions when representing others.

Overall, the results suggest that the leaning back of female 'representatives' from competition entry as compared to self-representing women is largely attributed to these women experiencing lower levels of confidence and being less likely to enter a competition when they guessed their performance to be less than the best. Furthermore, the gender gap in competition entry amongst other-representing individuals is explained by female 'representatives' being significantly less confident in their abilities as compared to male 'representatives'.

Were women 'representatives' correct in leaning back from entering the tournament? As already discussed, male 'representatives' were particularly energized in their performance. As a result, there was in fact a gender difference in the probability of winning the Task 2 tournament in the other-representative treatment condition. Of the twenty Task-2 tournaments, men and women won the same number of tournaments in the self-representation condition, while in the other-representation treatment, men won six more tournaments than women (men won 13 and women won 7). This difference is marginally significant (one-sided Fisher's exact test, p-value = 0.098). As in the self-representation condition (Section 4), to assess the probability of winning the tournament in the other-representation condition I randomly created four-person groups from the observed performance distributions. Conditioning only on gender, the probability of winning the tournament is 32.50 percent for a man and 17.5 percent for a woman; in a sample of forty men and forty women, this difference is marginally significant in a Fisher's one-tail test (p-value = 0.098). Conditioning on performance, in a sample of those who scored 13 or greater, the probability of winning the tournament was 50 percent for a man and 29.2 percent for a woman (Fisher's one-tail test, $p = 0.11$), in contrast to 36.4 percent for women in the self-representation condition. The probability of winning the tournament increased to 52.2 percent for a man and 35.3 percent for a woman among those who scored 14 or greater (Fisher's one-tail test, p-value = 0.23). Whereas, for a woman in the self-representation condition who scored 14 or greater, she would have had a 50 percent chance of winning the tournament. Moreover, the probability of winning the tournament increased to 63.2 percent for a man and 50 percent for a woman among those who scored 15 or greater; this difference is not significant (Fisher's one-tail test, p-value = 0.36). In contrast, for a woman in the self-representation condition who scored 15 or greater, she would have had a 62 percent probability of winning the tournament.

Therefore, given the higher performance of men under the other-representation condition, for every level of performance, women in the other-representation condition faced lower odds of winning the Task

Table 6
Probit regression of tournament choice in task 3, marginal effects.

	(1) Women	(2) Women	(3) Women	(4) Representatives	(5) Representatives	(6) Representatives
Piece-Rate	0.006	0.032	0.044	−0.023	−0.009	−0.033
	(0.025)	(0.027)	(0.027)	(0.028)	(0.040)	(0.027)
Tournament	0.061**	−0.004	0.001	0.0785***	0.064	0.061
	(0.027)	(0.014)	(0.025)	(0.030)	(0.047)	(0.039)
Represent	−0.365***	−0.075	−0.155			
	(0.120)	(0.075)	(0.138)			
Female				−0.264*	−0.073	−0.055
				(0.145)	(0.163)	(0.112)
Guessed 1 as Tournament Rank		0.904***	0.818***		0.446**	0.135
		(0.084)	(0.155)		(0.220)	(0.210)
Guessed 2 as Tournament Rank		0.681***	0.690***		−0.157	−0.136
		(0.134)	(0.147)		(0.226)	(0.183)
Submitted the Piece Rate			0.120			0.429*
			(0.128)			(0.224)
Observations	79	79	79	75	75	75
PsuedoR2	0.310	0.526	0.542	0.401	0.533	0.560

The table presents Probit regression results with marginal effects reported in percentage points. The dependent variable is Task-3 choice of compensation scheme (1 = tournament and 0 = piece rate). Piece-rate refers to Task-1 performance and Tournament refers to Task-2 performance. Represent refers to the treatment condition (1 = Other-Representation Treatment and 0 = Self-Representation Control condition). Guessed 1 Tournament Rank equals 1 if individual ranked their performance as the best in their group, and 0 otherwise. Guessed 2 Tournament Rank equals 1 if individual ranked their performance as the second best in their group, and 0 otherwise. Demographic variables such as age, race, and household income level as well as the month the experiment was conducted are also controlled for. Column 1 presents marginal effects evaluated at means for a woman in the self-representation treatment. Column 2 presents marginal effects evaluated at means for a woman in the self-representation treatment who believed that she ranked third in Task 2 performance. Column 3 presents marginal effects at means for a woman who submits to the tournament in the self-representation treatment. Column 4 presents marginal effects evaluated at means for a man in the other-representation treatment. Column 5 presents marginal effects evaluated at means for a man in the other-representation treatment who believed that he ranked third in Task 2 performance. Column 6 presents marginal effects at means for a man who submits to the tournament in the other-representation treatment. Ranking guesses of 4, were excluded from the analysis. *** Significance at the 1 percent level. ** Significance at the 5 percent level. * Significance at the 10 percent level.

2 tournament than women in the self-representation condition. Women in the other-representative condition may have expected the male 'representatives' to be more energized, which could have impacted their confidence and would have led them to be less likely to participate in the tournament if they did not believe that they were the top performer in the group. This could have been motivated by stereotype threat given the stereotypically male math task involved, possibly reducing women's overall confidence, and leading them to not become as energized as men in their performance. However, given that risk and feedback aversion do not appear to have played a role, it is more likely that women's concerns for their 'constituent' was driving their behavior.

6. Conclusion

This study examined the role that being a 'representative' (an important feature in leadership positions) plays on the competition behavior of men and women in a stereotypically male domain. In the treatment condition, female and male 'representatives' were asked to decide whether they choose to enter into a tournament that involved performing in a male-stereotyped task. These 'representatives' were responsible for the payoff of someone else, their 'constituents', in addition to their own. Whereas, in the control group, individuals faced the exact same tasks and decisions, but were only responsible for their own payoffs.

With a significant majority of study participants being Harvard University students, unlike previous competition experiments, men and women in the control group exhibited very similar performance and competition behaviors. This baseline places this study in a unique position to examine how the responsibility of representing another person's interests influenced the competitive behavior of similarly ambitious and qualified men and women, such as men and women in leadership positions. Comparing the treatment and control conditions, I found that not only were women not eager to enter competitions when charged with representing others, but they *leaned back*. Controlling for performance, female 'representatives' were significantly less likely to enter tournaments than male 'representatives' and self-representing women. In addition, female 'representatives' did not significantly increase their performance, while male 'representatives' were "particularly energized" and experienced a boost in their performance levels as compared to the self-representing men.

The leaning back of female 'representatives' from competition entry as compared to self-representing women, was largely driven by female 'representatives' lower levels of confidence and differences amongst women who guessed their performance to be less than the best. When women believed that they were the top performer in both the self-representing and other-representing conditions, they are just as likely to enter competitions. However, when women believed that they performed less than the best in their group, women in the other-representing condition were significantly less likely to enter the tournament. This suggests that women's other-regarding preferences may be at play, with women being concerned about their 'constituent' ending up with earnings of zero if they are not the top performer.

Also, compared to male 'representatives', female 'representatives' were significantly less confident in their abilities, which explained the gender gap in competition entry amongst other-representing individuals in the treatment condition. These insights are in line with existing field research that has found highly qualified women to have lower confidence in their abilities with regards to their political and entrepreneurial ambitions (Fox & Lawless, 2011; Lawless, 2012; Thebaud, 2010) – both, traditionally male domains where women's performance in competitions with men could also impact others.

These findings also connect the dots between the Bowles et al. (2005) negotiation research and the Niederle and Vesterlund (2007) competition research, informing the prevalent hypothesis that "just as women have been found to opt out of negotiations, and yet be willing to negotiate on behalf of others, it may be that they are more eager to compete when doing so benefits others" (Niederle & Vesterlund, 2011). It turns out that women had no problem entering competitions when only representing their own interests, but as 'representatives', they leaned back and shied away from competing.

These insights illuminate the impact of the key leadership responsibility of representing others' interests on the competition

behavior of men and women in stereotypically male domains. Perhaps the prospect of having to engage in future competition and represent others in traditionally male-dominated fields (a key feature of most high-level leadership positions) may deter women from pursuing positions that require the representation role. If so, this could help explain the persistent gender gaps in leadership. As participants in this study did not self-select into the representation role, this is an area future research can explore in greater detail.

Additionally, given these insights, perhaps we need to rethink the design of competitions instead of placing the burden on women to "lean in" and compete more. In fact, He, Kang and Lacetera (2021) test a "nudge" intervention and show that altering the choice architecture and making competition the default option (where individuals would need to opt out if they did not want to compete), eliminated gender differences in competition entry. Similarly, Erkal et al. (2022) also found that women are more likely to participate in a performance-based leadership selection process under an opt-out mechanism than an opt-in option, reducing the gender gap in leadership selection.

To help more qualified women become competitive in leadership both in terms of performance and competition entry when representing the interests of others, interventions to redesign competition environments and their applications to outside of the lab settings are possible areas for future research. Some examples include altering choice architecture, encouraging gender-blindness (Martin & Phillips, 2017), information nudges (e.g. to reduce the role of stereotype threat, share information when the task is introduced about the fact that men and women performed the same in a similar group of participants, as was the case in the control), reducing the pressure involved in performing when representing the interests of others (e.g. increasing the time allowed to perform), changing the incentives to one that women care about more (Cassar, Wordofa & Zhang, 2016; Cassar & Zhang, 2022), and providing feedback on actual and relative performance. Future research can build on the insights of this study and inform how to support 'representatives' so they can perform their best and make decisions that would maximize gains for everyone involved.

Funding

This research was supported by the Women and Public Policy Program, Harvard Kennedy School; and the Program on Negotiation (PON) Next Generation Grant, Harvard Law School.

CRediT authorship contribution statement

Maliheh Paryavi: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declarations of Competing Interest

None.

Data availability

Data will be made available on request.

Acknowledgements

I am deeply grateful to Iris Bohnet, Max Bazerman, Ryan D. Enos, and Torben Iversen for their invaluable feedback and support throughout this research. I am also thankful to Hannah Riley Bowles, whose negotiation research inspired this study, and Muriel Niederle and Lise Vesterlund for sharing the Z-tree program of their competition experiment. I also greatly appreciate the support of the Women and Public Policy Program and the Harvard Decision Science Lab at Harvard Kennedy School.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.socec.2023.102142.

References

Alan, S., Ertac, S., Kubilay, E., & Loranth, G. (2020). Understanding gender differences in leadership. *The Economic Journal, 130*(626), 263–289.
Amanatullah, E. T., & Morris, M. W. (2010). Negotiating gender roles: Gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others. *Journal of Personality and Social Psychology, 98*(2), 256–267.
Attali, Y., Neeman, Z., & Schlosser, A. (2011). *Rise to the challenge or not give a damn: Differential performance in high vs. low stakes tests*.
Balafoutas, L., & Sutter, M. (2010). *Gender, competition and the efficiency of policy interventions*.
Baldiga, K. (2014). Gender differences in willingness to guess. *Management Science, 60*(2), 434–448.
Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics, 116*(1), 261–292.
Beckmann, D., & Menkhoff, L. (2008). Will women be women? Analyzing the gender difference among financial experts. *Kyklos : Jahrbuch des Instituts fur Geschichte der Medizin an der Universitat Leipzig, 61*(3), 364–384.
Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology, 59*(5), 960–970.
Beyer, S., & Bowden, E. M. (1997). Gender differences in self-perceptions: Convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin, 23*(2), 157–172.
Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about gender. *American Economic Review, 109*(3), 739–773.
Born, A., Ranehill, E., & Sandberg, A. (2022). Gender and willingness to lead: Does the gender composition of teams matter? *Review of Economics and Statistics, 104*(2), 259–275.
Bowles, H. R. (2012). *Psychological perspectives on gender in negotiation*.
Bowles, H. R., Babcock, L., & McGinn, K. L. (2005). Constraints and triggers: Situational mechanics of gender in negotiation. *Journal of Personality and Social Psychology, 89*, 951–965.
Buser, T., Dreber, A., & Mollerstrom, J. (2017). The impact of stress on tournament entry. *Experimental Economics, 20*(2), 506–530.
Campbell, A. (2002). *A mind of her own: The evolutionary psychology of women*. Oxford, UK: Oxford University Press.
Cason, T. N., Masters, W. A., & Sheremeta, R. M. (2010). Entry into winner-take-all and proportional-prize contests: An experimental study. *Journal of Public Economics, 94*(9), 604–611.
Cassar, A., Wordofa, F., & Zhang, Y. J. (2016). Competing for the benefit of offspring eliminates the gender gap in competitiveness. *Proceedings of the National Academy of Sciences, 113*(19), 5201–5205.
Cassar, A., & Zhang, Y. J. (2022). The competitive woman: Evolutionary insights and cross-cultural evidence into finding the femina economica. *Journal of Economic Behavior & Organization, 197*, 447–471.
Catalyst. (2020). Women on Corporate Boards. Retrieved from https://www.catalyst.org/research/women-on-corporate-boards/.
Catalyst. (2021). Historical List of Women CEOs of the Fortune Lists: 1972-2021. Retrieved from https://www.catalyst.org/research/historical-list-of-women-ceos-of-the-fortune-lists-1972-2021/.
Chen, J., & Houser, D. (2017). Gender composition, stereotype and the contribution of ideas.
Chen, J., & Houser, D. (2019). When are women willing to lead? The effect of team gender composition and gendered tasks. *The Leadership Quarterly, 30*(6), Article 101340.
Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics, 129*(4), 1625–1660.
Coffman, K. B., Collis, M., & Kulkarni, L. (2020). *When to apply?* Harvard Business School.
Coffman, K. B., Flikkema, C. B., & Shurchkov, O. (2021). Gender stereotypes in deliberation and team decisions. *Games and Economic Behavior, 129*, 329–349.
Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature, 47*(2), 448–474.
Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin, 122*(1), 5–37.
Daly, M., & Wilson, M. (1983). *Sex, evolution, and behavior* (2nd ed.). Belmont, CA: Wadsworth Publishing Company.
Dargnies, M.-P. (2011). *Social identity and competitiveness*.
Dilmaghani, M. (2020). Gender differences in performance under time constraint: Evidence from chess tournaments. *Journal of Behavioral and Experimental Economics, 89*(101505).
Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.
Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics*, 1061–1073.

Erkal, N., Gangadharan, L., & Xiao, E. (2022). Leadership selection: Can changing the default break the glass ceiling? *The Leadership Quarterly, 33*(2), Article 101563.
Ertac, S., & Szentes, B. (2010). *The effect of performance feedback on gender differences in competitiveness: Experimental evidence*.
Exley, C. L., & Kessler, J. B. (2019). *The gender gap in self-promotion*. National Bureau of Economic Research (No. w26345).
Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171–178.
Fox, R. L., & Lawless, J. L. (2011). Gendered perceptions and political candidacies: A central barrier to women's equality in electoral politics. *American Journal of Political Science, 55*(1), 59–73.
Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics, 118*(3), 1049–1074.
Grosse, N. D., & Riener, G. (2010). *Explaining gender differences in competitiveness: Gender-task stereotypes*.
He, J. C., Kang, S. K., & Lacetera, N. (2021). Opt-out choice framing attenuates gender differences in the decision to compete in the laboratory and in the field. *Proceedings of the National Academy of Sciences, 118*(42).
Healy, A., & Pate, J. (2011). Can teams help to close the gender competition gap? *The Economic Journal, 121*(555), 1192–1204.
IPU Parline. (2021). *Global and regional averages of women in national parliaments, as of September 1, 2021*. Retrieved from: https://data.ipu.org/women-averages.
Kamas, L., & Preston, A. (2010). *Social preferences, competitiveness and compensation: Are there gender differences?*.
Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology, 43*(5), 825–832.
Lawless, J. L. (2012). *Becoming a candidate: Political ambition and the decision to run for office*. Cambridge University Press.
Lundeberg, M. A., Fox, P. W., & Punćcohaŕ, J. (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal of Educational Psychology, 86*(1), 114.
Marianne, B. (2011). New perspectives on gender. In *Handbook of labor economics, 4* pp. 1543–1590). Elsevier.
Markowsky, E., & Beblo, M. (2022). When do we observe a gender gap in competition entry? A meta-analysis of the experimental literature. *Journal of Economic Behavior & Organization, 198*, 139–163.
Martin, A. E., & Phillips, K. W. (2017). What "blindness" to gender differences helps women see and do: Implications for confidence, agency, and action in male-dominated environments. *Organizational Behavior and Human Decision Processes, 142*, 28–44.
Niederle, M. (2017). A Gender Agenda: A Progress Report on Competitiveness. *American Economic Review, 107*(5), 115–119.
Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, (122), 1067–1101.
Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives, 24*(2), 129–144.
Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics, 3*(1), 601–630.
Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = Male, Me = Female, Therefore Math (not equal to) Me. *Journal of Personality and Social Psychology, 83*(1), 44–59.
Paserman, M.D. (2007). *Gender differences in performance in competitive environments: Evidence from professional tennis players*.
Price, C. R. (2010). *Do women shy away from competition? Do men compete too much? : A (Failed) replication*.
Price, C. R. (2012). Gender, competition and managerial decisions. *Management Science, 58*(1), 114–122.
Price, J. (2008). Gender differences in the response to competition. *Industrial and Labor Relations Review, 61*(3), 320–333.
Pruitt, D. G., Carnevale, P. J. D., Forcey, B., & Van Slyck, M. (1986). Gender effects in negotiation: Constituent surveillance and contentious behavior. *Journal of Experimental Social Psychology, 22*(264–275).
Pulford, B. D., & Colman, A. M. (1997). Overconfidence: Feedback and item difficulty effects. *Personality and Individual Differences, 23*, 125–133.
Roberts, T.-A., & Nolen-Hoeksema, S. (1989). Sex differences in reactions to evaluative feedback. *Sex Roles, 21*(11/12).
Ruble, D. N., Martin, C. L., & Berenbaum, S. A (2006). Gender development. In W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology* (6th ed., pp. 858–932). New York: Wiley. Vol. 3.
Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology, 74*(3), 629–645.
Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: Hidden costs to women of a kinder, gentler image of middle managers. *Journal of personality and social psychology, 77*(5), 1004–1010.
Rutgers University. (2021). History of women in the U.S. congress. Retrieved from http://www.cawp.rutgers.edu/history-women-us-congress.
Saccardo, S., Pietrasz, A., & Gneezy, U. (2018). On the size of the gender difference in competitiveness. *Management Science, 64*(4), 1541–1554.
Sandberg, S. (2013). *Lean in: Women, work, and the will to lead*. New York: Alfred A. Knopf.
Shurchkov, O. (2012). Under pressure: Gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association, 10*(5), 1189–1213.
Soll, J. B., & Klayman, J. (2004). Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning Memory, and Cognition, 30*(2), 299–314.
Spencer, S. J., & Steele, C. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4–28.
Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811.
Sutter, M., & Rützler, D. (2010). *Gender differences in competition emerge early in life*.
Thebaud, S. (2010). Gender and entrepreneurship as a career choice: Do self-assessments of ability matter? *Social Psychology Quarterly, 73*(3), 288–304.
Wozniak, D., Harbaugh, W. T., & Mayr, U. (2009). *Choices about competition: Differences by gender and hormonal fluctuations, and the role of relative performance feedback*.