The gender gap in competitive behavior

Chumbley, JR, Egg, ME, Huber, D, Fehr, E

Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland.

**Abstract:** Women typically compete less than men, particularly when they lack prior information abouttheir relative ability   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Niederle & Vesterlund, 2007; Wozniak, 2009). A critical question is therefore whether women actively acquire this information. We recorded subjects' preference for contests with differing levels of performance feedback: public, private and none at all. Our data confirms that prior information indeed reduces the gender gap but reveals that women disproportionately avoid this information. This tendency does not exclusively reflect self-confidence, risk or ambiguity preferences. Our results suggest that a self-perpetuating information asymmetry may maintain the gender gap.

# 1. Introduction

It is often argued that women are less competitive than men e.g.   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Croson & Gneezy, 2009; Datta Gupta, Poulsen, & Villeval, 2005; Gneezy, Niederle, & Rustichini, 2003a; Niederle & Vesterlund, 2007), a conclusion which likely has undesirable consequences for gender equality in the labor market   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Wirth, 2009). This difference in competitiveness allegedly explains early career choices   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Kamas & Preston, 2012; Kleinjans, 2009) and reflects gender differences in aversion to risk   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Croson & Gneezy, 2009), stress or ’pressure’   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Shurchkov, 2012), family/professional identity conflict   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Cadsby, Servatka, & Song, 2011) or a different taste for the thrill of competition *per se*   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Niederle & Vesterlund, 2007). Here we focus on the roles of prior and posterior information in the gender gap. The former captures the impact of current knowledge, the latter quantifies the preference for new performance feedback.

Prior information should encourage underconfident or ambiguity-averse individuals to compete more   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Gneezy & Pietrasz, 2013). Accordingly, it resolves the gender gap   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Ertac & Szentes, 2011; Wozniak, 2009) in part by combating the male overconfidence   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Niederle & Vesterlund, 2007) and/or female underconfidence   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Cotton, McIntyre, & Price, 2010) which themselves exist for various reasons   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Croson & Gneezy, 2009; Datta Gupta et al., 2005; Große & Riener, 2010; Kamas & Preston, 2012) e.g. ’gender-task stereotype’  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Kleinjans, 2009). Secondly, by reducing ambiguity   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Ellsberg, 1961) about the probability of success, it may combat any gender differences arising from gender-dependent ambiguity preferences   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Borghans, Heckman, Golsteyn, & Meijers, 2009; Schubert, Brown, Gysler, & Brachinger, 2000). Despite these benefits of prior information, people may actively avoid it in order to protect self- or social esteem ([Maslow 1963](#_ENREF_2); [Josephs, Larrick et al. 1992](#_ENREF_1)). Various evidence suggests that women are more sensitive to negative performance feedback more than men, attributing it more to their general ability than chance   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Dweck, 2000; Roberts & Nolen-Hoeksema, 1989).

We isolated the impact of prior information (ambiguity) by comparing voluntary participation in ability-matched versus unmatched contests, i.e. the probability of competing in column 1 versus 2 of Figure 1*ii*. More specifically, because *unambiguous contests* were ability-matched, subjects knew they had a probability of beating their four competitors: In *ambiguous contests,* the risks were unclear. We isolated the *preference for public feedback* (social evaluation) by comparing participation in contests which differed in whether relative-performance feedback was public or private, i.e. the probability of competing in rows 1 versus 2 of Figure 1*ii*. This definition ensured that potential material and self-esteem losses/gains are constant and only the publicity varies. We isolated *the preference for private feedback* (self-evaluation) by comparing contests with private feedback versus no feedback at all, i.e. the relative probability of competing in rows 2 versus 3 of Figure 1*ii*. To assess whether preferences for prior and posterior information are related, we inspected the interaction terms of our factorial design: i.e. between or of Figure 1*ii*. We compared the strength of each preference in men versus women.

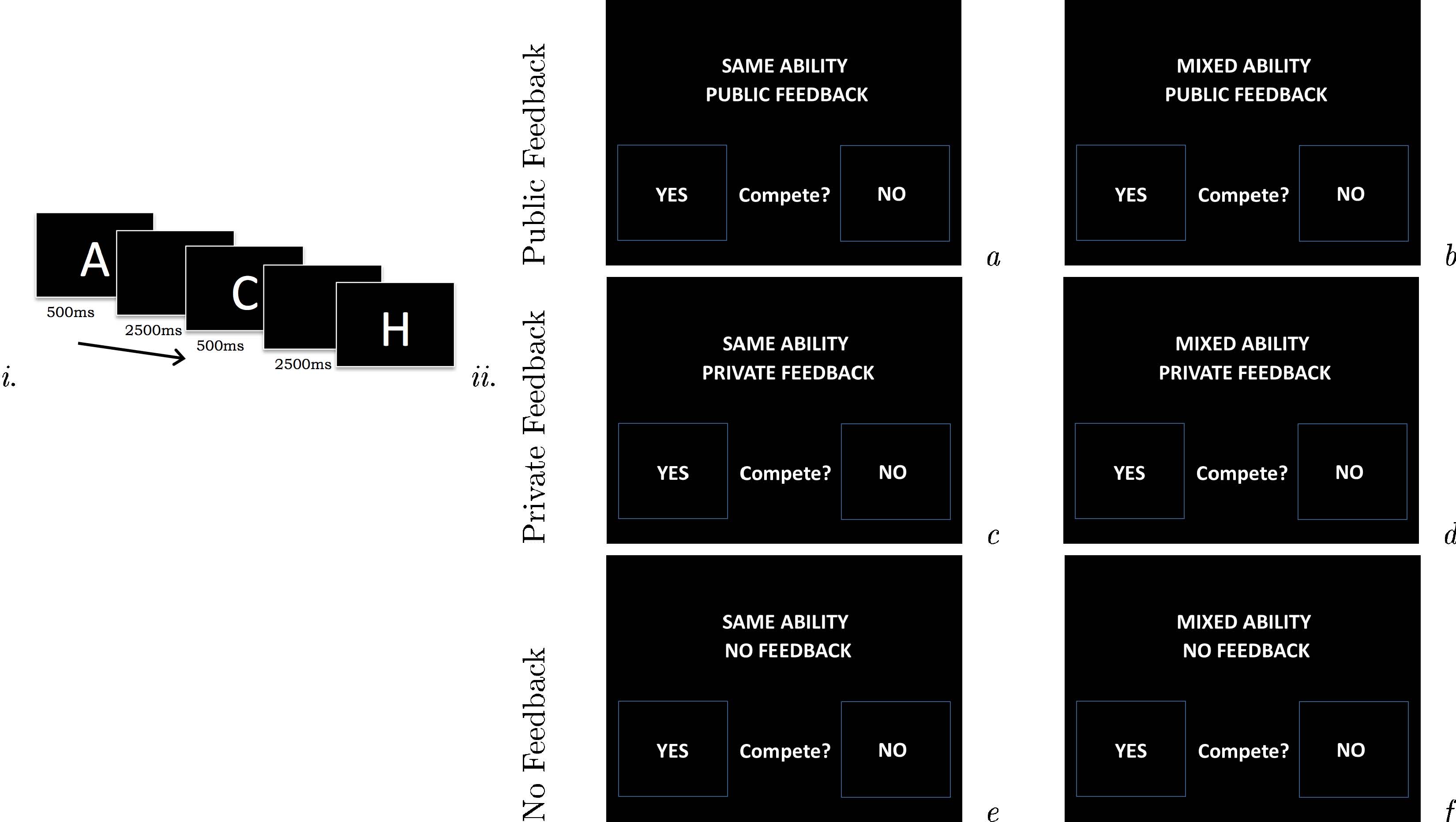


Figure 1: Contests and choices. i. Contests involved performance on the N-back test. Subjects observed a stream of letters on their computer screen and were required to press the space bar every time they saw any letter which had appeared N steps earlier in the sequence. This involved remembering and comparing new stimuli against their memory. N progressively ranged from one to four to make the task more challenging. N-back performance ability is correlated with IQ, a socially valued trait. ii. These six decision screens were presented in random order and subjects had unlimited time to decide. N-back contests varied in the nature of feedback and prior knowledge of one’s relative ability (ambiguity) according to a 3×2 factorial design.

# 2. Methods

## 2.1. Subjects

160 subjects completed informed consent (see Supplementary Material 1.1). Four were excluded due to a software failure to record their choices. The remaining 156 subjects (18-50 years, median age 23 years) comprised 77 women. The sample size was determined based on previous related studies and was fixed in advance. The study was approved by local ethics committee. Subjects were not deceived in any part of this study.

## 2.2. Procedure

On each day of testing, a group of 25 or 30 subjects were welcomed and invited to complete informed consent. We then introduced them to the N-back task verbally and photographed each subject’s face in portrait, before relocating subjects to a behavioral lab in the same building.

On subjects’ arrival, the behavioral lab had low ambient lighting and a wildlife photograph was projected onto the wall, thereby introducing subjects to the technology of public feedback. This photograph was removed once all subjects were seated. Subjects were seated at individual computer booths and twice performed the N-back working memory test. See Supplementary Material 1.2 for specifics of the N-back test we used. Once they were paid according to their absolute performance and once according to their relative performance within a randomly-composed group of five. *Absolute pay* was 0.10 Swiss Francs (CHF) per percentage point on the N-back test (1 CHF 1 USD). *Relative pay* was 0.50 CHF per percentage point for the winner and 0 CHF for the four non-winners. See Supplementary Material 1.3 for the verbatim instructions given to subjects, and Supplementary Material 1.2 for how their N-back score was calculated. Subjects did not receive any feedback about their status or pay. The order of these two performances - forced contest versus forced non-contest - was counter-balanced between subjects. After 25 minutes subjects freely chose whether to compete in each of six potential contests i.e. whether to be paid in absolute or relative terms. This choice was the main dependent measure in our analysis. Subjects were instructed that one contest would be selected at random for real play and pay.

 2. Timeline

Before making their choices, subjects were told that if they chose to compete in a contest they would be “grouped with four other people”. See Supplementary Material 1.4 for the verbatim instructions. They were not told whether these group members had themselves chosen to compete. Instead they were told that this information was irrelevant to their chances of success: “N-back performance ability is quite stable over different contexts, e.g. competition versus non-competition”. This instruction was based on our previous data, see Supplementary Material 1.7 for details.

Across different contests (see Figure 1) information about the outcome - i.e. players’ ranks - was either absent, displayed privately on the subject’s personal computer monitor or publically projected onto the wall, together with the subject’s photograph. Subjects were instructed to expect a “random and secret bonus” added to their total pay and were therefore unable to infer their rank in the first of these contests from their pay. See Supplementary Material 1.4 for the verbatim instructions given to subjects. In unambiguous hierarchies, subjects knew performance would be compared to four other people with the same ability, so they should have a one in five chance of winning. In ambiguous hierarchies, subjects knew nothing about their relative ability. These two cases parallel the distinction between “risk” and “ambiguity” in the terminology of behavioral economics   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Chumbley et al., 2012; Ellsberg, 1961). To evaluate subjects relative performance in each of these two cases, our computer program assigned each subject to two groups, each with five members. The mixed-ability group was composed completely at random, by partitioning the participants into groups arbitrarily, independently of each group member’s choice. The matched-ability group was composed of individuals with similar N-back ability, again independently of each group members choices: to do this we collected every subject’s N-back score on the forced contest in real time using a basic client-server network, ranked subjects and then stratified them into groups of five. For any participant who chose to compete in an unambiguous or ambiguous contest, we compared their final performance with the final performance of the four other individuals specified in these two partitions, respectively. The six choices were presented in random order and subjects indicated their choice to compete or not by pressing “E” or “I” on a qwerty keyboard.

After their choices, but before the final N-back test, subjects reported how they expected to perform in a mixed-ability contest, by indicating their expected rank from one to five in this ambiguous hierarchy. They were instructed that they would win 5 CHF if they were correct, incentivizing them to disclose their expected rank honestly. We did not ask them to estimate their rank in matched contests, which should be close to three on average, if each person’s rank is determined by chance. Additionally, subjects chose between gambles designed to reveal their preference between a. sure versus risky payoffs b. sure versus ambiguous payoffs and c. risky versus ambiguous payoffs. The order of these gambles was randomized and subjects were told that they would be paid according to their choice. See Supplementary Material 1.5 for the verbatim instructions given to subjects. After all subjects had completed their choices, subjects were instructed - on their screen - which contest had been randomly selected. They then performed their final N-back test.

Following the final N-back performance the respective feedback was given to subjects. They then completed some computerized questionnaires, assessing the following relevant personal information: risk attitudes, personality measures on reward- and punishment-sensitivity scales (BIS/BAS), gender and age, see Supplementary Material 1.6 for details.

## 2.3. Statistical methods

### 2.3.1. Gender differences in competing

Subjects chose whether or not to enter each of six contests. They were not explicitly asked about their preferences between contests. Instead we can infer this from differential behavior between contests. Our analysis focuses on these *preferences between contests*, i.e. the relative probability of entering two contests that differ only on the experimentally-controlled dimensions of Figure 1*ii*. In this way, while the choice to compete in any single type of contest may be caused by a host of factors, we can isolate specific effect of prior and posterior information by comparing subjects’ behavioral responses to ambiguous contests, public contests and no-feedback contests with their response to a baseline contest, the private-matched contest. We formalized these comparisons in the following model.

The binary choice of subject to compete in contest is denoted . Each subject’s preferences for feedback and ambiguity were defined relative to their probability of entering the private, matched contest, itself denoted in Equation 1. This reference treatment is shown in cell of Figure 1*ii*. The parameter captures their additional probability of entering contests with public feedback; captures their additional probability of entering contests with no feedback. The parameter captures ambiguity preferences *per se*, the additional probability with which they enter unmatched contests, relative to the baseline (matched) contest. While in principle the preference for posterior feedback may depend on prior ambiguity, we found no significant interactions between the rows and columns of Figure 1*ii*, and omit these terms from the models considered below.

We used mixed effects logistic regression to quantify between-subject variation in these preferences as a function of gender , an indicator variable which equals for females and otherwise, see Equation 2. Between-subject variability not explained by gender was modeled as Gaussian random effects where the off diagonal components of capture within-subject correlations (repeated measures).

|  |  |  |
| --- | --- | --- |
|  |  | (1) |
|  |  | (2) |

Despite our focus on preferences between contests, for completeness we quantified gender differences within each cell, row and column of Figure 1*ii* and in total as follows. To identify gender differences in each single cell we used logistic regression with a dummy variable for sex. To quantify gender differences in each row and column of Figure 1*ii*, and in total we used binomial regression. These models have the form where N - the number of contests considered - equals 1, 2 or 3 depending on whether we are considering an individual cell, row or column. The single parameter  β captures the additive effect of gender on the probability of competing. We refer to such parameters as β in the results section, where the context should indicate to which model we refer.

### 2.3.2. Gender differences in confidence or risk-taking

We used logistic regression to identify gender differences in risk and ambiguity preferences in a *non-competitive* context. We used ordered logistic regression to identify gender differences in confidence - self-reported predicted rank on a scale of one to five.

### 2.3.3. Gender differences in confidence or risk-taking: do these underpin competitive preferences?

We wanted to examine potential gender differences in competitiveness after accounting for confidence and/or domain-general preferences for risk and ambiguity (see previous section). We therefore augmented the right hand side of Equation 2 with two new sets of covariates. First, as a proxy to each subject’s confidence we used their self-reported prediction about their rank, . Second, as a proxy for their risk and ambiguity preferences in a non-competitive context we used their binary choices in three gambles: these choices revealed their preference between sure vs. risky financial payoffs, sure vs. ambiguous payoffs and risky vs. ambiguous payoffs. The binary components of indicate these three choices respectively. Thus, we augemented Equation 2

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

where is now a matrix of parameters Encoding the impact of each of these variables on the three preferences plus , the baseline competing probability. We then considered nested sub models with either or as between-subject covariates.

# 3. Results

## 3.1. Gender differences in competing

We first report gender differences in each cell, row and column of Figure 1*ii*, and in total. Men competed more than women in both ambiguous private (, , ) and ambiguous public (, , ) contests. There were no significant differences in any other cell. See Figure 2*.* Men competed significantly more than women in total, (, , ), in public contests, i.e. row 1 (, , ), and in unmatched contests, i.e. column 2 (, , ). There were no significant gender differences in other rows or columns.

We then considered gender-specific preferences *between* contests. As defined in the introduction and Equation 2 this permits us to better isolate the causal role of information in competitive behavior, and to identify gender differences in role of information. Men had a preference against public feedback (, , ). They were also averse to having no feedback: they preferred private feedback to no feedback (, , ). Men had a stronger preference for ambiguous contests than women: this gender difference was reduced (eliminated) by ability-matching (, , ), see figure 3. There were no gender differences in the preference for public feedback or no feedback, albeit a trend towards men having a stronger preference for private information (, , ).

To further probe gender differences in feedback sensitivity, we then considered a more coarse measure of preference for feedback. Specifically, we contrasted an individual’s participation in public contests (which simultaneously convey information to both to oneself and others) with their participation in uninformative contests (which convey neither). To do this we contrasted rows 1 versus 3 of Figure 1*ii*, simply by changing the baseline condition in equation 1 from the private-matched contest to the no-feedback-matched contest, i.e. from cell to cell of Figure 1*ii*. Re-estimating this model revealed that women have a stronger preference against public contests relative to uninformative contests[[1]](#footnote-1) (, , ).

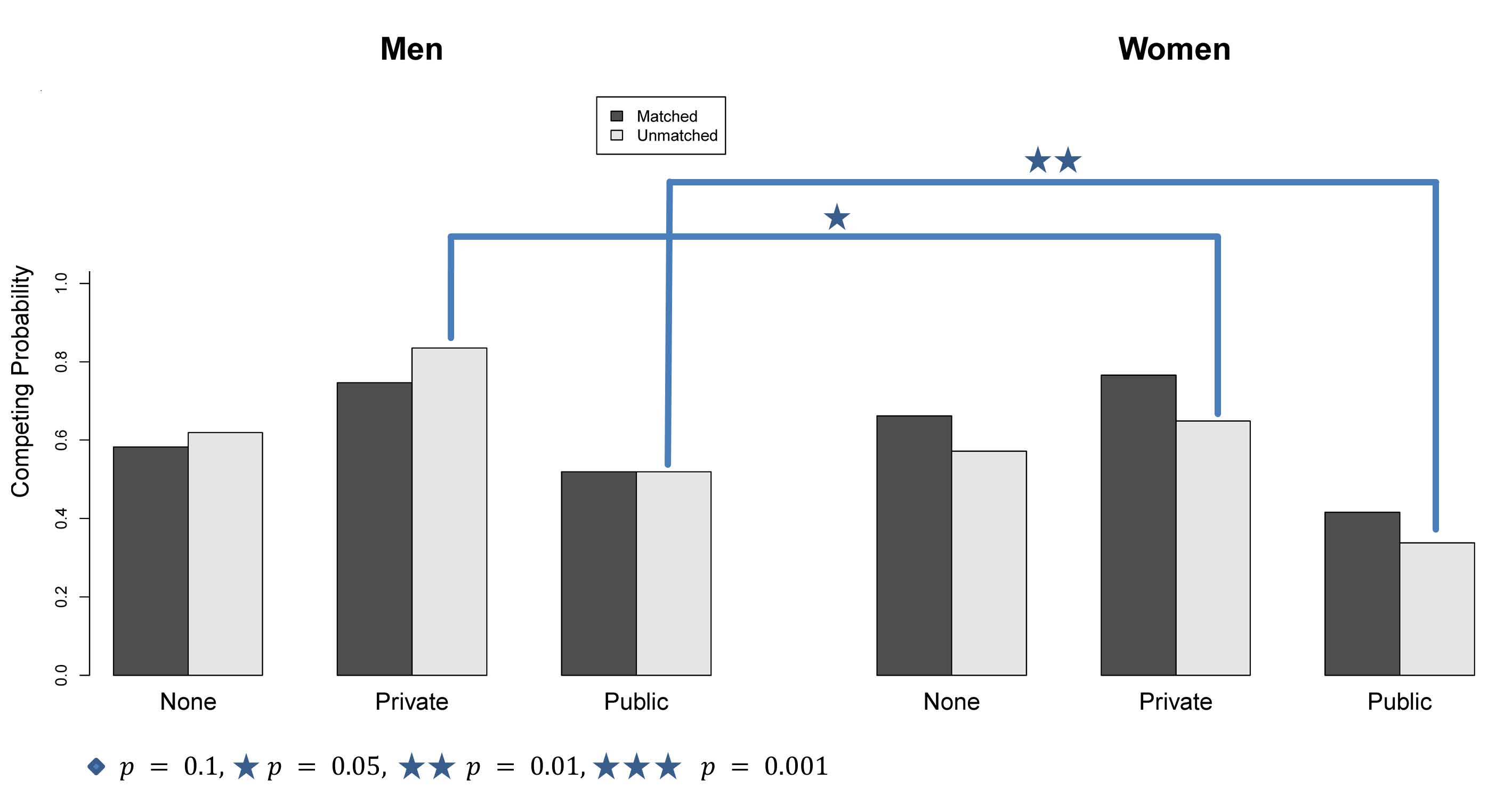


Figure 2: Preferences for feedback and ambiguity in men and women. This plot gives the fraction of men and women who enter each of the 6 contests depicted in Figure 1*ii*. As quantifed in the main text and Figure 4, women have an aversion to ambiguous (not ability-matched) and public contests. In Figure 4 below, we therefore depict them in isolation, i.e. depict row data, summing over columns etc.

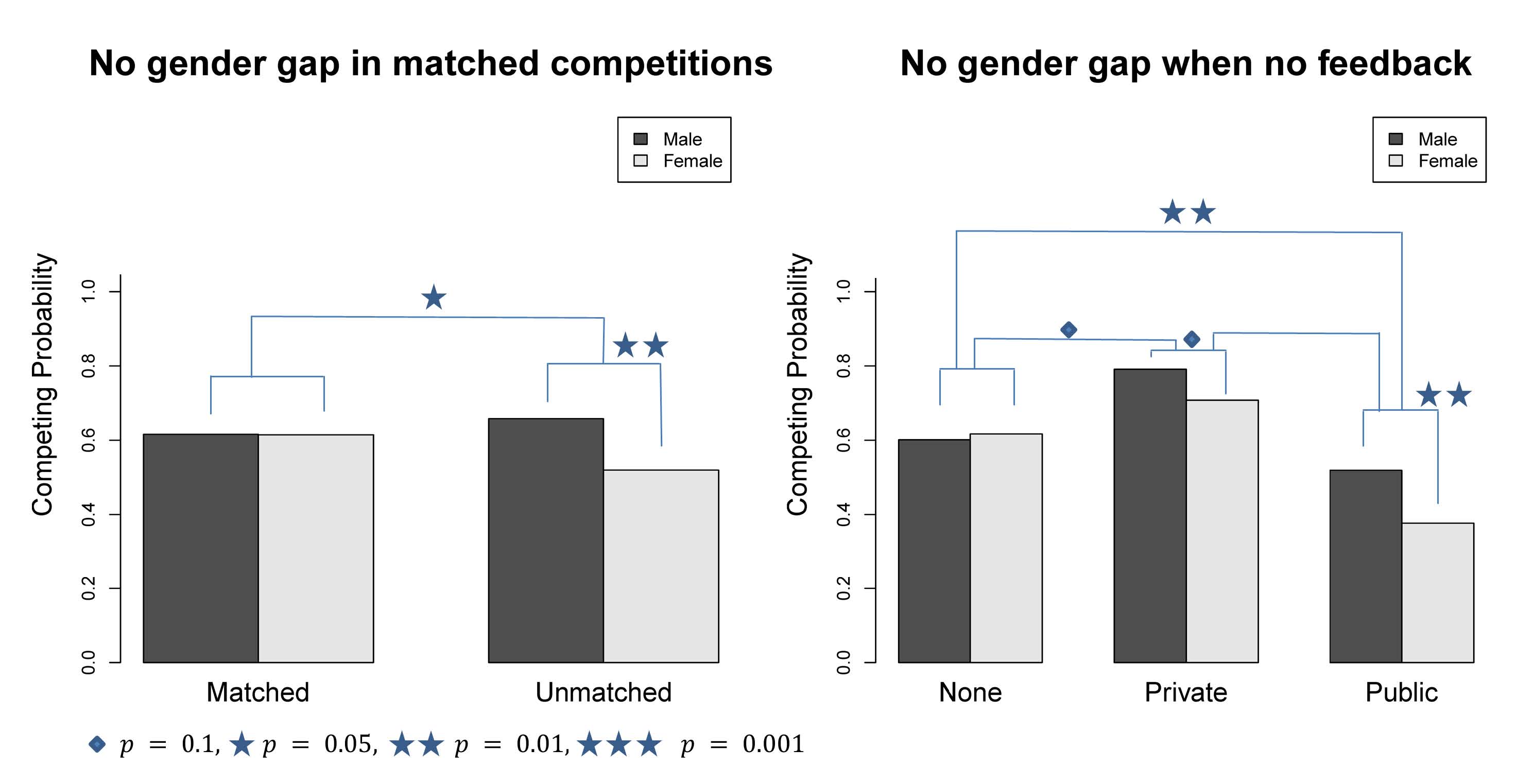


Figure 3: Preferences for feedback and ambiguity in men and women. The left panel gives gender differences in entrance to ambiguous versus unambiguous contests -i.e. columns of Figure 1*ii*. The right panel gives gender differences in the preference for private feedback and public feedback, relative to uninformative contests, respectively.

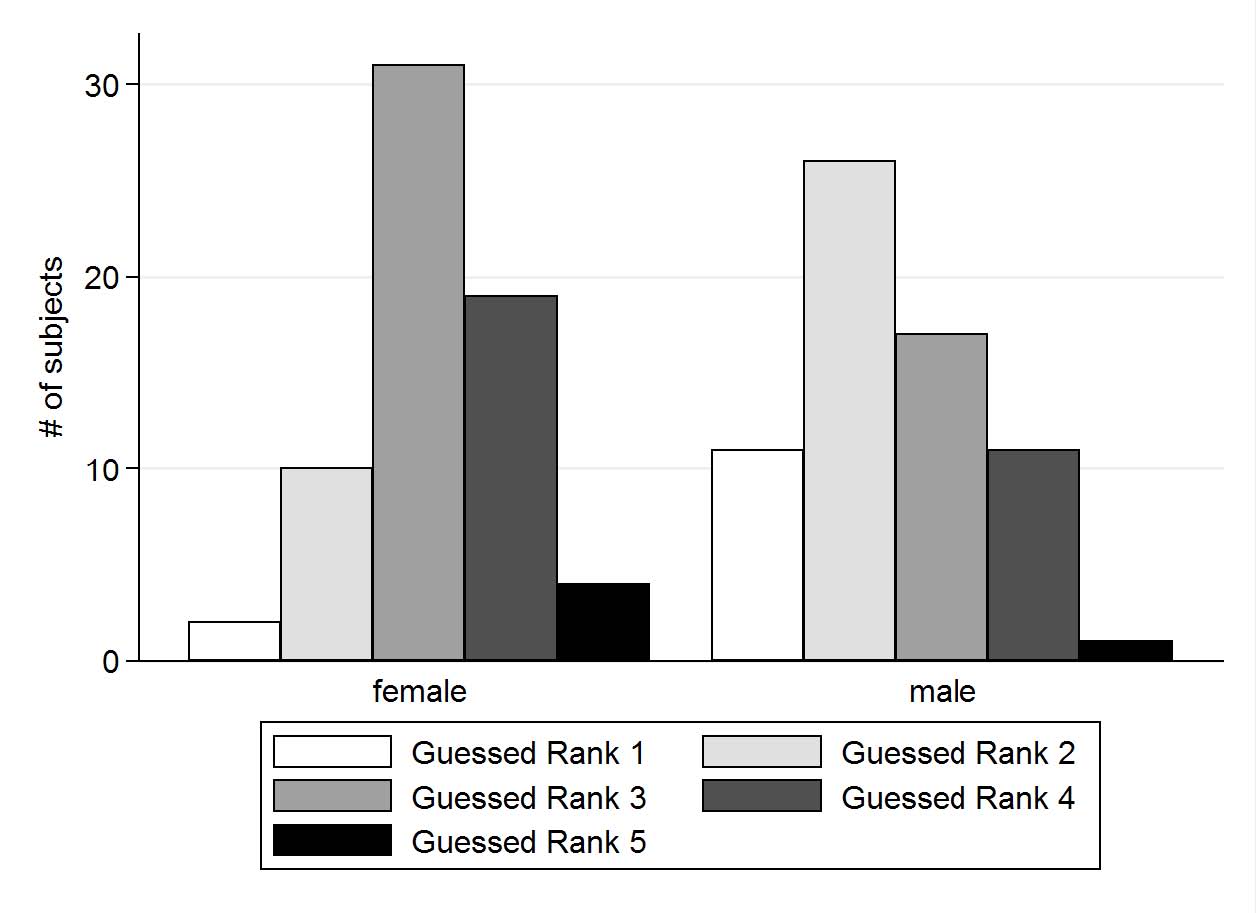


Figure 4: Guessed rank in an ambiguous contest.

## 3.2. Gender differences in confidence or risk-taking

Women and men had a comparable preference for (non-competitive) risk, as concluded from an insignificant effect of sex on gambling choice using logistic regression. Men however had higher confidence faced with ambiguous contests (see 5), as concluded from an ordered logistic regression with a dummy for sex: men predicted significantly higher rank for themselves (95% CI ).

On average, men predicted their rank to be 2.47. A t-test revealed that this is significantly different from 3 (). This means men expected to perform above average. In contrast, women on average predicted themselves to rank 3.2. This is significantly greater than 3, thus they expected to perform beneath average ().

## 3.3. Gender differences in confidence or risk-taking: do these underpin competitive preferences?

We wondered whether gender-specific preferences for ambiguity and feedback are explained away by gender differences in risk preferences, ambiguity preferences or confidence (self-reported prediction about rank). We estimated nested sub models of Equation 3, including the vector of non-competitive risk/ambiguity preferences or competitive confidence . The sub models arise simply by setting the zero sub-matrices of in Equation 3, a fact which permits us to use a single subscripts notation to identify each of the effects below, e.g. always indicates the impact of belief on preferences for a competitive ambiguity.

Subjects’ risk preference significantly predicted their preference for competitive ambiguity (, ,), but gender differences in ambiguity aversion remained significant in this model (, ,). Subjects confidence also significantly predicted a preference for competitive ambiguity (,,): however including explained away gender differences in preference for competitive ambiguity. There was no residual gender difference after accounting for .

As discussed above, we quantified the more coarse measure of *preference for feedback.* In the previous section we showed that women were more feedback averse, defined in this manner. This gender difference remained significant after accounting for non-competitive risk/ambiguity preferences, i.e. after including the vector in the model[[2]](#footnote-2) (,,). Conversely, this gender difference in the preference for feedback was not significant after accounting for confidence . Instead, confidence itself significantly predicted lower feedback aversion (,,) and a greater preference for competitive ambiguity (,,). In summary, non-competitive risk and ambiguity preferences did not explain gender differences in competitive behaviour. In contrast, gender differences appear to reflect gender-specific confidence.

The preceding analyses show that confidence significantly explains gender differences in preference for prior and posterior information. To establish whether confidence is *necessary*, we looked for gender differences in matched contests, in which confidence - beliefs about relative ability - should play no role: by experimental design both male and female subjects should expect to win with probability . We did this by estimating a reduced form of Equation 2 on only data from the unmatched contests: This reduced model lacked ambiguity parameters and corresponding sub matrices of in Equation 3. The first analysis showed that women retained a stronger preference against feedback than men, even when experimentally equating their beliefs about relative ability in this way (,,). Thus while confidence matters, it may not be the only factor influencing gender differences in preference for feedback.

## 3.4. No gender differences in performance

As Gneezy, Niederle, & Rustichini (2003) found, competition itself may create gender differences in performance (in mixed-gender experiments). Therefore, we aimed for a gender-neutral task and selected the N-back.

A t-test revealed no significant difference in the male versus female N-back performance (0.856 and 0.842 respectively). Furthermore, there was no gender stereotype associated with the N-back in our sample. We asked participants if they thought the task men or women perform better on the task in general, or whether there were no gender differences. The majority thought it was gender neutral, and the few opting for one gender compensate each other. This is especially different to the usually used math tasks (see for example Niederle & Vesterlund (2007), which have a gender-stereotype favoring men.

## 3.5. No detectible effect of menstrual cycle

Some previous work points to a role for female hormones in competitiveness. Post-menopausal women compete as much as men   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(J. A. Flory, Leonard, & List, 2011; J. Flory, Gneezy, Leonard, & List, 2012) and pre-menopausal women compete differently throughout the menstrual cycle, though the pattern is not entirely consistent   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Buser, 2012; Wozniak, 2009). We therefore divided women into those in the low ( versus high ( hormone phase and examined differences in competitive behavior between these groups. Specifically, from the date of their last menstruation and the expected date of their next menstruation   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Chen, Katušcák, & Ozdenoren, 2012) we calculated their position in a normalized menstrual cycle (normed to 28 days to accommodate individual differences in cycle length). All females using no hormonal contraceptives are assigned to the high hormone group when they are between day 11 and 26 of their (normed) menstrual cycle. Females using hormonal contraceptives between day 8 and 28 were assigned to the high hormone group. We found no differences in competitive behavior between high and low hormone individuals in any of the six conditions and no differences in preferences for ambiguity, private or public feedback.

# 4. Discussion

Contests have consequences for social reputation, self-esteem and material wealth. The first two consequences derive from posterior information about the contest’s outcome *per se*: public or private feedback about one’s status relative to the group. Prior information is also critical when deciding whether to enter a contest. We looked at the gender differences in these preferences for prior and posterior information. We show the gender gap decreases when subjects are given prior information, but increases specifically when they must attain it through voluntary participation in informative contests. The prospect of feedback triggers avoidant social behavior, particularly in women, beyond that explained by other potentially aversive aspects of a contest, such as the size of potential material losses and risk or ambiguity preferences.

Why do men and women differ in these preferences? Like   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
Kamas & Preston (2012) and Niederle & Vesterlund (2007), we found that men were overconfident and as like   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
Cotton et al. (2010) we found that women were underconfident. These differences largely explain their relative preferences for ability-matched and unmatched contests. These differences in confidence are also an important explanation for gender differences in preference for feedback. In principle, gender differences in confidence may be reduced by reducing women’s under-confidence or reducing men’s over-confidence.   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
Niederle & Vesterlund (2008) and   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
Balafoutas & Sutter (2010) suggest quota-like restrictions or preferential treatment in competitions to encourage women. Alternatively,   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
Wozniak (2009) found that men’s overconfidence and women’s underconfidence reduce by giving them some feedback about their relative performance before the competition. However we further found that gender differences in preference for feedback may not only be driven by confidence differences: they persist even when male and female confidence are experimentally equated, i.e. in the matched contests depicted in column 1 of Figure 1*ii*, i.e. even under this condition of minimal ambiguity, women have a stronger preference against feedback. This suggests that the aversion to public feedback does not stem exclusively from underconfidence, i.e. pessimism about ones winning probability. Another source for the gender gap according to   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
Buser (2012) and Wozniak (2009) is the influence of the menstrual cycle on women. In contrast to them we did not find hormone phase to predict willingness to compete.

# 5. Acknowledgements

We thank Sebastian Stephan. We used Cogent 2000 and Cogent Graphics FIL/ICN/LON (John Romaya).

**References**

Balafoutas, L., & Sutter, M. (2010). Gender, competition and the efficiency of policy interventions.

Borghans, L., Heckman, J. J., Golsteyn, B. H., & Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, *7*(2-3), 649–658.

Buser, T. (2012). The impact of the menstrual cycle and hormonal contraceptives on competitiveness. *Journal of Economic Behavior & Organization*, *83*(1), 1–10.

Cadsby, C. B., Servatka, M., & Song, F. (2011). How Competitive are Female Professionals? A Tale of Identity Conflict. *A Tale of Identity Conflict*.

Chen, Y., Katušcák, P., & Ozdenoren, E. (2012). Why can’t a woman bid more like a man? *Games and Economic Behavior*.

Chumbley, J., Flandin, G., Bach, D., Daunizeau, J., Fehr, E., Dolan, R., & Friston, K. (2012). Learning and Generalization under Ambiguity: An fMRI Study. *PLoS Computational Biology*, *8*(1).

Cotton, C., McIntyre, F., & Price, J. (2010). *Causes of Gender Differences in Competition: Theory and Evidence*.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 448–474.

Datta Gupta, N., Poulsen, A., & Villeval, M.-C. (2005). Male and female competitive behavior-experimental evidence. *GATE Working Paper No. WP 05-12*.

Dweck, C. (2000). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of experimental economics results*, *1*, 1061–1073.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 643–669.

Ertac, S., & Szentes, B. (2011). The effect of information on gender differences in competitiveness: experimental evidence. TÜSAD-Koç University Economic Research Forum working paper series.

Flory, J. A., Leonard, K. L., & List, J. (2011). Outgrowing the Gender Gap in Competitiveness.

Flory, J., Gneezy, U., Leonard, K., & List, J. (2012). *Sex, competitiveness, and investment in offspring: On the origin of preferences*. The Field Experiments Website.

Gneezy, U., Niederle, M., & Rustichini, A. (2003a). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, *118*(3), 1049–1074.

Gneezy, U., & Pietrasz, A. (2013). When Half of the Men Are More Competitive than All Women.

Große, N. D., & Riener, G. (2010). *Explaining gender differences in competitiveness: gender-task stereotypes*. Jena economic research papers.

Hartig, J., & Moosbrugger, H. (2003). Die “ARES-Skalen” zur Erfassung der individuellen BIS-und BAS-Sensitivität. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *24*(4), 293–310.

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning-implications for training and transfer. *Intelligence*, *38*(6), 625–635.

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

Kleinjans, K. J. (2009). Do gender differences in preferences for competition matter for occupational expectations? *Journal of Economic Psychology*, *30*(5), 701–710.

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

Niederle, M., & Vesterlund, L. (2008). Gender differences in competition. *Negotiation Journal*, *24*(4), 447–463.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, *25*(1), 46–59.

Roberts, T.-A., & Nolen-Hoeksema, S. (1989). Sex differences in reactions to evaluative feedback. *Sex Roles*, *21*(11-12), 725–747.

Schubert, R., Brown, M., Gysler, M., & Brachinger, H. W. (2000). *Gender specific attitudes towards risk and ambiguity: An experimental investigation*. Institut für Wirtschaftsforschung, Eidgenössische Technische Hochschule.

Shurchkov, O. (2012). Under pressure: Gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*(5), 1189–1213.

Wirth, L. (2009). *Breaking through the glass ceiling: Women in management*. International Labour Organization (ILO).

Wozniak, D. (2009). Choices About Competition: Differences by gender and hormonal fluctuations, and the role of relative performance feedback.

# 1. Supplementary Material

## 1.1. Consent

## 1.2. N-Back test

In the *N-Back task*, subjects saw a stream of stimuli and responded whenever a stimulus was the same as N steps before. Specifically, subjects were shown a random sequence drawn from a pool of eight different letters (A-H). Each letter was shown for 500ms with an interval of 2500ms between two letters. Subjects were required to report ’target’ letters by pressing the space key in the 3000ms interval before the next letter was shown. If the letter shown was not a target, a non-response was required. Subjects first performed a 1-back test to ensure they understood the task. In three blocks of increasing difficulty, N was 2, 3 and 4. Each block contained 20 stimuli and 4 targets, resembling the parameters of   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Jaeggi et al., 2010; Owen, McMillan, Laird, & Bullmore, 2005). The whole N back task lasted approximately 15 minutes. A subject’s score was computed as the average of the true positive and true negative rate. Obtained scores lied between 60% and 100%.

## 1.3. Instructions (forced)

## 1.4. Instructions (free choice)

## 1.5. Instructions guessed rank and choice of envelope

## 

## 1.6. Questionnaires

### General information

The first questions asked subjects’ age, gender, nationality, study field, highest level of education completed and disposable income, followed by questions on their smoking and drinking behavior and about working hours during nights or shifts. We asked women about their menstrual cycle and hormonal contraceptives usage.

### Risk preferences

To measure self-reported risk preferences, we used a test from   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Datta Gupta et al., 2005). Subjects answered 16 questions concerning the likelihood to engage in risky activities. Summing up the answers resulted in a score for each subject, in which a higher score meant the subject was more risk loving.

The risk preferences differ among subjects and may therefore explain some between-subject variance in our task (see   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Eckel & Grossman, 2008; Niederle & Vesterlund, 2007)).

### Action Regulating Emotion Systems (ARES)

ARES captures the sensitivities of two emotion systems, the BIS and BAS sensitivities. BIS I looks at the anxiety/nervousness and BIS II at frustration/sadness. BAS I contains questions evaluating the drive behind goal-directed behavior. BAS II measures the responses toreward attainment. BIS captures punishment sensitivity and BAS, reward sensitivity. We used the short version of the ARES-scales which contains 20 items from   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
(Hartig & Moosbrugger, 2003). We translated the German version and if subjects a choice between English and German. The english version is given below.

**Instructions:** The following questionnaire consists of a set of statements, which describe oneself. Each statement may apply to you to a different extent ("Disagree", "Rather Disagree", "Rather Agree", and "Agree"). To fill in the questionnaire, please place a cross in the corresponding box. Please answer every statement. Should you feel unsure what to indicate, cross the one answer that suits you the most.

* It gives me a kick when things work out as planned.
* I get frustrated pretty easily if something doesn’t work out as I had hoped.
* Even small things make me really happy.
* I become flustered quickly when I realize that I did something wrong.
* I sadden quickly if I don’t attain a goal I was aiming for.
* You can barely stop me once I set myself a goal.
* If someone criticizes me, I get insecure and nervous.
* Even little everyday mishaps can really frustrate me.
* The prospect of success energizes me.
* If I feel that what I am doing goes wrong, I get anxious and insecure quickly.
* It stimulates me when I feel that I can attain a personal goal.
* I am easily delighted.
* I can get really sad if something doesn’t work out the way I wanted it to.
* Even little mishaps unsettle me considerably.
* It makes me very happy to achieve a goal I strove for.
* I get rather seldom really excited about something.
* If I do something wrong, I immediately fear the consequences.
* Even little mishaps really disappoint me.
* I rarely get excited, even when I get something that I really wanted.
* Even little incentives can strongly motivate me.

## 1.7. Equivalent N-back performance in competitive versus non-competitive contexts

We used an assumption that N-back performance is equivalent between competitive and non-competitive contexts. This assumption was used to evaluate subjects’ relative performance in the final N-back contest within groups composed irrespective of the competitive context (i.e. independently of whether group members chose to compete or not). On average, the subjects solved 81.90 % of the task correctly when they were paid according to their absolute performance and 82.33 % when they were paid according to their relative performance. In support of our assumption that N-back ability is stable across contexts, this difference was not statistically significant ().

**Supplementary References**

Josephs, R. A., R. P. Larrick, C. M. Steele and R. E. Nisbett (1992). "Protecting the self from the negative consequences of risky decisions." Journal of personality and social psychology **62**(1): 26.

Maslow, A. H. (1963). "The need to know and the fear of knowing." The Journal of General Psychology **68**(1): 111-125.

1. Here indicates that this parameter parameter comes from the re-specified model (with the different reference contest). [↑](#footnote-ref-1)
2. Here indicates that this parameter parameter comes from the re-specified model. [↑](#footnote-ref-2)