Supplementary Material

Competition and Gender Inequality:

A comprehensive analysis of effects and mechanisms

Klarita Gërxhani, Jordi Brandts, Arthur Schram

January, 2023

Appendix A: Experimental Designs and Procedures

In this appendix, we provide a detailed description of the experimental design and procedures, as well as the power analysis that guided our choice of sample size. We do so for each of the studies in turn. This project was started in 2017, before pre-registration became the dominant norm for experimental papers in the social sciences. For this reason, the experiments were not pre-registered. Note, however, that this paper and our experimental designs follow naturally from our previous paper (Schram et al. 2019). The experiment in Study 1 is designed in a way that consistently follows from the original paper and all other experiments follow from Study 1. Moreover, the tests we use (predominantly Fisher/Pitman permutation tests) are the same as used there.

STUDY 1

Experimental Design

For Study 1 we conducted sessions with six treatments involving performance on a cognitive task under different combinations of rivalry and ranking. In each of the six treatments there are two parts which involve three types of participants: *A*s, *B*s and *C*s. In most sessions there are six *A*s, six *B*s and one *C*.[1] Instructions for each type are presented in Appendix B. In part 1, *A*s and *B*s work individually on a cognitive task and *C*s are passive. Treatments vary in the monetary incentives *A*s and *B*s face and in the information they and others receive related to their status ranking. This is explained below.

The cognitive task in part 1 is the same in all sessions and is taken from Weber and Schram (2017). Participants are presented with a sequence of pairs of 10x10 matrices filled with two-digit numbers. These matrices appear at the lower half of their computer monitor (Figure A1). For each pair of matrices each participant is asked to individually search to find the highest number in the left matrix and the highest number in the right matrix and to calculate the sum of these two numbers. This sum must be entered in the window at the center-top of the monitor.[2] After a number has been entered, two new matrices appear, regardless of whether the sum was correct or not. The task continues for 15 minutes. At any time during work on the task each participant can see on the screen their own cumulated number of correct summations as well as the remaining time. Participants are not informed about the performance of any other participant.

The instructions emphasize the importance of doing well in this task by mentioning that it has been shown to correlate positively with success in professional life.[3] Participants were told that we would provide evidence of this claim upon request after the experiment.[4]

---

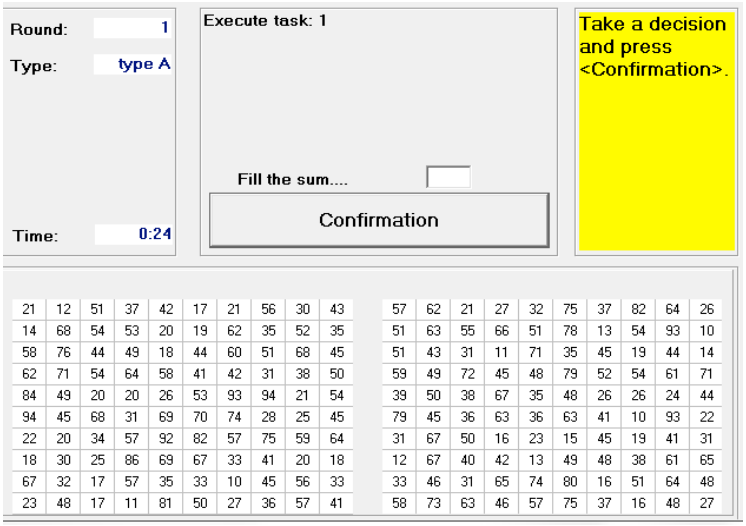
[1] As will be explained below, there were also sessions that did not require A or C participants; in these sessions, there were two groups of six *B*s. The lower panel of Table 1 summarizes the participant types.

[2] As described in the note to Table 2 in the main text, we treated any participant who repeatedly entered numbers larger than 200 as outliers. There were only four such outliers (less than 1% of the participants).

[3] This emphasis was made to stress the importance of status ranking based on the performance in the particular task we used. Additional tests reported in Schram et al. (2019) show that this priming does not induce stereotype threat. In fact, excluding it does not significantly affect results. Nevertheless, we include it here to maintain consistency across related projects.

[4] For this purpose, we had available copies of Koedel and Tyhurst (2012), a résumé study linking math skills to labor market outcomes that provides such evidence.

Figure A1: Screenshot Part 1



Notes. As explained in Schram et al. (2019), the instructions inform participants that numbers were 'randomly generated'. Drawing from a uniform distribution would give a high probability of very high sums. To avoid this, for each cell, we first drew a random number between 40 and 99, say X. Then, we drew a random number (uniformly) between 10 and X. This gives a far lower probability of high numbers (the chance of a number being 75 or more is approximately 0.06).

The way in which participants are rewarded for task performance is one of our treatment variables. In a similar vein to Gneezy et al. (2003), we use two different payment schemes. The first is an individual piece-rate payment, with each correct answer yielding €1. The second is a tournament payment scheme, where only the two participants with the highest score in a group of six receive payment for each correct answer, while the other four receive nothing. To keep average payment constant across treatments, the payment per correct answer with tournament payment is three times that of piece-rate payment: €3. The idea underlying this treatment variation is that the tournament payoff creates a rivalry for resources, while the piece rate does not. For this reason, we use the acronyms *nRfR* (no rivalry for resources) and *RfR* (rivalry for resources) for the piece-rate and tournament incentive treatments, respectively.

Our second treatment variation is used to study the status-ranking dimension of competition, building on the design of Schram et al. (2019). This is applied in part 2, where we vary whether or not participants receive ranking feedback. One group of *B* players receives no such feedback; we call this the no-ranking treatment (*nR*). This means that these *B* players are passive in part 2; the *C* player also remains passive in part 2. Other players do receive feedback, which may be one of two types. Recall that status ranking has two distinct characteristics. It informs an individual of her own ranking vis-à-vis others and it informs others of her ranking. By varying the feedback participants receive, we isolate the former. This allows us to differentiate between the effects the two characteristics might have. In the first type of feedback participants are only informed about their own ranking. In particular, in some sessions, the B-players are privately given this information. We call this the private-ranking treatment, *PR*. The

second type of feedback is provided in the social-status ranking treatment (*SR*), which involves the following. After finishing the instructions, each *A* player is individually taken to the *C* player, who does not take part in the real-effort task and whose task consists in listening. The *A* player reads aloud a text stating that she will return after the task to report her performance score. This is done to create the anticipation of having to later report to the *C* player. After finishing the task, the *A* player is informed of her score and rank amongst the group of six *A* players. She is then individually taken to the *C* player a second time, where she reads a text stating her score and rank. The *C* player, a peer, is the same for all *A*s in a session, so that this person will end up knowing the rank of each of the participants in the real-effort task. In all cases, the ranking condition is common information. Importantly, both private and social-status ranking information consist in knowing one's own position in the ranking, but not the complete ranking of all relevant participants. The only person who has this complete knowledge is the *C* player.

We crossed tournament pay and piece-rate pay with the three ranking treatments in a full-factorial design, yielding the total of six treatments. Table A1 provides an overview of these treatments and of the player types (A, B, and C).

Procedures

Study 1 was run at the BLESS laboratory of the University of Bologna, Italy in 34 sessions with 13 participants each in the period November 16 - December 7, 2017.[5] 408 active (A and B players) participants (207 men, 201 women) and 24 passive C players were recruited from the BLESS participant pool for sessions that lasted 60-75 minutes.[6] Average earnings were approximately €18.

Table A1: Overview of treatments and participant types

Treatment	Acronym	Participant type	Information	Payoff
no rivalry for resources, no ranking,	nRfR/nR	B	none	piece rate
no rivalry for resources, private ranking,	nRfR/PR	B	rank	piece rate
no rivalry for resources , social-status ranking	nRfR/SR	A	rank	piece rate
rivalry for resources, no ranking,	RfR/nR	B	none	Tournament
rivalry for resources, private ranking,	RfR/PR	B	rank	Tournament
rivalry for resources , social-status ranking	RfR/SR	A	rank	Tournament
Participant types				
Type A	In sessions with social-status ranking; do the task; report score to the C-player			
Type B	In all sessions; do the task; report to no one			
Type C	In sessions with social-status ranking; do not do task; only hear reports by A-players			

Notes. 'Information' denotes whether a participant is told her rank within the group. In 'piece-rate' payoffs, every participant is rewarded for the own score. In 'tournament' only the top two performers in a group are rewarded.

Power

Based on the social-status ranking treatment in Schram et al. (2019), a power of 80% and significance level 5% would require a sample of 12 men and 12 women if means are tested with a t-test. We have more treatments than just *SR*, of course, and therefore need more observations. On the other hand, unless indicated otherwise, we use

[5] We thank Marco Casari for making the laboratory available and Lorenzo Golinelli, Mario Spiezio, and Giorgio Monti for their assistance in organizing the experiments.

[6] To avoid signaling our interest in gender differences, we did not recruit separately per gender. For this reason, the numbers show some random fluctuation, but the minimum of 25 observations per gender was realized in all cases.

permutation t-tests using Monte-Carlo resampling with 10,000 repetitions (henceforth, PtT) to compare means. This non-parametric Fisher-Pitman test has substantially higher power than standard t-tests (cf. Appendix C in Schram et al. 2019) and therefore requires fewer observations. To the best of our knowledge however, there is no standard tool available to do a precise power analysis for these tests. In the end, we decided to stay on the safe side of these countervailing considerations and aimed at a minimum of 25 men and 25 women per treatment. Numbers of observations per treatment are presented in Table 2 of the main text.

STUDY 2

Experimental Design

As explained in the main text, we elicited beliefs about gender differences in performance in the absence of any competitive setting (*nRfR/nR*), under (only) social-status ranking (*nRfR/SR*), and under (only) rivalry for resources (*RfR/nR*).

Each session consists of four parts. In the first, participants did the summation task with piece-rate remuneration of €0,50 per correct answer. This allowed them to get acquainted with the task. In the following three parts, we explained that previous participants (i.e., those of Study 1) (i) had done the same task for €1,00 per correct answer; or (ii) had done the same task for €1,00 per correct answer and then had to report their rank to a peer; or (iii) had done the same task and received €3,00 per correct answer (only) if they were in the top two in a group of six.[7] In each part, they were asked to predict whether the mean score of men was better than that of women or vice versa. We excluded the possibility of predicting exactly equal mean scores because this is an event with extremely low probability. One of these three parts was randomly chosen at the end of the experiment and every participant that had predicted correctly in that part received an additional payoff of €5.00.

Procedures

The sessions were run on December 10-11, 2019 at the BLESS laboratory in Bologna, Italy. We recruited 96 participants (48 men, 48 women) for four sessions that lasted 50-60 minutes each. Average earnings were €17.82.

Power

Even though our hypothesis is directional (i.e., people believe that men perform better under competition than women do), we base the power analysis and the statistical analysis in the main text on two-sided tests. The test here is a binomial test of whether the fraction of a population that believes that men and women perform differently is 0.5. To detect an actual proportion of 0.3 with power 80% and significance level 5%, we needed 49 men and 49 women. The realized 48 participants per cell give a power of 0.79 (calculated on http://powerandsamplesize.com/Calculators/Test-1-Proportion/1-Sample-Equality).

STUDY 3

Experimental Design

To implement a treatment where competitive success has no negative impact on others, we focus on the private ranking environment without rivalry, *nRfR*/*PR*. A priori, there is

[7] To avoid order effects, each of the four sessions had a different order of (i), (ii), (iii) in parts 2-4. The order does not affect the reported beliefs, so we pool the data in our analysis. More details are available upon request.

no reason to believe that a warmth stereotype would play a different role in distinct dimensions of competition. In all cases, one's success is at the expense of another's failure. We choose to investigate the role of warmth stereotype in *nRfR/PR*, because this is a competitive environment where an alternative without a negative impact on others is easiest to implement. With private ranking but without rivalry, earning a high rank by definition makes other participants be ranked lower than they would have otherwise been. To circumvent this impact on others, we created a new treatment where this ranking is not vis-à-vis others in the same session. Instead, to determine a private rank, we randomly selected for each participant five other participants from previous *nRfR/PR* sessions and anonymously ranked her performance in relation to theirs. Importantly, these others were not informed about this ranking. We denote this new treatment as *private historic ranking* (*nRfR/PHR*).

Procedures

For this new treatment, we recruited 65 participants (34 men, 31 women) for three sessions on Feb. 19, 2019 at BLESS in Bologna, Italy, that lasted approximately 60 minutes. Participants earned on average €16.50.

Power

For this study, the same power analysis holds as for Study 1, where we aimed at 25 men and 25 women. Because the lack of a warmth stereotype may diminish gender differences in performance, we opted to increase the number of participants to at least 30 each.

STUDY 4

Experimental Design

To test whether the gender composition affects behavior under competition, we ran treatments with either only private ranking (*nRfR/PR*), only social-status ranking (*nRfR/SR*) or rivalry for resources with private ranking (*RfR/PR*). We decided to add information about private ranking to the *RfR* because rivalry always involves some information about private ranking (one learns whether or not one is amongst the top two in a group of six). As a consequence of adding *PR* to *RfR*, we also included the *nRfR/PR* treatment. A comparison of gender composition effects in *RfR/PR* and *nRfR/PR* then allows us to establish the effect of rivalry alone.

Procedures

We organized 16 additional sessions at BLESS (Bologna, Italy), eight of which were all men, and the other eight only women.[8] These were run between Nov. 28 and Dec. 10, 2019. Average earnings were €17.10.

Power

[8] For these sessions, we were careful not to reveal that we were recruiting only (wo)men for a session. To avoid signaling our research question, we also did not stress the same-gender setting. Nevertheless, the gender composition was very obvious, in a natural way, when participants were waiting to be invited into the laboratory, when registering and signing the consent forms, and when being seated. Moreover, A-players (those subjected to social-status ranking) could infer the gender of others when these others were taken one by one to visit the C-player (before performing the task). Acknowledging that we cannot know with certainty whether participants noticed that they were in same-gender sessions, we decided that the disadvantage of them perhaps not noticing was smaller than the disadvantage of possibly inducing experimenter-demand effects.

Once again, the same power analysis holds as in Study 1. We therefore aimed at 25 men and 25 women per treatment cell. The numbers of participants varied between 29 and 36 per cell.

Appendix B: Experimental Instructions (Translated into Italian for the Experiment)

I. No Rivalry For Resources

I.1 A-Players (Status Ranking)

Part 1

Welcome to this experiment.

You will receive 5 euro for your participation in today's experiment. Depending on your decisions and the decisions of other participants in today's experiment, you may earn more money. You will be paid privately at the end of the experiment. In the experiment you will remain anonymous. Your decisions will only be linked to your station id and not to your name in any way. Neither the other participants, nor the experimenter will be able to see how much you personally earn. The experiment will take approximately 1 hour.

The participants in this experiment have been randomly divided into three groups. Six of you are of type A, six are of type B and one is of type C. The participants of type A and of type B are in this room, while the participant of type C is in a different room.

You are of type A.

The experiment is divided into two stages. You will receive instructions for each stage when it starts. We guarantee that everything we tell you in these instructions will proceed precisely as described. If you have any doubts about whether we are acting in the way described in the instructions, we will be happy to show you at the end of the experiment that this is the case.

We now start with stage 1.

In stage 1 the participants of type A and of type B will all independently perform a task during 15 minutes.

This is an important task that is often used to measure people's talents. Many scientific studies have found that people who do well in a task like this are more successful in professional life than people who do less well. You will not be told, however, what is typically a good or a bad score for this task.

The task is as follows. You will see two matrices on the computer screen. Each matrix has 10 rows and 10 columns and is filled with randomly generated numbers. Your job is to find the largest number in each of the matrices and then to add them up. You are not allowed to use calculators, but you can use the paper and pencil that you have found on your desk.

After entering a sum the computer will tell you whether it is correct or incorrect (please note that the time will continue to run while you see this result). Subsequently, irrespective of whether your answer is correct or incorrect, a new pair of matrices will appear. This means that for each pair, you have only one attempt to provide the correct answer. However, there will be new matrices as long as you are within the 15 minutes limit. We will allow a maximum of 50 attempts. Any matrices you try to solve after the 50^{th} will not count towards your earnings.

After the task has been completed, we will inform you about your score and how your score ranks amongst the six participants of type A. To do this, the experimenter at the computer server will compare the score of all six participants of type A. To ensure your anonymity, this experimenter is seated in a way that he cannot identify you. He can link your score to your station identification, but cannot see who is sitting at each particular station.

For each correct sum you will receive 1 euro and for each incorrect sum you will receive 0 euros. The total number of euros you have gained will be visible on the screen at the end of this stage.

Remember that studies have found that people who do well in a task like this are more successful in professional life. You will not know how people typically perform in this task. Nor will you know how other participants scored in the task today.

After this stage, you will be asked to inform the participant of type C about your score. You will tell her your score and your rank amongst the participants of type A. Note that each of the A players will report to the same C player.

You will have to go to a separate room where this participant will be waiting. This participant does not know what task you did and what the score means. He or she has only been told that a higher score is thought to lead to a more successful professional life. Importantly, this participant will hear your score and how it ranks to the other A participants. Only the participant of type C will be able to make this comparison.

As explained, the experimenter at the computer server will not be able to link scores to individuals. He will prepare closed envelopes that will be distributed by a different experimenter. The first set of envelopes will only be distributed to participants of type A. This envelop will contain your score, your rank, and the text you will need to read to the participant of type C. The other experimenter will not know the contents of the envelop he is handing to you.

After all participants of type A have reported to the participant of type C, the experimenter at the computer server will prepare envelopes for all participants. These include your earnings and a receipt for you to sign. Once again, the experimenter handing out the envelopes does not know its contents, so your earnings in today's experiment remain anonymous.

We would like you to see the type C participant that you will present your score and rank to. For this reason, each of you will now first leave the room and read aloud –to the C participant– a text that will be given to you. Note that each of you will be going to the same C participant.

Part 2
This brings us to the end of the first stage of the experiment. The experimenter at the computer server will now prepare the envelopes with information about your score and rank amongst the participants of type A. In the meantime, we ask you to fill out a brief questionnaire.

Now you will be asked to inform the participant of type C about your score and your rank. Recall that every type A player is visiting the same participant. Each participant of type A will be accompanied by the experimenter to the room in which the participant of type C is waiting. There, each A participant will open the envelop and read aloud the text to the C participant. Then all A participants will return to their desks. Participants of type B will remain seated during this process and are not informed about what the A participants will do.

Remember that the C participant does not know what task you did and what the score means. He or she has only been told that a higher score is thought to lead to a more successful professional life. Importantly, this participant will know your score and how it ranks to the other A participants.

I.2 B-Players (No Status Ranking)

Part 1

Welcome to this experiment.

You will receive 5 euro for your participation in today's experiment. Depending on your decisions and the decisions of other participants in today's experiment, you may earn more money. You will be paid privately at the end of the experiment. In the experiment you will remain anonymous. Your decisions will only be linked to your station id and not to your name in any way. Neither the other participants, nor the experimenter will be able to see how much you personally earn. The experiment will take approximately 1 hour.

The participants in this experiment have been randomly divided into three groups. Six of you are of type A, six are of type B and one is of type C. The participants of type A and of type B are in this room, while the participant of type C is in a different room.

You are of type B.

The experiment is divided into two stages. You will receive instructions for each stage when it starts. We guarantee that everything we tell you in these instructions will proceed precisely as described. If you have any doubts about whether we are acting in the way described in the instructions, we will be happy to show you at the end of the experiment that this is the case.

We now start with stage 1.

In stage 1 the participants of type A and of type B will all independently perform a task during 15 minutes.

This is an important task that is often used to measure people's talents. Many scientific studies have found that people who do well in a task like this are more successful in professional life than people who do less well. You will not be told, however, what is typically a good or a bad score for this task.

The task is as follows. You will see two matrices on the computer screen. Each matrix has 10 rows and 10 columns and is filled with randomly generated numbers. Your job is to find the largest number in each of the matrices and then to add them up. You are not allowed to use calculators, but you can use the paper and pencil that you have found on your desk.

After entering a sum the computer will tell you whether it is correct or incorrect (please note that the time will continue to run while you see this result). Subsequently, irrespective of whether your answer is correct or incorrect, a new pair of matrices will appear. This means that for each pair, you have only one attempt to provide the correct answer. However, there will be new matrices as long as you are within the 15 minutes limit. We will allow a maximum of 50 attempts. Any matrices you try to solve after the 50th will not count towards your earnings.

After the task has been completed, we will inform you about your score [*Only in Private-Ranking Treatment:* and how your score ranks amongst the six participants of type B. To do this, the experimenter at the computer server will compare the score of all six participants of type B]. To ensure your anonymity, this experimenter is seated in a way that he cannot identify you. He can link your score to your station identification, but cannot see who is sitting at each particular station.

For each correct sum you will receive 1 euro and for each incorrect sum you will receive 0 euros. The total number of euros you have gained will be visible on the screen at the end of this stage.

Remember that studies have found that people who do well in a task like this are more successful in professional life. You will not know how people typically perform in this task. Nor will you know how other participants scored in the task today.

The type A participants will now leave the room, each for a few minutes. Please remain seated quietly until this has been completed.

Part 2
This brings us to the end of the first stage of the experiment.

In this stage all B participants are required to remain quietly seated at their desks and wait till the end of stage 2.

I.3 C-players

Welcome to this experiment.

Your role in today's experiment is a passive one. You will not be asked to make any decisions.

Your only task is to hear the results of a task performed by one of the other participants. You will not be informed about the content of this task. All you need to know is that it is an important task that is often used to measure people's talents. Many scientific studies have found that people who have a high score in a task like this are more successful in professional life than people who have a low score. You will not be told, however, what is typically a good or a bad score for this task.

You will be seated in this room. While you are waiting for the other participants, feel free to read anything you like, or to browse your phone. Before the other participants have started their task six of them will be taken to you, one at a time. Each will read a text to you and then return to the laboratory. After all the other participants have completed their tasks, the same six will come here again, one at a time. They will read to you their score and how this ranks amongst the six participants.

You may not speak or interact with the other participants in any way.

For your role in today's experiment, you will receive 10 euros on top of the 5 euros show-up fee. You will be paid and dismissed after the participants have told each of you their scores.

Finally, please treat the room you will be in with respect. You are a guest here, so please do not touch anything that is not yours.

I.4 Texts

Text to read to C-players by A-players before the summation task

---

Your station id is ……..
You must go to room # …….

Please read the following text to the participant waiting for you in that room:

"I will go back and do a task. After I have done so, I will come back and tell you my score. I will also tell you how my score ranked amongst the six participants."

Texts we give to A-participants after the summation task

Form 1 (For own use)

This is for your information, only.

Your station id is ……..

Your score on the task is ……

This is the …... highest score amongst the six participants of type A.

Form 2 (To read to C player)

Your station id is ……..
You must go to room # …….

Please read the following text to the participant waiting for you in that room:

"My score on the task I did was …….. With this score, I was ranked …… amongst the six participants."

Texts we give to B-participants after the summation task

This is for your information, only.

Your station id is ……..

Your score on the task is ……

[*In Private-Ranking treatment only:* This is the …... highest score amongst the six participants of type B.]

II. Rivalry For Resources

II.1 A-Players (Status Ranking)

Part 1

Welcome to this experiment.

You will receive 5 euro for your participation in today's experiment. Depending on your decisions and the decisions of other participants in today's experiment, you may earn more money. You will be paid privately at the end of the experiment. In the experiment you will remain anonymous. Your decisions will only be linked to your station id and not to your name in any way. Neither the other participants, nor the experimenter will be able to see how much you personally earn. The experiment will take approximately 1 hour.

The participants in this experiment have been randomly divided into three groups. Six of you are of type A, six are of type B and one is of type C. The participants of type A and of type B are in this room, while the participant of type C is in a different room.

You are of type A.

The experiment is divided into two stages. You will receive instructions for each stage when it starts. We guarantee that everything we tell you in these instructions will proceed precisely as described. If you have any doubts about whether we are acting in the way described in the instructions, we will be happy to show you at the end of the experiment that this is the case.

We now start with stage 1.

In stage 1 the participants of type A and of type B will all independently perform a task during 15 minutes.

This is an important task that is often used to measure people's talents. Many scientific studies have found that people who do well in a task like this are more successful in professional life than people who do less well. You will not be told, however, what is typically a good or a bad score for this task.

The task is as follows. You will see two matrices on the computer screen. Each matrix has 10 rows and 10 columns and is filled with randomly generated numbers. Your job is to find the largest number in each of the matrices and then to add them up. You are not allowed to use calculators, but you can use the paper and pencil that you have found on your desk.

After entering a sum the computer will tell you whether it is correct or incorrect (please note that the time will continue to run while you see this result). Subsequently, irrespective of whether your answer is correct or incorrect, a new pair of matrices will appear. This means that for each pair, you have only one attempt to provide the correct answer. However, there will be new matrices as long as you are within the 15 minutes limit. We will allow a maximum of 50 attempts. Any matrices you try to solve after the 50th will not count towards your earnings.

After the task has been completed, we will inform you about your score [*Only in Private-Ranking Treatment:* and how your score ranks amongst the six participants of type A. To do this, the experimenter at the computer server will compare the score of all six participants of type A]. To ensure your anonymity, this experimenter is seated in a way that he cannot identify you. He can link your score to your station identification, but cannot see who is sitting at each particular station.

This experimenter will also determine the two players of type A who have the highest score. If there is a tie, then the higher score is assigned to the participant who reached that score first.

Only the two participants of type A with the highest score will be paid for their score in the summation task. Each will receive 3 euro for each correct solution that he or she had. This is added to the show-up fee of 5 euro.

This means that the other four participants of type A will only receive the 5 euro show-up fee today.

Remember that studies have found that people who do well in a task like this are more successful in professional life. You will not know how people typically perform in this task. Nor will you know how other participants scored in the task today.

After this stage, you will be asked to inform the participant of type C about your score. You will tell her your score and your rank amongst the participants of type A. Note that each of the A players will report to the same C player.

You will have to go to a separate room where this participant will be waiting. This participant does not know what task you did and what the score means. He or she has only been told that a higher score is thought to lead to a more successful professional life. Importantly, this participant will hear your score and how it ranks to the other A participants. Only the participant of type C will be able to make this comparison.

As explained, the experimenter at the computer server will not be able to link scores to individuals. He will prepare closed envelopes that will be distributed by a different experimenter. The first set of envelopes will only be distributed to participants of type A. This envelop will contain your score, your rank and the text you will need to read to the participant of type C. The other experimenter will not know the contents of the envelop he is handing to you.

After all participants of type A have reported to the participant of type C, the experimenter at the computer server will prepare envelopes for all participants. These include your earnings and a receipt for you to sign. Once again, the experimenter handing out the envelopes does not know its contents, so your earnings in today's experiment remain anonymous.

We would like you to see the type C participant that you will present your score and rank to. For this reason, each of you will now first leave the room and read aloud –to the C-particpant– a text that will be given to you. Note that each of you will be going to the same C participant.

Part 2

This brings us to the end of the first stage of the experiment. The experimenter at the computer server will now prepare the envelopes with information about your score and rank amongst the participants of type A. In the meantime, we ask you to fill out a brief questionnaire.

Now you will be asked to inform the participant of type C about your score and your rank. Recall that every type A player is visiting the same participant. Each participant of type A will be accompanied by the experimenter to the room in which the participant of type C is waiting. There, each A participant will open the envelop and read aloud the text to the C participant. Then all A participants will return to their desks. Participants of type B will remain seated during this process and are not informed about what the A participants will do.

Remember that the C participant does not know what task you did and what the score means. He or she has only been told that a higher score is thought to lead to a more successful professional life. Importantly, this participant will know your score and how it ranks to the other A participants.

II.2 B-Players

Part 1

Welcome to this experiment.

You will receive 5 euro for your participation in today's experiment. Depending on your decisions and the decisions of other participants in today's experiment, you may earn more money. You will be paid privately at the end of the experiment. In the experiment you will remain anonymous. Your decisions will only be linked to your station id and not to your name in any way. Neither the other participants, nor the experimenter will be able to see how much you personally earn. The experiment will take approximately 1 hour.

The participants in this experiment have been randomly divided into three groups. Six of you are of type A, six are of type B and one is of type C. The participants of type A and of type B are in this room, while the participant of type C is in a different room.

You are of type B.

The experiment is divided into two stages. You will receive instructions for each stage when it starts. We guarantee that everything we tell you in these instructions will proceed precisely as described. If you have any doubts about whether we are acting in the way described in the instructions, we will be happy to show you at the end of the experiment that this is the case.

We now start with stage 1.

In stage 1 the participants of type A and of type B will all independently perform a task during 15 minutes.

This is an important task that is often used to measure people's talents. Many scientific studies have found that people who do well in a task like this are more successful in professional life than people who do less well. You will not be told, however, what is typically a good or a bad score for this task.

The task is as follows. You will see two matrices on the computer screen. Each matrix has 10 rows and 10 columns and is filled with randomly generated numbers. Your job is to find the largest number in each of the matrices and then to add them up. You are not allowed to use calculators, but you can use the paper and pencil that you have found on your desk.

After entering a sum the computer will tell you whether it is correct or incorrect (please note that the time will continue to run while you see this result). Subsequently, irrespective of whether your answer is correct or incorrect, a new pair of matrices will appear. This means that for each pair, you have only one attempt to provide the correct answer. However, there will be new matrices as long as you are within the 15 minutes limit. We will allow a maximum of 50 attempts. Any matrices you try to solve after the 50th will not count towards your earnings.

After the task has been completed, we will inform you about your score [*Only in Private-Ranking Treatment:* and how your score ranks amongst the six participants of type B. To do this, the experimenter at the computer server will compare the score of all six participants of type B]. To ensure your anonymity, this experimenter is seated in a way that he cannot identify you. He can link your score to your station identification, but cannot see who is sitting at each particular station.

This experimenter will also determine the two players of type B who have the highest score. If there is a tie, then the higher score is assigned to the participant who reached that score first.

Only the two participants of type B with the highest score will be paid for their score in the summation task. Each will receive 3 euro for each correct solution that he or she had. This is added to the show-up fee of 5 euro.

This means that the other four participants of type B will only receive the 5 euro show-up fee today.

Remember that studies have found that people who do well in a task like this are more successful in professional life. You will not know how people typically perform in this task. Nor will you know how other participants scored in the task today.

The type A players will now leave the room, each for a few minutes. Please remain seated quietly until this has been completed.

Part 2
This brings us to the end of the first stage of the experiment.

In this stage all B participants are required to remain quietly seated at their desks and wait till the end of stage 2.

II.3 C-players

Welcome to this experiment.

Your role in today's experiment is a passive one. You will not be asked to make any decisions.

Your only task is to hear the results of a task performed by one of the other participants. You will not be informed about the content of this task. All you need to know is that it is an important task that is often used to measure people's talents. Many scientific studies have found that people who have a high score in a task like this are more successful in professional life than people who have a low score. You will not be told, however, what is typically a good or a bad score for this task.

You will be seated in this room. While you are waiting for the other participants, feel free to read anything you like, or to browse your phone. Before the other participants have started their task six of them will be taken to you, one at a time. Each will read a text to you and then return to the laboratory. After all the other participants have completed their tasks, the same six will come here again, one at a time. They will read to you their score and how this ranks amongst the six participants.

You may not speak or interact with the other participants in any way.

For your role in today's experiment, you will receive 10 euros on top of the 5 euros show-up fee. You will be paid and dismissed after the participants have told each of you their scores.

Finally, please treat the room you will be in with respect. You are a guest here, so please do not touch anything that is not yours.

II.4 Texts

Text to read to C-players by A-players before the summation task

---

Your station id is ……..
You must go to room # …….

Please read the following text to the participant waiting for you in that room:

"I will go back and do a task. After I have done so, I will come back and tell you my score. I will also tell you how my score ranked amongst the six participants."

Texts we give to A-participants after the summation task

Form 1 (For own use)

This is for your information, only.

Your station id is ……..

Your score on the task is ……

This is the …... highest score amongst the six participants of type A.

Form 2 (To read to C player)

Your station id is ……..
You must go to room # …….

Please read the following text to the participant waiting for you in that room:

"My score on the task I did was …….. With this score, I was ranked …… amongst the six participants."

Texts we give to B-participants after the summation task

This is for your information, only.

Your station id is ……..

Your score on the task is ……

[*In Private-Ranking treatment only:* This is the …... highest score amongst the six participants of type B.]

Appendix C: Additional Tests

Bayesian Analysis for Hypothesis 1.0

We want to test the alternative hypotheses that there is no gender difference in performance in the absence of both dimensions of competition. We base our analysis on linear regressions of the number of correct summations on a constant term and a dummy indicating the participant's gender. We do so for both *nRfR/nR* and *nRfR/SR*. As expected, this gives an insignificant gender effect in *nRfR/nR* (coefficient = –1.20, $p = 0.247$), and a significant (coefficient = –3.45, $p < 0.001$) for *nRfR/SR*. The Bayesian analysis for hypothesis 1.0 requires an assumption about the prior distribution of the effect of gender on performance in *nRfR/nR* (as measured by the regression coefficient). To formulate a 'null' hypothesis, we use the results for *nRfR/SR* and assume a normal distribution for the coefficient with mean and standard deviation determined by the corresponding *nRfR/SR* regression. This basically assumes that the gender effect on performance in *nRfR/nR* is the same as in *nRfR/SR*. For our own hypothesis of no gender difference, we use an alternative hypothesis that the gender effect centers around 0 (no effect), assuming a normal prior distribution with standard deviation 1 (our conclusions are robust to choosing standard deviation 0.1 instead). This setup allows us to calculate the posterior odds ratio of the alternative hypothesis (no gender effect) being correct to the 'null' hypothesis (same effect as in *SR*) being correct. Assuming that both models are equally likely a priori, this posterior ratio is almost 3:1.

Difference-in-Difference Private-Public Ranking

We test whether the gender difference with private ranking is different than the gender difference with public ranking. We do so separately for the cases with and without rivalry for resources. Observe in Table 2 of the main text that the gender difference increases from 2.71 to 3.45 without rivalry and decreases from 4.89 to 2.94 with rivalry. We restrict the data to the (141) observations in *nRfR/PR* and *nRfR/SR* and the (143) cases in *RfR/PR* and *RfR/SR*, respectively. In each case, we regress performance on a gender dummy, and the interaction between gender and the *SR* treatment. The results are shown in Table B1. In both cases, women perform significantly worse than men in *nRfR/PR*, confirming the results from the PtT tests reported in the main text. The interaction term with *nRfR/SR* is statistically insignificant in both regressions. We conclude that the gender effect of private ranking is not statistically significantly different than the gender effect of public ranking.

Table B1: Public versus Private Ranking

	No Rivalry for Resources		Rivalry for Resources	
Constant	12.59	(0.41)***	12.59	(0.46)***
Female	–2.74	(0.75)***	–4.36	(0.80)***
Female x Public Ranking	–0.67	(0.87)	0.93	(0.96)
N	141		143	

Bayesian Analyses for Empirical Questions 1.1 and 1.2

We first consider the effect of adding rivalry for resources to private ranking; that is, we compare *RfR/PR* to *nRfR/PR*. In the substitutes model, the effect of gender is the same in both cases. The complements model assumes that the *additional* effect of rivalry is the same as the effect of rivalry when there is no status ranking, that is, the gender effect in *RfR/nR*.

To test this, we use linear regressions of performance on gender. Denote by $\beta_1, \beta_2, \beta_3$ the coefficient for gender in *RfR/nR, nRfR/PR,* and *RfR/PR*, respectively. The complements model then predicts that $\beta_3 = \beta_1 + \beta_2$, while the substitutes model predicts $\beta_3 = \beta_2$. Our analysis shows an odds ratio of 2.3:1 in favor of the complements model. We thus conclude that adding rivalry for resources to private ranking is more likely to increase the gender difference than to have no effect. We also test the reverse, that is, the effect of adding private ranking to rivalry for resources. In this case, the substitutes model predicts $\beta_3 = \beta_1$. Here, the odds ratio is about 1.2:1, which means that it is more or less equally likely that (i) adding private ranking increases the gender difference observed under rivalry for resources and (ii) adding private ranking has no effect.

Finally, the analysis for public ranking follows the same lines and yields the odds ratio's reported in Table 4 of the main text. Details are available upon request.

Bayesian Analysis for Hypothesis 3

We base our analysis on linear regressions of the number of correct summations on a constant term and a dummy indicating the participant's gender. We do so for *nRfR/nR* and *nRfR/PR*. The former gives an insignificant gender effect in *nRfR/nR* (coefficient = –1.20, p = 0.247), the latter a significant gender effect (coefficient = –2.71, p = 0.003). The Bayesian analysis for hypothesis 3 requires an assumption about the distribution of the effect of gender on performance in *nRfR/PHR* (as measured by the regression coefficient). For our hypothesis of no effect, we use the results for *nRfR/nR* and assume a normal distribution for the coefficient with mean and standard deviation determined by the corresponding *nRfR/nR* regression. We compare this to a case with gender difference, based on nRfR/PR. The latter assumes a normal distribution for the coefficient with mean and standard deviation determined by the corresponding *nRfR/PR* regression. This setup allows us to calculate the posterior odds ratio of the two hypotheses being correct. Assuming that both models are equally likely a priori, this posterior ratio is almost 43:1 in favor of the gender difference in *nRfR/PHR* being like in *nRfR/nR.*