Women dislike competing against men*

Diogo Geraldes†

March 2016

[CLICK HERE FOR THE LATEST VERSION]

Abstract

A prominent finding in the literature on gender competition is that women are less inclined to compete in comparison to men. In this paper, we conduct a laboratory experiment to examine the importance of the belief in the sex of potential competitors on men's and women's decision to enter into competition. Specifically, we test whether women have a weaker preference to compete per se, or rather just shy away from competing against men. The results support the latter hypothesis. When given the possibility to choose a competitor's sex, or when being in the lab surrounded only by female participants, the percentage of women entering into competition is high and similar to the figures commonly reported for men. Moreover, only women are sensitive to the different cues we provide concerning the sex of potential competitors, and their competitiveness is largely driven by their beliefs in other women's competitive attitude. Our findings have important policy implications for the labor markets and educational programs in which women are under-represented. They suggest that persuasive references to recent female applicants and/or hiring of female staff while advertising a position could be more effective to promote women's participation than on-going interventions highlighting women's under-representation.

Keywords: male-typed, perception of gender composition, potential competitors, winning beliefs
JEL classification: C91, C92, D03, D81, D84, J16

---

*The author thanks for invaluable comments the participants of the Yale SOM Whitebox Conference, ESA European Meeting, Political Economy Seminar (Columbia University), The Choice Lab Seminar (NHH - Norwegian School of Economics), M-BEES and IMEBE congress. The author is particularly indebted for the comments of Jordi Brandts, Alessandra Casella, Matthew Embrey, Zsombor Méder, Joshua Miller, Ronald Peeters, Ernesto Reuben, Arno Riedl, Martin Strobel and Bertil Tungodden. The author is very grateful to Lise Vesterlund for providing her z-Tree program for the work task.
†Maastricht University, Department of Economics – Section AE1. Address: Maastricht University, Department of Economics – Section AE1, P.O. Box 616, Maastricht 6200 MD, The Netherlands. E-mail: diogo.geraldes@gmail.com

1. Introduction

"Let me tell you about my trouble with girls…Three things happen when they are in the lab: You fall in love with them, they fall in love with you, and when you criticize them they cry".[1] These recent controversial remarks of the Nobel Laureate Tim Hunt anecdotally illustrate a pervasive phenomenon at work; men often employ gender stereotypes to patronize women. The growing concern with gender stereotypes at the workplace is an undeniable fact, as attested by the massive attention paid to the recent sexism case of Ellen Pao in Silicon Valley. While the demand side of the market has already started to address gender problems at the workplace (e.g., McKinsey's gender diversity program), and some research has been investigating its effects from a demand side perspective (e.g., Reuben, Sapienza and Zingales [2014]), very little is known about how gender stereotypes affect the supply side. In this paper, we study whether (and how) women's belief in the sex of potential competitors influences their decision to enter into a competitive male-typed domain.[2]

The seminal work of Niederle and Vesterlund (2007) has initiated a flourishing literature investigating men's and women's self-selection into competitive environments (see Niederle and Vesterlund [2011] for a review of the literature). In this literature the prevalent procedure to measure competitive preference is to ask participants to choose between a non-competitive and a competitive payment scheme for a subsequent work task performance. The main finding is that men are significantly more inclined to choose the competitive payment. Accordingly, this literature suggests that a possible reason for the well-documented gender gap in wages and positions at the workplace as well as in educational choices is that men have a stronger preference to compete compared to women.[3] An open key question, however, is whether the accumulating evidence of women's lower inclination to compete is due to an aversion to competition per se, or rather just due to the fact that women dislike competing against men. In this paper, we attempt to fill this gap by examining women's belief in the sex of potential competitors, a presumably important element influencing their entry into competition that research to date has tended to neglect.

A common procedure in many studies that report women's lower inclination to choose the competitive option is to conduct mixed-sex laboratory sessions without making reference to the gender composition of participants (e.g., Niederle and Vesterlund [2007]; Gneezy, Leonard and List [2009]; Cason, Masters and Sheremeta [2010]; Dohmen and Falk [2011]; Almås, Cappelen, Salvanes, Sorensen and Tungodden [2012]; Wozniak, Harbaugh and Mayr [2014]). Hence, if female participants perceive the gender composition in the lab and, accordingly, correctly anticipate potential male competitors, then the claim that women dislike competition per se might be compromised. If women dislike competing against men, the stylized fact that women shy away from competition might be a result of the mixed-sex composition in the lab rather than an intrinsic weaker preference for

[1]Hunt, T. (2015). Remarks at the World Conference of Science Journalists.

[2]Throughout this paper, male-typed refers to domains/situations/contexts/environments/areas/fields in which men are stereotypically believed to be more skilled and/or overrepresented.

[3]For instance, although female Fortune 500 CEOs reached recently a historic high of 4.8% (June 2014), the figure is still dismal for women (source: fortune.com). For further evidence on the gender gap at the workplace see, e.g., Blau, Ferber and Winkler (2013). For evidence on gender differences in educational choice and its correlation with competitiveness see Buser, Niederle and Oosterbeek (2014).

competition. Thus, for drawing accurate conclusions, elucidating the relationship between women's willingness to compete and their belief in the sex of potential competitors is of utmost importance, particularly for policy makers concerned with women's willingness to enter into competitive male-typed domains.

Our study has two primary purposes. First, we conduct a controlled laboratory experiment to investigate whether women have a weaker preference to compete per se, or rather just shy away from competition against men. To this end, we measure men's and women's self-selection into competition across conditions entailing different cues for the gender composition of potential competitors. Specifically, we ask a participant to choose between a non-competitive and a competitive payment scheme for a subsequent performance of a math work task in three conditions: i. *Replication condition*: the gender composition of participants present in the lab is symmetric; ii. *Choice of sex condition*: the gender composition of participants present in the lab is symmetric, but a participant can choose the sex of the competitor while deciding to compete; iii. *All women condition*: only female participants are present in the lab. Our conjecture is that women's belief in the sex of potential competitors influences their decision to enter into competition. In particular, we hypothesize that women are not less competitive than men but, at least in male-typed domains, they dislike facing a male competitor. A second purpose of this paper is to assess whether women's willingness to compete across the three conditions could enlighten us to tailor policy intervention to encourage women to compete.

A central element of our experiment is that we do not inform participants about the gender composition in the lab before eliciting their choice of payment scheme, but rather rely on participants' perception of the gender composition, which we elicit at a later moment in the experiment. With this procedure we aim not only to minimize the chance of participants realizing the goal of the experiment but also to capture the spirit of relevant decisions outside the lab as, for instance, deciding whether to apply for a job. Moreover, the focus of this paper is on a male-typed domain because the most representative labor markets and educational programs in which the gender gap is a serious concern (i.e., high-level business positions, STEM fields) are stereotypically believed as male-typed domains.[4] To create a male-typed environment in the lab we choose a math work task because we expect female participants to be aware of the well-known stereotype that men are better at mathematics, and there is empirical and theoretical insight to expect women to dislike competing against men in this situation.

Psychology research shows evidence that stereotype threat triggers women to feel apprehensive by the possibility of being judged by the negative stereotype about their group in a (non-competitive) mixed-sex math context (e.g., Spencer et al. [1999]). Stereotype threat is a situational predicament where an individual is at risk of confirming a negative stereotype about his/her group. The effect mostly associated to stereotype threat is the impairment of an individual's performance in situations where poor performance confirms the stereotype (see Inzlicht and Schmader [2013] for a review of the literature). Therefore, an explanation based on stereotype threat is consistent with the influential results of Gneezy, Niederle and Rustichini (2003). While investigating

[4]We could argue that it is interesting per se to investigate whether the stylized result that women shy away from competition also holds in a female-typed domain. In fact, as the seminal work of Niederle and Vesterlund (2007) as well as many of the follow-up studies that replicate their finding use a male-typed work task, some studies have already used a female-typed work task instead. These studies have shown that the gender gap in competition entry is reduced (e.g., Kamas and Preston [2012]), which indicates that gender-task stereotypes play a role in women's self-selection into competition.

exogenously set competition within a male-typed domain, these authors show that men and women perform equally well under competitive incentives in single-sex groups, but women perform worse when performing in mixed-sex groups.

Hence, a conceivable side effect of repetitive exposure to a math stereotype threat is that women anticipate a poor math performance when on the verge of a mixed-sex competition. Accordingly, a reason for expecting women to shy away from competition against men in a math context could "simply" be a decreased confidence level. The model on self-stereotyping of Bordalo, Gennaioli and Shleifer (2014) formally proposes a confidence mechanism to predict women's lower inclination of women to enter a math competition. When a woman compares herself to men in a situation involving a math competition, she underestimates her probability to win because the stereotype that men are better at mathematics amplifies believed differences in ability between the sexes. The model prediction is therefore that women enter less into competition than they would do in case their confidence level was not biased by negative self-stereotyping.

We consider an alternative, but not necessarily competing, reason for expecting women to shy away from competition against men in a math context. If women are prone to a math stereotype threat and they are chronically exposed to this threat, we conjecture that a logical reaction of women is to develop a preference to avoid situations where their math ability can be compared to men's. Accordingly, women should be particularly susceptible to dislike situations involving a math competition against men because the mixed-sex competition should heighten women's awareness of gender, thus amplifying the situational predicament of stereotype threat.

The closest previous work to our study is scant. Gneezy et al. (2009) notably acknowledge the possible importance of female participants deducing the gender distribution of potential competitors[5], and Niederle et al. (2013) recognize quotas as making the competition more sex-specific. To our knowledge, however, only a few studies have attempted to manipulate the gender composition while investigating entry into competition. Using mixed- and single-sex groups, Booth and Nolen (2012) study boys and girls from either coeducational or single-sex schools. Their results suggest that girls are more willing to compete against other girls, but the figure is not broken down by the type of education. Gupta, Poulsen and Villeval (2013) conduct balanced mixed-sex sessions in which they randomly match participants in mixed- and single-sex pairs. Importantly, in each pair, a participant is automatically the winner in case a participant is the only one to choose competition. The authors find that women's choice of the competitive payment does not depend on the sex of their pair. However, since the decision to choose competition also depends on outguessing, we cannot conclude that this result is driven by women's competitive preference. Hence, despite these two attempts[6], the connection between women's decision to enter into competition and the sex of potential competitors remains unclear. Moreover, a key difference

[5]Using an elegant experiment, Gneezy et al. (2009) find that women of a matrilineal society in India (Khasi tribe) and men of a patriarchal society in Tanzania (Maasai tribe) similarly enter into competition. However, as the majority of Khasi participants in their experiment are women the authors admit the possibility that Khasi female participants could have deduced the gender distribution of potential competitors. Accordingly, the authors acknowledge that their interpretation for the result [nurture] might be compromised if women are more likely to compete against women regardless of their society.

[6]Sutter and Rützler (2010), in an interesting study about how men's and women's competitive preference evolve early in life, also include mixed- and single-sex groups in their design to investigate boys and girls. However, the figure they report of girls' entry into competition is not broken down by single- and mixed-sex groups.

between these studies and our experiment is that they inform participants about the gender composition in the lab before eliciting participants' choice of payment scheme.

Our paper also speaks to recent research that has started investigating institutional mechanisms to promote women to compete more. Motivated by the on-going debate about affirmative action policy, Balafoutas and Sutter (2012) and Niederle, Segal and Vesterlund (2013) find that instituting the competitive option more favorable to women in the form of quotas significantly increases women's entry into competition. As a result, the gender gap in competition entry decreases, or even reverses. In addition, Balafoutas and Sutter (2012) show that preferential treatment for women in the form of a head start has a similar effect. However, they also find that preferential treatment in the form of a repetition of the competition in case a man wins – an institution that from our perspective does not turn the competition more sex-specific – does not significantly alter women's entry into competition.

Our results provide evidence that women are not less competitive than men, but rather dislike competing against men. In line with the literature, we find that women's willingness to compete (38%) is significantly lower than the one of men (69%) when male and female participants are present in the lab and there is no possibility of choosing the sex of a competitor. In stark contrast, we find that women's percentage of competition entry is high (67%) and similar to the figure for men (76%) when male and female participants are present in the lab, but they can choose the sex of the competitor while deciding to compete. Likewise, women's entry into competition is high (71%) when there is no possibility of choosing the sex of a competitor but only female participants are present in the lab. Since the competitive payment scheme we offer in each condition is exactly the same, collectively these results strongly indicate that the significant gender gap in the replication condition is not due to women shying away from competition per se, but rather women shying away from competition against men. Corroborating this interpretation, we observe that women who compete in the choice of sex condition mostly choose a female competitor (77%).

We find little support for negative self-stereotyping in our setting. Although we show that women hold the stereotype that men are better at mathematics, we find that differences in confidence between men and women do not explain much of the gender gap in the replication condition. Moreover, women's preference for a female competitor in the choice of sex condition is not explained by a higher confidence to compete against women than against men. Notably, rather than observing significant confidence differences between men and women in the replication condition, we observe instead a significantly lower entry into competition among low-confident women compared to low-confident men. We reach the same conclusion when comparing low-confident women in the replication condition to low-confident women in the other two conditions. High-confident women, on the other hand, similarly enter into competition across the three conditions. These results indicate that women's dislike of competition against men primarily affects the entry into competition of low-confident women.

The interpretation we give to women's behavior across conditions is substantiated by their belief in the sex of potential competitors. To directly test our conjecture that women's belief in the sex of potential competitors influences their willingness to compete, we elicit two beliefs. First we elicit if (and how) women perceive the gender composition of participants present in the lab. Second, we elicit women's belief in the likelihood of each

sex to enter into competition. We show that in each condition almost all women correctly perceive the actual gender composition of participants present in the lab. Moreover, in support of our hypothesis, we find in the replication condition that women are significantly less likely to enter into competition the more male participants women believe are among the participants who choose competition.

Less intuitive is our finding that women's dislike of competition against men is not driven by women's belief in men's likelihood to enter into competition, but rather by women's belief in other women's likelihood to enter into competition. In every condition we find a significant inclination of women to enter into competition the more they believe other women compete, but women's average belief in other women's likelihood to enter into competition is lower in the replication condition. In fact, we show that this difference greatly explains the lower entry into competition of women in the replication condition compared to women in the other two conditions.

Our experiment allows us to distil distinctive policy interventions aiming to encourage women to enter into competitive male-typed domains. First, our findings suggest that policy interventions that induce an increase in women's confidence can be effective, but not as plainly as we could expect. The evidence we present shows that it is not that women are less confident than men while deciding to enter a mixed-sex competition; it is just that low-confident women are more reluctant to compete compared to low-confident men. This implies that policy interventions should particularly target low-confident women. Second, our findings on women's belief in the sex of potential competitors offer a new view on how to induce women to compete. In the light of our results, we prescribe that highlighting female role models can increase women's willingness to enter into competitive male-typed domains. Namely, we suggest that information highlighting women's underrepresentation that is commonly observed on job advertisements in a representative male-typed domain – the economics academic career[7] – should be replaced by references to recent female applicants and/or hiring of female staff.

The remainder of the paper is organized as follows. In section 2 we describe the design of the experiment. In section 3 we present and discuss results on work task performance, choice of payment scheme and risk attitude. In section 4 we analyze participant's confidence level. In section 5 we assess participants' belief in the sex of potential competitors. In section 6 we discuss the policy implications of the results. In Section 7 we conclude.

2. Design of the experiment

To investigate men's and women's self-selection into competition we conduct a series of experiments with three conditions: the *replication* condition, the *choice of sex* condition and the *all women* condition.

2.1. Experimental procedure

The experiment is computerized using z-Tree software (Fischbacher [2007]) and conducted in the Behavioral and Experimental Economics Laboratory (BEElab) at Maastricht University. All instructions are presented on-

[7]Economics is the least friendly science toward women among social sciences, and the gender gap gets larger at each upper stage of the academic ladder. In 2013, only 28% of assistant professors, 25% of associate professors with tenure, and 12% of full professors were women (source: 2013 annual report of the Committee on the Status of Women in the Economics Profession, a committee of the American Economic Association).

screen. We recruit the participants via ORSEE software (Greiner [2004]) from the BEElab subject pool. The participants are 204 in total (58, 90 and 56 participants in the replication condition, the choice of sex condition and the all women condition, respectively), have predominantly a business and economics background[8] and, by and large, do not know each other at the session level.[9] The average age is 22. Each session lasts on average 60 minutes. Average earnings are € 12.65. Participants' decisions and interactions stay anonymous. Moreover, we do not inform participants about the gender composition in the lab before eliciting their preference to compete in any of the conditions. Importantly for our purposes, however, before being invited to take a seat in a private computer cubicle, participants spend a few minutes in a welcome room where they can see each other.

2.2. Experimental conditions

2.2.1 Replication condition: an equal number of men and women participate in each session of this condition.[10] We subsequently describe the different steps of this condition in sequential order.

- ***Practice round***: The experiment starts with a practice round to familiarize a participant with the work task. The work task consists of solving consecutive five two-digit numbers addition problems (Niederle and Vesterlund [2007]; e.g., 84 + 94 + 17 + 19 + 39). This work task requires real effort, and previous results indicate no significant gender difference in productivity (e.g., Niederle and Vesterlund [2007]). In this round a participant practices the work task for 2 minutes and receives no payment.[11]

- ***Round 1: Baseline Performance***: A participant performs the work task during 5 minutes under a piece-rate payment scheme. Specifically, a participant earns € 0.50 per correct answer. This step serves to elicit a participant's work task productivity.

- ***Choice of payment scheme for a subsequent work task performance***: A participant chooses between a non-competitive and a competitive payment scheme for a subsequent work task performance. If a participant chooses the non-competitive alternative, a participant is paired with a randomly chosen participant among the participants who have also chosen the non-competitive alternative. At the end of the experiment, one of the two is chosen with a 50% probability for actual payment of the subsequent work task performance.[12] The chosen participant earns € 1 per correct answer. The other participant earns nothing regardless of his/her number of correct answers. If a participant chooses the competitive alternative, a participant subsequently competes under a winner-take-all tournament against a randomly chosen participant among the participants who have also chosen

[8]78% of the participants have a business and economics background. The remaining participants have a background in law, health sciences, psychology, neuroscience, arts and culture, European studies or computer science.
[9]Using the debriefing question: "How many of the participants in this experimental session do you consider as a person you know?", we find that participants know, on average, 0.81, 0.71 and 0.64 persons in the replication, choice of sex and all women conditions, respectively.
[10]We conduct two sessions in this condition. The procedure to ensure an equal number of male and female participants per session is described in the Online Appendix.
[11]During the practice round as well as during subsequent performance rounds, the following holds while a participant performs the work task: i. After providing an answer, a participant is immediately informed whether the answer is correct and a new problem is instantaneously displayed on the screen; ii. A participant's total number of problems correctly solved is permanently displayed on-screen; iii. A participant does not receive feedback on other participants' performance.
[12]In case an odd number of participants choose the non-competitive payment, the unmatched participant is also paid with a 50% probability, but this participant being paid does not imply that another participant is not paid.

to compete.[13] In this pairwise competition, the participant who correctly solves more problems earns € 1 per correct answer and the other earns zero. In case of a tie, one of the two is chosen as the winner with a 50% probability.

- ***Winning belief***: Elicitation of a participant's belief in winning the subsequent competition (details of this elicitation provided in section 4).
- ***Round 2: Performance under the chosen payment scheme***: A participant performs the work task during 5 minutes under the chosen compensation.
- ***Perception of the gender composition***: Elicitation of a participant's perception of the gender composition of participants present in the lab (details of this elicitation provided in section 5).
- ***Belief in the inclination of each sex to enter into competition***: Elicitation of a participant's belief in the likelihood of each sex to enter into competition (details of this elicitation provided in section 5).
- ***Stereotype-based beliefs***: Elicitation of a participant's belief in men's and women's performance in round 1, in which all participants perform under a piece-rate payment scheme. (details of this elicitation provided in section 4).
- ***Risk attitude***: A participant answers to the question: "How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, in which the value 0 means: 'not at all willing to take risks' and the value 10 means: 'very willing to take risks' ".[14]
- ***Questionnaires***: A participant answers demographics and debriefing questions.

We inform participants beforehand that one of the two monetarily incentivized performances is randomly chosen with equal probability at the end of the experiment for actual payment. A participant's total payment in the experiment equals the sum of the earnings for the randomly chosen performance, the earnings associated to the answers to some questions (we indicate in later sections the specific questions that are monetarily incentivized) plus a € 3 show up fee. A participant only learns his/her earnings after the questionnaires.

Our measure of competitiveness builds on the existing literature. However, the non-competitive alternative we offer is different. The standard in the literature is to offer a plain piece-rate as a non-competitive payment alternative. A possible reason to observe choice differences between a plain piece-rate and a competitive payment scheme is differences in risk attitude because the latter incentive makes payment uncertain. Therefore, to mitigate the influence of risk attitude on the choice of payment scheme, we offer instead a non-competitive payment alternative that also makes payment uncertain, although not dependent on the performance of other participants.

[13]In case an odd number of participants choose competition, a randomly chosen performance among the matched participants who have chosen competition is used a second time to determine the earnings of the unmatched participant. If only one participant chooses competition (something that did not happen), the competition is not implemented, and the participant is paid according to the non-competitive payment.

[14]In a field experiment Dohmen, Falk, Huffman, Sunde, Schupp and Wagner (2011) show that responses to this general risk question reliably predict incentivized lottery choices.

Another aspect of the design that we should highlight is the implementation of a real-time competition.[15] This competition setting allows us to inform a participant that choosing the competitive alternative means to subsequently compete against another participant who has also chosen to compete. As we argue in later sections, controlling for this belief is essential to elicit the proper confidence level of a participant as well as the belief of a participant in the sex of potential competitors.

2.2.2. Choice of sex condition: an equal number of men and women participate in each session of this condition.[16] The only important difference compared to the replication condition is when a participant chooses a payment scheme for a subsequent work task performance. In this step, a participant is rather asked to choose between three payment alternatives for a subsequent work task performance. The first alternative is the same non-competitive payment scheme offered in the replication condition. The other two alternatives offer a competitive payment scheme. A participant who chooses the second alternative subsequently competes under a winner-take-all tournament against a randomly chosen *man* among the men who have also chosen to compete. A participant who chooses the third alternative subsequently competes under a winner-take-all tournament against a randomly chosen *woman* among the women who have also chosen to compete.[17] The only difference between the two competitive alternatives is the sex of the competitor. We emphasize that the competitive payment scheme is exactly the same; it is a winner-take-all tournament.

2.2.3. All women condition: The same design as in the replication condition. The distinctive element is that only women participate in this condition.[18]

3. Basic experimental results

Throughout the remainder of the paper we report two-sided tests, except if: i. Subjacent to the test is our hypothesis that women dislike competing against men; ii. Testing whether participants hold the stereotype that men are better at mathematics; iii. Testing the self-stereotyping theory. We explicitly indicate when a test is one-sided.

3.1. Piece-rate baseline performance

We start by verifying that there is no significant gender difference in productivity for the work task. In Table 1, we document participants' piece-rate baseline performance. We observe that men's and women's performance

[15]The procedure extensively used in the literature to determine the earnings of a participant who chooses competition is to compare a participant's subsequent competitive performance to a past exogenously elicited competitive performance of other participants in the experiment. Yet, we are not the first to implement real-time competition. See, e.g., Dohmen and Falk (2011).

[16]We conduct three sessions in this condition. The procedure to ensure an equal number of male and female participants per session is described in the Online Appendix.

[17]Participants who choose competition are matched to participants with equivalent choices (e.g., a man who chooses a female competitor is matched to a woman who chooses a male competitor). For unmatched participants we use a similar procedure to the one described in footnote 13 (see the Online Appendix for the detailed procedure in the choice of sex condition). For the cases in which the competition cannot be implemented (something that did not happen), the participant is paid according to the non-competitive payment.

[18]We conduct two sessions in this condition.

Table 1. Piece-rate baseline performance (in average number of correct answers)
Note: p-values derived from Mann-Whitney tests of equality of distributions between men and women. Standard deviation in parentheses.

	Replication	Choice of Sex	All Women
Men	10.31 (4.05)	10.29 (5.07)	n.a.
Women	10.28 (3.50)	10.38 (3.76)	10.45 (4.78)
p-value	0.919	0.418	n.a.

is similar in the replication and the choice of sex conditions, respectively. Moreover, the same conclusion is reached when comparing the distribution of women's performance in the all women condition to the distribution of men's performance in the other two conditions (Mann-Whitney tests yield a p-value ≥ 0.542 for both comparisons). Therefore, the baseline performance data do not corroborate the stereotype about women's lower ability in mathematics.[19]

3.2. Choice of payment scheme for a subsequent work task performance

In Table 2, we summarize men's and women's self-selection into competition. In the replication condition

Table 2. Choice of a winner-take-all tournament (in percentage)
Note: p-values derived from Fisher's exact tests of contingency between men and women

	Replication	Choice of Sex	All Women
Men	69%	76%	n.a.
Women	38%	67%	71%
p-value	0.017[a]	0.486	n.a.

[a]one-sided test

69% of men choose to compete whereas only 38% of women do so. This observed gender gap in competition entry is substantial and in line with the literature. In regression (1) of Table 3 we show that this result is robust when controlling for individual baseline performance and risk score. The regression shows that women have a significant 32 percentage point lower probability than men to choose the competitive payment scheme.

In the choice of sex condition 76% of men and 67% of women select the competition. This gender gap in competition entry is insignificant, and robust when controlling for individual productivity and risk attitude (see

[19]Round 2 performance data do not corroborate the stereotype either. The average number of correct answers in that round is 11.45 and 11.41 for men and women in the replication condition, respectively; 11.51 and 11.58 for men and women in the choice of sex condition, respectively; 11.80 for women in the all women condition.

Table 3. Probit models of payment choice

Note: The table reports marginal effects. Robust standard errors in brackets. *** significant at 1%.

Dependent variable: *1 if payment choice is a winner-take-all tournament*			
	Replication	**Choice of Sex**	**All Women**
	(1)	(2)	(3)
1 if female	-0.32***	-0.11	n.a.
	[0.123]	[0.089]	
Baseline performance	0.05***	0.04***	0.04***
	[0.016]	[0.010]	[0.010]
Risk score	-0.007	0.016	0.022
	[0.028]	[0.017]	[0.025]
Observations	58	90	56
Pseudo R^2	0.164	0.120	0.133

regression (2) of Table 3). Recall that the two competitive alternatives in this condition involve a winner-take-all tournament, which is exactly the same payment incentive that we offer as a competitive alternative in the replication condition. Hence, if women had a weaker preference to compete per se, we would have also observed a low percentage of competition entry among the women in the choice of sex condition. Instead we observe a striking 76% increase in the percentage of women who choose competition in the choice of sex condition compared to the replication condition. In point of fact, the narrowing of the gender gap in entering into competition in the choice of sex condition is due to a significant increase of women who choose to compete (38% vs. 67%, Fisher's exact test, one-sided, p = 0.014) since the competition entry of men is only insignificantly higher in the choice of sex condition (69% vs. 76%, Fisher's exact test, p = 0.597).

As for the choice of the competitor's sex, we observe in Table 4 that both men and women reveal preference

Table 4. Choice of the competitor's sex (in percentage)

Note: p-values derived from binomial tests under the null hypothesis that the choices of male and female competitor are equally likely to occur.

	Male competitor	**Female competitor**	**p-value**
Men	35%	65%	0.121
Women	23%	77%	0.003[a]

[a]one-sided test

for a female competitor, but this inclination is not significant for men. In addition, regression analysis shows that both men's and women's choice of the competitor's sex is not driven by their performance and risk attitude (see Online Appendix).

These results are consistent with our conjecture that the sex of potential competitors influences women's decision to enter into competition. Since the number of male and female participants present in the lab is

perfectly balanced in both conditions, a possible interpretation of these results is that the mere circumstance that women cannot avoid the possibility of competing against men in the replication condition is deterring women from entering into a competitive environment. Recall that in neither condition participants are informed about the gender composition of participants present in the lab before choosing a payment scheme. Still, a participant could perceive the gender composition because, before taking a seat in a private computer cubicle, a participant can see the other participants in the welcome room.

To further investigate women's behavior we conduct the all women condition, in which the setting is exactly the same as in the replication condition but only female participants are present in the lab. In this condition we find that 71% of women enter into competition. This represents a striking 87% increase in the percentage of women who compete compared to the replication condition (38% vs. 71%, Fisher's exact test, one-sided, $p = 0.003$). Compared to the choice of sex condition, the increase in the percentage of competing women is not significant (67% vs. 71%, Fisher's exact test, $p = 0.667$). Hence, women's high percentage of competition entry in the all women condition is also consistent with our conjecture.

Finally, a remark on risk attitude is in order. Although we design the experiment to measure competitiveness in a setting in which the impact of risk attitude should be negligible, we nonetheless measure participants' risk preference. In regressions (1) – (3) of Table 3, we see that the risk score predictor is statistically insignificant in each condition, which indicates that participants' choice of payment scheme is not determined by their risk attitude.[20]

4. Stereotypes, confidence level and the "irrational" competitors

In this section we investigate whether there is evidence to establish a connection between the stereotype that men are better at mathematics and the choice for competition. Specifically, we investigate whether we can discriminate between two possible explanations grounded in stereotype threat; confidence to compete vs. preference to compete. A common element to both explanations is that they assume that individuals of the targeted group hold the stereotype that men are better at mathematics. Therefore, we start by verifying that participants do indeed hold the stereotype.

4.1. Stereotype-based beliefs

To verify that participants hold the stereotype we elicit participants' estimation of the performance of male and female participants in the 5 minutes round 1 piece rate performance. Specifically, a participant answers two questions: i. "How many addition problems did male participants correctly solve, on average, in the 5 minutes round 1 piece-rate performance (excluding yourself if you are a man)?"; ii. The same question with reference to female participants. The questions are monetarily incentivized using a scoring rule that yields a higher reward the more accurate the estimate is (see Online Appendix for the full instructions and details of the monetary

[20]The risk score is also insignificant when we regress specification (1) and (2) of Table 3 separately for men and women.

Table 5. Average estimate of the gender gap in performance (in number of correct answers)
Note: *** significant at 1% derived from one-sided Wilcoxon signed-rank tests. Standard deviation in parentheses.

	Replication	Choice of Sex	All Women
Men	1.1*** (2.70)	0.8*** (2.37)	n.a.
Women	2.1*** (2.52)	1.7*** (2.22)	1.6*** (3.06)

incentives).[21]

In the first row of Table 5 we present men's average estimate of the gender gap in performance by reporting the average difference between the men's estimates of the performance for male and female participants. In the second row we present the corresponding figures for women. In each condition participants hold the stereotype. That is, they do believe that men have, on average, a significant advantage to perform the addition task compared to women. Most importantly, the participants of the targeted group – women – hold the stereotype.[22]

Hence, women's competition entry behavior across conditions accommodates a preference-based explanation grounded in stereotype threat. As we discussed in the introduction, if women are chronically exposed to a math stereotype threat relatively to men, a logical reaction of women is to develop a preference to avoid situations where their math ability can be compared to men. Moreover, women's dislike to be compared to men is amplified in our setting because we add a competitive dimension. In the light of this explanation, women's entry into competition should be low in the replication condition and they should prefer a female competitor in the choice of sex condition. In the all women condition there is no stereotype threat (assuming women correctly perceive that men do not participate). As a result, women's entry into competition should be high.

However, as the stereotype is about women's lower ability in mathematics, an alternative mechanism could "simply" be that women's exposure to a situation evoking the stereotype triggers negative self-stereotyping, which undermines women's confidence to compete against men. Accordingly, women's belief in winning the competition should be considerably lower in the replication condition because women cannot avoid the possibility of a mixed-sex competition in case they choose competition. In contrast, women's belief in winning the competition is not biased downwards in the choice of sex condition because negative self-stereotyping bias women's confidence against men but not against other women. In the all women condition, negative self-stereotyping does not come to mind at all (assuming women correctly perceive the absence of male participants). Therefore, according to the self-stereotyping model, the lower percentage of competition entry that we observe

[21]In the all women condition we ask participants the same two questions, but with reference to male and female participants of a previous session of the experiment so that we avoid deception.
[22]The "magnitude" of holding the stereotype is not significantly different between men and women in the replication condition (1.1 vs. 2.1, Mann-Whitney test, $p = 0.113$), and is marginally significantly stronger for women in the choice of sex condition (0.8 vs. 1.7, Mann-Whitney test, $p = 0.054$).

among women in the replication condition is just a rational response of women as payoff maximizers to their belief in a lower probability of winning the competition.

To test self-stereotyping, however, the stereotype-based beliefs that we reported are not appropriate. The stereotype-based beliefs surely indicate how a female participant compares the work task performance of other female participants relatively to male participants. However, these beliefs do not necessarily indicate how a female participant believes in her *own* work task performance relatively to the *other participants who choose to compete*, which is in our view the proper measure of a participant's confidence level in the moment a participant is asked to choose a payment scheme.

4.2. Winning belief: the normative view

Before reporting the confidence level results, we first motivate the beliefs that we elicit in the experiment to sort out the impact of possible differences in confidence level on the choice of payment scheme.

Consider an arbitrary participant *i* of the experiment who maximizes expected payoff. Participant *i*'s preferences over monetary outcomes are represented by a twice-differentiable utility function $u_i(n_i)$ in which $n_i \geq 0$ represents the number of correct answers that worth € 1 each, $u_i(0) = 0$ and $u_i'(n_i) > 0$ for $n_i > 0$. In addition, participant *i*'s cost to correctly solve addition problems is represented by a strictly convex function $c_i(n_i)$ in which $c_i(0) = 0$, $c_i'(0) = 0$, $c_i'(n_i) > 0$ and $c_i''(n_i) > 0$ for $n_i > 0$. Hence, if participant *i* chooses the non-competitive payment scheme for the subsequent performance, participant *i* faces the following problem:

$$\max_{n_i} \ \pi_i^{nc}(n_i) \quad with \quad \pi_i^{nc}(n_i) = 0.5 \cdot u_i(n_i) - c_i(n_i) \qquad (1)$$

in which 0.5 represents the probability of a participant being paid for the subsequent performance. Alternatively, if participant *i* chooses a winner-take-all tournament payment scheme for the subsequent performance, participant *i* faces the following problem:

$$\max_{n_i} \ \pi_i^{com}(n_i) \quad with \quad \pi_i^{com}(n_i) = bwin_i \cdot u_i(n_i) - c_i(n_i) \qquad (2)$$

in which $bwin_i$ represents the belief in the probability of winning the competition and, thus, the belief in the probability of participant *i* being paid for the subsequent performance. In the replication and all women conditions, $bwin_i$ represents the belief in the probability of winning the competition against a participant randomly chosen among the participants who have chosen to compete. In the choice of sex condition, participant *i* considers two versions of problem (2). In one version, $bwin_i$ represents the belief in the probability of winning the competition against a man randomly chosen among the men who have chosen to compete. In the other version, $bwin_i$ represents the belief in the probability of winning the competition against a woman randomly chosen among the women who have chosen to compete. In the latter condition participant *i* only considers the higher $bwin_i$ of these two when deciding whether to compete.

Proposition 1: *Suppose* $u_i(n_i)$ *is linear or strictly concave. Then, the solutions of problem (1) and (2) are positive and unique. That is,* $n_i^{nc*} = \arg\max_{n_i} \ \pi_i^{nc} > 0$ *and* $n_i^{com*} = \arg\max_{n_i} \ \pi_i^{com} > 0$.

Proposition 2: *Suppose* $u_i(n_i)$ *is linear or strictly concave. Then:*

(a) *If* $bwin_i > 0.5$ *then* $\pi_i^{com}\left(n_i^{com*}\right) > \pi_i^{nc}\left(n_i^{nc*}\right)$

(b) *If* $bwin_i < 0.5$ *then* $\pi_i^{com}\left(n_i^{com*}\right) < \pi_i^{nc}\left(n_i^{nc*}\right)$

(c) *If* $bwin_i = 0.5$ *then* $\pi_i^{com}\left(n_i^{com*}\right) = \pi_i^{nc}\left(n_i^{nc*}\right)$ (see proofs in the Online Appendix)

Hence, according to proposition 2, the choice of payment scheme of a risk neutral or risk averse participant only depends on a participant's winning belief.[23] Specifically, a participant who believes to have a higher (lower) than 50% chance of winning the competition should choose the winner-take-all tournament (non-competitive payment). Therefore, eliciting a participant's belief in winning the competition allows us not only to learn a participant confidence level but also to evaluate if a participant's actual choice of payment scheme is a rational response to a participant's winning belief in terms of expected payoff maximization.

4.3. Winning belief: the actual impact of confidence level

We elicit the winning belief right after participants make their choice of payment scheme. In the replication and all women conditions we ask participants the following question: "Consider only the other participants who have just chosen to compete in the subsequent performance round. How many of these participants (in percentage terms) did correctly solve less addition problems than you in the 5 minutes round 1 piece-rate performance?". In the choice of sex condition, we ask participants two versions of this question. One version with reference to only the male participants who have chosen to compete; the other version with reference to only the female participants who have chosen to compete. The questions are monetarily incentivized using a scoring rule that yields a higher reward the more accurate the estimate is (see Online Appendix for the full instructions and the details of the monetary incentives).

Following the previous subsection, the elicitation of the winning belief should incorporate two elements: i. Elicit a participant's belief in his/her *own* performance relatively to *only* the *other participants who have chosen to compete*; ii. Participant's estimate is about his/her expected relative *competitive* performance in the *subsequent* performance round. The question we use only incorporates the former element because incorporating the latter element poses a serious problem.[24] As a proxy for the second element we use a participant's estimate of

[23] It is reasonable to assume that participants are risk neutral or risk averse since economics experiments eliciting risk attitude do not find support for risk loving preferences, which corresponds to the case of a convex utility function (see Croson and Gneezy [2009]).

[24] If the question were instead "Consider only the other participants who have just chosen to compete in the subsequent performance round. How many of these participants (in percentage terms) do you expect to correctly solve less addition problems than you in the subsequent performance round?" a too salient hedging opportunity would arise because we

his/her relative performance in the piece-rate round, i.e., we rely on the assumption that participants' belief in the relative performance among the participants who choose competition is similar in both rounds. Spearman rank correlations are consistent with this assumption [replication condition: piece-rate performance (PR) vs. competitive performance (COM), $\rho = 0.731$; choice of sex condition: Men: PR vs. COM, $\rho = 0.749$ | Women: PR vs. COM, $\rho = 0.775$; all women condition: PR vs. COM, $\rho = 0.847$].

In Table 6, we document the winning beliefs. If the self-stereotyping hypothesis were correct, the women's

Table 6. Winning belief (in average percentage)

Note: In the choice of sex condition, in which we separately elicit the wining belief against a man and against a woman, a participant *i*'s winning belief refers to the higher of the two. Standard deviation in parentheses.

	Replication	**Choice of Sex**	**All Women**
Men	44.5% (24.9)	44% (25.6)	n.a.
Women	41.5% (24.4)	44.9% (25.7)	44.6% (24.7)

belief in winning in the replication condition, in which they cannot avoid the possibility of competing against men, should be substantially lower not only than men's belief in winning, but also than women's belief in winning in both the choice of sex and all women conditions. But in fact none of these comparisons is significantly different (41.5% vs. 44.5%, Mann-Whitney test, one-sided, $p = 0.396$; 41.5% vs. 44.9%, Mann-Whitney test, one-sided, $p = 0.264$; 41.5% vs. 44.6%, Mann-Whitney test, one-sided, $p = 0.278$). Although women in the replication condition are directionally less confident, the difference in confidence level is not significant. Hence, we do find support for the self-stereotyping prediction.

Regression analysis corroborates this finding. Comparing regressions (1) and (4) of Table 7, we see that after adding the winning belief as a predictor, the gender gap in the probability of entering into competition (which is measured by the female predictor) in the replication condition only slightly decreases. In addition, comparing regressions (2) and (5) of Table 7, we see that the gender gap in competition entry in the choice of sex condition remains insignificant after including the winning belief as a predictor. In a nutshell, although regressions (4) – (6) of Table 7 show that the winning belief helps to predict entry into competition in each condition, differences in men's and women's confidence level do not explain much of the observed gender difference in entry into competition in the replication condition.[25]

Nonetheless, we find an interesting difference related to confidence level. In Table 8, we break down the entry into competition by low- and high-confident participants. In the replication condition, we observe a

monetarily incentivize the question using a scoring rule that makes a participant's earning higher the more accurate a participant's estimate is. This rule would, therefore, give an incentive for a participant to report an estimate of zero and subsequently solve no problem in order to ensure the maximum payment.

[25]In stark contrast to the winning beliefs, we do not find any significant effect of the stereotype-based beliefs when we add them as a covariate in the regressions of Table 7 (as well as in the regression analysis that we subsequently present in the paper). Therefore, we dismiss this variable in the regression analysis that we present for the sake of relevance.

Table 7. Probit models of payment choice II

Note: The table reports marginal effects. Robust standard errors in brackets. ***, ** and * significant at 1%, 5% and 10%, respectively. In Panel A, we again report the results of Table 3 for ease of comparison. *Winning belief* is a variable that ranges from 0-100.

	Dependent variable: *1 if payment choice is a winner-take-all tournament*					
	Panel A			**Panel B**		
	Replication	**Choice of Sex**	**All Women**	**Replication**	**Choice of Sex**	**All Women**
	(1)	(2)	(3)	(4)	(5)	(6)
1 if female	-0.32***	-0.11	n.a.	-0.28***	-0.11	n.a.
	[0.123]	[0.089]		[0.106]	[0.084]	
Baseline performance	0.05***	0.04***	0.04***	0.03*	0.02*	0.02
	[0.016]	[0.010]	[0.010]	[0.017]	[0.012]	[0.013]
Risk score	-0.007	0.016	0.022	-0.004	0.017	0.024
	[0.028]	[0.017]	[0.025]	[0.023]	[0.017]	[0.023]
Winning belief				0.009***	0.006***	0.005**
				[0.002]	[0.002]	[0.002]
Observations	58	90	56	58	90	56
Pseudo R^2	0.169	0.120	0.133	0.388	0.197	0.190

Table 8. Choice of a winner-take-all tournament split by low- and high-confident participants

Note: *bwin* refers to participants' winning belief. p-values derived from Fisher's exact tests of contingency between men and women. The choices of the participants with a winning belief equal to 50% are: in the replication condition, one woman competes and two women do not compete; in the choice of sex condition, one woman and one man compete and one woman does not compete; in the all women condition, three women compete.

	Replication		**Choice of Sex**		**All Women**	
	bwin < 50	*bwin* > 50	*bwin* < 50	*bwin* > 50	*bwin* < 50	*bwin* > 50
Men	56%	91%	62%	95%	n.a.	n.a.
Women	17%	88%	56%	84%	61%	82%
p-value	0.018[a]	0.678[a]	0.779	0.603	n.a	n.a

[a]one-sided test

considerable lower entry into competition among low-confident women compared to low-confident men. The same is observed when comparing low-confident women in the replication condition to low-confident women in the other two conditions. High-confident women, on the other hand, similarly enter into competition across the three conditions. Therefore, collectively these findings show that the lower inclination of women to enter into competition in the replication condition is driven by the low-confident women.

4.4. Confidence and the choice of sex

The results so far show that women shy away from competition only when they cannot avoid the possibility of competing against men. Since this behavior is not explained by women's lower confidence level relatively to men, we understand the results as evidence of women's dislike of competition against men. However, women's aversion to compete in the replication condition seems to hold only for low-confident women. As a robustness check for whether women's dislike of competition against men is specific to low-confident women, we further examine the women's choice of the competitor's sex taking into consideration women's confidence level.

In Table 9, we show that both low- and high-confident women overwhelmingly, and equally, prefer to compete against a woman. Moreover, regression analysis shows that women's choice of the competitor's sex is not driven by a higher confidence level to compete against a woman (see Online Appendix for details). Looking at men, we observe that both low- and high-confident men's choice of the competitor's sex is not significantly different from a random choice.

Table 9. Choice of the competitor's sex split by low- and high-confident participants
Note: *bwin* refers to participants' winning belief. p-values derived from binomial tests under the null hypothesis that the choices of male and female competitor are equally likely to occur.

		Male competitor	Female competitor	p-value
Men	*bwin* > 50	35%	65%	0.332
	bwin < 50	38%	63%	0.454
Women	*bwin* > 50	20%	80%	0.018[a]
	bwin < 50	21%	79%	0.029[a]

[a]one-sided test

Our conclusion is therefore that women dislike competing against men independently of their confidence level. However, if women cannot avoid the possibility of competing against men (as is the case in the replication condition), high-confident women enter into competition because their belief in a high probability of winning the competition more than compensates their dislike of competition against men.

4.5. The "irrational" competitors

In this subsection we examine whether a participant's choice of payment scheme is a rational response to a participant's winning belief in terms of expected payoff maximization. More specifically, a participant is considered to have a rational response if a participant's actual choice of payment scheme coincides with the choice predicted by proposition 2. Hence, for each participant, we compare a participant's actual choice of payment scheme to the choice predicted by proposition 2 once considering a participant's winning belief.

In column (3) of Table 10, we report the percentage of participants whose actual choice coincides with the predicted choice. We observe that 86% of women in the replication condition have a rational response to their

Table 10. Rational responses in terms of expected payoff maximization

Note: Predicted Choice refers to the percentage of competitive choices predicted by proposition 2 given participants' winning belief; **Rational Response** refers to the percentage of participants whose actual choice coincides with the predicted choice. We again report the percentage of actual competitive choices for ease of comparison.

		Predicted Choice[a] (1)	**Actual Choice** (2)	**Rational Response** (3)
Replication	Men	38%	69%	62%
	Women	33%	38%	86%
Choice of Sex	Men	41%	76%	62%
	Women	42%	67%	62%
All Women	Women	42%	71%	59%

[a]The few participants that estimate a probability of winning equal to 50% (3 women in the replication condition, 2 women and 1 men in the choice of sex condition and 3 women in the all women condition) are treated as equally likely to choose between the two payment alternatives.

winning belief. The percentage of rational responses in the other four cases — men in the replication condition, men and women in the choice of sex condition and women in the all women condition — is significantly lower compared to women in the replication condition (Fisher's exact tests of contingency yield p-value ≤ 0.07 for any of the four comparisons). The lower percentage of rational responses in these four cases seems to be due to an excessive entry into competition. If all participants had a rational response, then only 38% of men in the replication condition, 41% of men and 42% of women in the choice of sex condition, and 42% of women in the all women condition would choose competition (see column (1) in Table 10).

To examine excess entry we analyze the choices of the participants who did not have a rational response. There are two possible types of "irrational" response in our setting: i. The "*irrational" competitor*, i.e., a participant who chooses competition but whose expected payoff maximization choice corresponding to his/her winning belief is the non-competitive payment; ii. The "*irrational" non-competitor*, i.e., a participant who chooses not to compete but whose expected payoff maximization choice corresponding to his/her winning belief is competition.

The results in Table 11 show that the occurrence of both types of "irrational" response is sparse among women in the replication condition. In the other four cases, however, we observe that the percentage of "irrational" competitors is considerable, and clearly higher than the percentage of "irrational" non-competitors. McNemar tests for the significance of changes validate these observations. The occurrence of both responses is not significantly different for women in the replication condition whereas in each of the other four cases the "irrational" competitor response is significantly more often (McNemar tests yield p-value = 0.625 for women in

Table 11. "Irrational" competitors and " irrational" non-competitors

Note: "Irrational" competitors (non-competitors) refers to the participants who choose competition (no competition) but whose choice should be the non-competitive payment (competitive payment) in terms of expected payoff maximization.

	Replication		Choice of Sex		All Women	
	"Irrational" competitors	"Irrational" non-competitors	"Irrational" competitors	"Irrational" non-competitors	"Irrational" competitors	"Irrational" non-competitors
Men	35%	3%	36%	2%	n.a.	n.a.
Women	10%	3%	31%	7%	34%	7%

the replication condition and p-value ≤ 0.013 for any of the other four cases).[26]

In short, we observe that in all cases but for women in the replication condition the percentage of "irrational" competitors is substantial. We regard an "irrational" competitor response as stronger evidence of a participant's preference to compete than if a participant's choice to compete is in accordance with expected payoff maximization.[27] Hence, we argue that the observed low percentage of "irrational" competitors only among women in the replication condition is evidence that women dislike competing against men.

5. Uncovering the belief in the sex of potential competitors

We conclude our analysis by examining participants' belief in the sex of potential competitors. Since a participant is not informed about the gender composition of participants present in the lab, we first examine if (and how) a participant perceives the gender composition of participants present in the lab. Second, we elicit a participant's belief in the likelihood of each sex to enter into competition. Finally, we investigate whether the belief in the sex of potential competitors influences a participant's decision to enter into competition.

5.1. Perception of the gender composition

To assess if (and how) participants perceive the gender composition of participants present in the lab we ask them the following question: "How do you perceive the gender composition of participants in this experimental session? Please choose the option that better describes your perception:". The options available are "I did not notice the gender composition of participants"; "Only female participants"; "Mainly female participants"; "Balanced composition"; "Mainly male participants"; "Only male participants". This elicitation is not monetarily incentivized because rewarding an accurate perception would incentivize the participants who did not notice on the gender composition to guess it.

We start by analyzing if men and women have noticed the gender composition. The results in Table 12 indicate that the overwhelming majority of women reveal that they did notice the gender composition whereas

[26]The McNemar test analyzes whether the two possible types of "irrational" response happen with the same frequency. The null hypothesis for each of the 5 cases we test is that the proportion of "irrational" competitors and non-competitors is the same. This test is not sensitive to the number of rational competitors and rational non-competitors.

[27]Reuben, Wiswall and Zafar (2013) share a similar view.

Table 12. Unawareness of the gender composition (in percentage)

Note: The percentage refers to the men (women) who chose the alternative "I did not notice the gender composition of participants in this experimental session".

	Replication	Choice of Sex	All Women
Men	45%	53%	n.a.
Women	14%	13%	16%

Table 13. Women's perception of the gender composition (in percentage)

Note: In each condition, the percentages are computed relatively to the women who reveal to have noticed the gender composition.

	Only female	Mainly female	Balanced composition	Mainly male	Only male
Replication	0%	4%	88%	8%	0%
Choice of Sex	0%	8%	90%	2%	0%
All Women	87%	11%	0%	2%	0%

for men this aspect has not distinctively attracted their attention. Hence, we proceed the analysis considering only women for the sake of relevance and brevity (see Online Appendix for the corresponding analysis for men).

In Table 13, we report the specific gender composition of participants present in the lab that was perceived by the women who reveal to have noticed the gender composition. We see that in each condition almost all women correctly perceive the actual gender composition of participants present in the lab. This fact makes us confident that the beliefs we present in the following subsection, which are elicited based on the actual gender composition of participants present in the lab, capture the true beliefs of women in the moment they are asked to choose a payment scheme.

5.2. Belief in the likelihood of each sex to enter into competition

To elicit participants' belief in the sex of potential competitors we ask participants the following two questions (presented on the same screen): i. "For your information there are 15 male participants in this experimental session. How many male participants do you think have chosen to compete (excluding yourself if you are a male participant and chose competition)?"; ii. The same question with reference to female participants. The questions are monetarily incentivized using a scoring rule that yields a higher reward the more accurate the estimate is (see Online Appendix for the full instructions and the details of the monetary incentives). In the all women condition we evidently could only ask the second question, which we adjust for the higher number of female participants present in the lab.

Table 14. Women's belief in the likelihood of men and women to enter into competition

Note: Belief probability men enter equals: (estimate of male participants who enter into competition/total male participants in the session)*100; **Belief probability women enter** equals: [estimate of female participants who enter into competition/(total female participants in the session-1)]*100. We report the average of these variables.

	Replication	**Choice of Sex**	**All Women**
Belief probability men enter	70.6% (15.1)	72% (15.9)	n.a.
Belief probability women enter	51.5% (18.0)	64.4% (15.9)	65.4% (16.2)

In the subsequent analysis we include all female participants. The results are qualitatively the same if we exclude the few female participants who reveal not to have noticed the gender composition. In Table 14, we report women's belief in the likelihood of each sex to enter into competition. We observe that women's belief in men's likelihood to enter into competition is essentially the same in the replication and the choice of sex condition (70.6% vs. 72%, Mann-Whitney test, $p = 0.806$). Women's belief in other women's likelihood to enter into competition is similar in the choice of sex and all women conditions (64% vs. 65.4%, Mann-Whitney test, $p = 0.897$), but significantly lower in the replication condition (51.5% vs. 64.4%, Mann-Whitney test, $p < 0.01$; 51.5% vs. 65.4%, Mann-Whitney test, $p < 0.01$). In other words, in the condition in which women cannot avoid the possibility of facing a male opponent in case they compete, women's belief in other women's willingness to compete is, on average, substantially lower.

Moreover, in regressions (1) and (2) of Table 15, we observe that women's belief in the percentage of men

Table 15. Probit models of payment choice III (only women)

Note: The table reports marginal effects. Robust standard errors in brackets. *** and ** significant at 1% and 5%, respectively. *Belief percentage men among competitors* indicates the believed proportion of men among the participants who choose competition, i.e., it equals: [estimate of male participants who enter into competition/(estimate of male participants who enter into competition + estimate of female participants who enter into competition)]*100. *Winning belief* is a variable that ranges from 0-100.

Dependent variable: *1 if payment choice is a winner-take-all tournament*		
	Replication (1)	**Choice of sex** (2)
Baseline performance	0.03 [0.040]	0.02 [0.021]
Risk score	-0.02 [0.043]	0.01 [0.026]
Winning belief	0.010*** [0.002]	0.005** [0.002]
Belief percentage men among competitors	-0.024** [0.012]	-0.014 [0.009]
Observations	29	45
Pseudo R^2	0.468	0.191

among the participants who choose competition significantly predicts women's choice for competition in the replication condition but not in the choice of sex condition. The insignificance of this predictor in the choice of sex condition is logical because the proportion of men among the participants who choose competition does not affect the chance of facing a male competitor in case a participant chooses to compete. The significance of this predictor in the replication condition corroborates the finding that women dislike competing against men because in that condition this predictor captures women's belief in the probability of facing a male competitor. Therefore, regression (1) of Table 15 specifically indicates that, controlling for performance, risk score and winning belief, women who belief in a higher probability of facing a male competitor are significantly less likely to choose competition.

Less intuitive is the more detailed result we present in regression (4) of Table 16, in which we separately

Table 16. Probit models of payment choice IV (only women)

Note: The table reports marginal effects. Robust standard errors in brackets. *** and ** significant at 1% and 5%, respectively. *Winning belief*, *Belief probability men enter* and *Belief probability women enter* are variables that range from 0-100.

	Dependent variable: *1 if payment choice is a winner-take-all tournament*					
	Panel A			**Panel B**		
	Replication (1)	**Choice of Sex** (2)	**All Women** (3)	**Replication** (4)	**Choice of Sex** (5)	**All Women** (6)
Baseline performance	0.03 [0.033]	0.03 [0.021]	0.02 [0.013]	0.03 [0.038]	0.03 [0.016]	0.01 [0.011]
Risk score	0.01 [0.035]	0.01 [0.026]	0.02 [0.023]	-0.01 [0.041]	-0.03 [0.025]	0.01 [0.024]
Winning belief	0.010*** [0.002]	0.005** [0.003]	0.005** [0.002]	0.010*** [0.002]	0.005** [0.002]	0.005** [0.002]
Belief probability men enter				-0.008 [0.007]	0.005 [0.004]	n.a.
Belief probability women enter				0.011** [0.005]	0.014*** [0.003]	0.010*** [0.003]
Observations	29	45	56	29	45	56
Pseudo R^2	0.340	0.166	0.190	0.492	0.366	0.322

report the impact of women's belief in men's and other women's likelihood to enter into competition. We see that women's lower inclination to compete in the replication condition is not driven by women's belief in men' likelihood to compete, but rather by women's belief in other women's likelihood to enter into competition. Moreover, regressions (5) and (6) of Table 16 show that in the choice of sex and all women conditions the more a woman believes other women enter into competition, the more likely a woman enters into competition herself. Finally, important for the policy implications that we draw in the next section, comparing regression (4) – (6) to regressions (1) – (3) of Table 16, we further observe that the women's belief in other women likelihood to enter into competition and the women's winning belief predictors are orthogonal.

Summing up, in each condition women's willingness to compete is higher the more they believe other women

Table 17. Probit models of payment choice V (only women, pooling the three conditions)
Note: The table reports marginal effects. Women in the choice of sex condition are the base group. Robust standard errors in brackets. *** and * significant at 1% and 10%, respectively. *Winning belief* and *Belief probability women enter* are variables that range from 0-100.

Dependent variable: *1 if payment choice is a winner-take-all tournament*			
	(1)	(2)	(3)
Replication	-0.25***	-0.23***	-0.11
	[0.092]	[0.086]	[0.084]
All Women	0.05	0.04	0.02
	[0.088]	[0.084]	[0.074]
Baseline performance	0.04***	0.02*	0.01*
	[0.010]	[0.010]	[0.010]
Risk score	0.01	0.02	-0.01
	[0.017]	[0.016]	[0.015]
Winning belief		0.006***	0.006***
		[0.002]	[0.001]
Belief probability women enter			0.010***
			[0.002]
Observations	130	130	130
Pseudo R^2	0.164	0.247	0.360

also compete. However, women's belief in other women's likelihood to enter into competition in the replication condition is, on average, significantly lower. Put together, these finding indicate that women compete less in the replication condition because they believe fewer women compete. The specification we use in Table 17, in which we pool the three conditions, clearly demonstrates this connection. Comparing regression (1) – (3) of Table 17 we observe that adding women's belief in the likelihood of other women to enter into competition as a predictor to the set of controls makes insignificant the difference in self-selection into competition between women in the replication condition and women in the choice of sex condition. Moreover, we obtain the same conclusion when we regress specifications (1) – (3) of Table 17 using either women in the all women condition as the base group or women in the choice of sex and all women conditions pooled as the base group. Thus, differences in women's belief in other women willingness to compete greatly explain the observed lower inclination of women to enter into competition in the replication condition compared to women in the choice of sex and all women conditions.

6. Policy implications

6.1. Women's willingness to compete and their confidence level

Our results indicate that implementing single-sex competition or allowing women to choose the sex of their competitors promote women to compete more. Such institutional changes, however, are usually out of

question.[28] Nevertheless, understanding the competitive behavior of women in these two situations helps us to gain insight into interventions that could encourage women to compete in the more realistic case in which they cannot avoid the possibility of competing against men.

In the regression (1) – (3) of Table 16, we observe that higher confidence significantly predicts more entry of women into competition in every condition. Moreover, the predictive magnitude and significance of women's winning belief as well as the explanatory power of the specification after adding the winning belief as a predictor is stronger in the replication condition.[29] Our results indicate, therefore, that a policy intervention that increases the confidence level of women should boost their entry into competition, and this kind of intervention should be particularly effective in situations where women cannot avoid the possibility of competing against men. Furthermore, the results in section 4 indicate that women on the verge of a mixed-sex competition are not less confident than men, but rather low-confident women enter much less into competition compared to low-confident men. Thus, interventions that boost women's confidence level should particularly target low-confident women because they are the more susceptible to change their competitive behavior.

Our results may therefore explain why a less invasive intervention as advice for women in Brandts, Groenert and Rott (2014)[30] is ineffective, but a strong intervention as preferential treatment for women in the form of a head start in Balafoutas and Sutter (2012) boosts women's competition entry in a mixed-sex context. However, the results of Balafoutas and Sutter (2012) also illustrate the likely side effect that would emerge if that kind of policy were implemented outside the lab; comparing Balafoutas and Sutter (2012) major head start treatment to their treatment without intervention, women's competition entry increases from 31% to 69%, but men's competition entry decreases from 64% to 50%.[31]

In a nutshell, there is room to intervene via women's confidence level. Our results particularly suggest that boosting the confidence of low-confident women is critical to make women compete more in a mixed-sex context. More research is needed though to test specific interventions that could encourage women to compete without discouraging men.[32]

6.2. Women's willingness to compete and their belief in the sex of potential competitors

The evidence we present shows a systematic inclination of women to enter into competition the more they believe other women are entering. Moreover, women's higher percentage of competition entry in the choice of

[28]Notable exceptions exist, though. In chess, for instance, women may choose to compete in mixed-sex or only-women tournaments.

[29]Regressing Panel A models of Table 16 without the winning belief predictor yield a pseudo R^2 of 0.105, 0.116 and 0.133 in the replication, choice of sex and all women conditions, respectively.

[30]Brandts et al. (2014) concerned with the reverse discrimination costs that affirmative action policy could entail, investigate whether advice by more experienced individuals could be an alternative to promote women to compete more. They find that this intervention does not significantly increase women's competition entry.

[31]Regressing the models (1) and (2) of Table 16 for men shows that higher confidence significantly predicts more entry into competition. Thus, lower confidence may explain men's lower entry into competition when there is a head start for women.

[32]From an efficiency perspective, we could argue that discouraging men is actually good if it drives out the less productive men. Yet, we leave aside that perspective in our analysis. Our experimental setting is not appropriate to quantify the cost/gain associated to participants' entry into competition based on participants' work task productivity because we purposely align the payoffs of the payment alternatives we offer.

sex and all women conditions compared to the replication condition reflects the fact that women in the former conditions believe that significantly more women compete. However, as aforementioned, institutional changes to implement single-sex competitions or to allow women to choose the sex of their competitors are usually not realistic.

Yet, an institutional intervention in mixed-sex competitions that in all likelihood induces women to perceive the competition more sex-specific is affirmative action in the form of quotas. Niederle et al. (2013), implementing a competition in mixed-sex groups with two possible winners, find that women increase their competition entry from 31% to 83% when a quota ensuring the better performing woman among the winners is introduced. However, an intervention in the form of quotas also illustrates the likely side effect that would emerge if that kind of policy were implemented outside the lab; men decrease their competition entry from 74% to 45% when the female quota is introduced.

Without wanting to mean that the use of a quota cannot be justified in any circumstance, we suggest an alternative institutional intervention to promote women's entry into male-typed domains that is unlikely to affect men's behavior. In the light of our results, we advocate that drawing women's attention to women who seek to work and/or are working in male-typed domains is a very promising mechanism to persuade women to enter into these domains.

To illustrate our idea, we take as a reference the economics academic job market. In this male-typed domain, job advertisements aiming to increase the application of women commonly highlight that women are underrepresented, which implicitly suggests that women fail to apply (see, e.g., https://econjobmarket.org/). According to our results, however, this kind of framing should be avoided. Rather than highlighting that women are a minority, we suggest that job advertisements seeking more female applicants should instead include positive information about recent female applicants and/or hiring of female staff. For example, a framing such as "The organization # wants to increase its proportion of female staff in areas where women are underrepresented. Women are therefore explicitly encouraged to apply", which is commonly found in economics academic job market advertisements, should be replaced by something like "The recent considerable increase of female employees has been contributing to our aim to promote a balanced representation of men and women among our staff. Therefore, we strongly encourage women to keep applying". Although a lot can still be done to improve women's representation, the recent decades have been witnessing an increase of women working in male-typed domains. Thus, employers certainly can find information about past female applicants and/or female hires to be used as a positive cue in job advertisements.

In brief, we advocate that highlighting female role models rather than on-going interventions that highlight women's underrepresentation is the appropriate intervention to boost the participation of women in male-typed domains. Moreover, this kind of intervention would be less likely to affect the behavior of men (and, thus, less controversial) compared to affirmative action in the form of quotas. Last but not least, our results also indicate that women's belief in other women's willingness to compete positively influences their inclination to compete independently of the confidence level, which makes this type of intervention complementary to policy intervention targeting women's confidence level.

7. Concluding remarks

The present study was designed to determine whether different cues to the sex of potential competitors alter women's willingness to compete in a male-typed domain. Our hypothesis was that women do not have a weaker preference to compete per se, but rather just dislike competing against men.

Our findings shed new light on the burgeoning literature on gender competition. To the best of our knowledge we provide the first attempt to explicitly test in the lab the connection between women's belief in the sex of potential competitors and their decision to enter into competition. Our results demonstrate that women dislike competing against men in a male-typed domain. Specifically, we find that women's unwillingness to compete in a situation where women cannot avoid the possibility of a mixed-sex competition in case they enter into competition is significantly driven by women's belief in a higher proportion of men among potential competitors. In contrast, when women can avoid the possibility of a mixed-sex competition in case they enter into competition, women compete as much as men do.

These findings accommodate an explanation based on stereotype threat. More specifically, women likely develop over time a dislike for competition against men in male-typed domains as a response to the repeated exposure to the situational predicament created by stereotype threat in these situations. Surprisingly, however, we find the interesting result that women's lower inclination to enter into a mixed-sex competition in a male-typed domain is not driven by their belief in men's likelihood to enter into competition, but rather by women's belief that other women are less inclined to enter into competition in this situation (which in turn increases women's belief in potential male competitors). Perhaps women behave according to what they believe other women do just because they are a woman themselves. This would be consistent with the psychology literature on gender typing (e.g. Bem [1981]) and the sociology literature on homophily (e.g., McPherson, Smith-Lovin and Cook [2001]), which is the tendency of people to interact with individuals who resemble them; future research will have to test the relevance of these motives.

In addition, we confirm previous findings that women's confidence level predicts their willingness to enter into competition. However, we find little support for negative self-stereotyping (Bordalo et al. [2014]), i.e., for a lower confidence level of women when a mix-sex competition is at stake. In this situation, we observe instead a remarkable lower inclination to enter into competition of low-confident women compared to low-confident men. We should also highlight that we use a different approach to measure men's and women's confidence level; we elicit participants' confidence self-assessment relatively to *only* the participants who enter into competition. Hitherto the standard approach has been to measure participants' confidence level relatively to all participants present in the lab, regardless of whether they choose competition or not. Hence, our results indicate that men's and women's confidence level relatively to individuals with a preference to compete are more aligned than their confidence level relatively to individuals in general.

Our results have important policy implications. A growing literature that investigates institutional mechanisms to promote women to compete more (e.g., Balafoutas and Sutter [2012]; Niederle et al. [2013]) shows that preferential treatment in the form of quotas or a head start for women boosts women's competition

entry (but also decreases men's competition entry). Our findings suggest an alternative course of action to encourage women to compete that is unlikely to affect men's willingness to compete and, consequently, to be less controversial. While attempting to encourage women to enter into competitive male-type domains, we should highlight women who do enter or seek to enter into these domains, rather than highlighting women's underrepresentation. Namely, we advocate that the inclusion of information in job advertisements pointing out women who do apply and/or get hired in male-typed domains is a very promising intervention to explore.

In closing, some words of caution are in order. First, in contrast to the literature, our results are derived in a setting in which women's risk attitude does not influence their decision to enter into competition. Therefore, whether our findings might be transferable to a setting in which risk attitude is important is an open question that our data cannot answer. Second, some studies quantify the cost/gain associated to a participants' entry into competition based on participants' work task productivity (see, e.g., Niederle et al. [2013], Balafoutas and Sutter [2012], Almås et al. [2012]). Our setting, though, is not suitable to perform such kind of analysis because we purposely align the payoffs of the payment alternatives we offer. Third, our study only considers the supply-side. Even if supply-oriented interventions are effective in increasing women's willingness to enter into male-typed domains, this would be insufficient in case employers have a strong bias to favor men in the hiring process of male-typed domains (see, e.g., Reuben et al. [2014]). Therefore, parallel interventions in the demand-side that make the hiring process as much gender blind as possible are certainly desirable (see, e.g., Goldin and Rouse [2000]).[33] Finally, our analysis and discussion was directed to understand how to promote women to compete more in a mixed-sex context. However, there are many circumstances in which being competitive could not be desirable, as for example working in mixed-sex teams. From this perspective, our results indicate that men compete too much, not that women compete too little. Hence, for mixed-sex contexts in which cooperation is the most desirable attribute, the relevant question to be studied seems to be how we could promote men to be less competitive rather than women to be more competitive.

Supplementary material

The Online Appendix contains the supplementary material of this article.

.

[33]Gender bias seems to be present not only in the hiring process but also in the promotion process. Leading organizations like McKinsey, BCG, Goldman Sachs and Catalyst are increasingly searching for mechanisms to tackle this problem.

REFERENCES

Almås, I., Cappelen, A., Salvanes, K., Sorensen, E., & Tungodden, B. (2012). Willingness to Compete in a Gender Equal Society. Working Paper.

Balafoutas, L., & Sutter, M. (2012). Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory. *Science , 335*, 579-82.

Bem, S. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review , 88(4)*, 354-64.

Blau, F., Ferber, M., & Winkler, A. (2013). *The Economics of Women, Men, and Work.* Englewood Cliffs, NJ: Prentice Hall.

Booth, A., & Nolen, P. (2012). Choosing to compete: How different are girls and boys? *Journal of Economic Behavior & Organization , 81(2)*, 542-55.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2014). Stereotypes. Working Paper.

Brandts, J., Groenert, V., & Rott, C. (2014). The Impact of Advice on Women's and Men's Selection into Competition. *Management Science* , forthcoming.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, Competitiveness and Career Choices. *The Quarterly Journal of Economics , 129(3)*, 1409-47.

Cason, T., Masters, W., & Sheremeta, R. (2010). Entry into Winner-Take-All and Proportional-Prize Contests: An Experimental Study. *Journal of Public Economics , 94(9-10)*, 604-11.

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature , 47(2)*, 1–27.

Dohmen, T., & Falk, A. (2011). Performance Pay and Multi-dimensional Sorting: Productivity, Preferences and Gender. *American Economic Review , 101(2)*, 556-90.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. (2011). Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. *Journal of the European Economic Association , 9(3)*, 522-50.

Fischbacher, U. (2007). z-Tree Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics , 10(2)*, 171–78.

Gneezy, U., Leonard, K., & List, J. (2009). Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society. *Econometrica , 77(5)*, 1637-64.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics , 118(3)*, 1049-74.

Goldin, C., & Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *American Economic Review , 90(4)*, 715-41.

Greiner, B. (2004). The online recruitment system ORSEE 2.0—A guide for the organization of experiments in economics. *Working Paper Series in Economics 10, University of Cologne* .

Gupta, N., Poulsen, A., & Villeval, M. (2013). Gender matching and competitiveness: experimental evidence. *Economic Inquiry , 51(1)*, 816-35.

Inzlicht, M., & Schmader, T. (2012). *Stereotype threat: Theory, process, and application*. New York: Oxford Univ. Press.

Kamas, L., & Preston, A. (2010). Are Women Really Less Competitive than Men. Working Paper.

McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology , 27*, 415-44.

Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away from Competition? Do Men Compete Too Much? *Quarterly Journal of Economics , 122(3)*, 1067-1101.

Niederle, M., & Vesterlund, L. (2011). Gender and Competition. *Annual Review in Economics , 3*, 601–30.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness. *Management Science , 59(1)*, 1-16.

Reuben, E., Sapienza, P., & Zingales, L. (2014). How Stereotypes Impair Women's Career in Science. *Proceedings of the National Academy of Sciences , 111(12)*, 4403-08.

Reuben, E., Wiswall, M., & Zafar, B. (2013). Preferences and Biases in Educational Choices and Labor Market Expectations: Shrinking the Black Box of Gender. Working Paper.

Spencer, S., Steele, C., & Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology , 35(1)*, 4–28.

Sutter, M., & Rützler, D. (2014). Gender Differences in the Willingness to Compete Emerge Early in Life and Persist. *Management Science* , forthcoming.

Wozniak, D., Harbaugh, W., & Mayr, U. (2014). The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices. *Journal of Labor Economics , 32(1)*, 161-98.