The gender gap in competitiveness: women shy away from competing with others, but not from competing with themselves

By Johanna Mollerstrom and Katharina Wrohlich

Women are less willing than men to compete against others. This gender gap can partially explain the differences between women's and men's education and career choices, and the labor market disparities that result. The experiments presented here show that even though women are less willing than men to compete against others, they are just as willing as men are to take on the challenge of improving themselves and competing against their own past performance. For organizations striving towards gender equality, this opens up new possibilities for institutionalizing competitive pressure.

Despite significant improvements in gender equality over the past decades, women are still underrepresented at top positions worldwide and are more often found in lower-paid professions.[1] It has been suggested that one reason for this gender discrepancy is that women are less willing than men to engage in competitive behavior. Conducting a field test of such a hypothesis is difficult because of selection issues, and researchers have thus turned to experimental methods (Box 1).

In a seminal paper, it was shown that the hypothesis did indeed hold in the laboratory: in an experiment, women were less willing than equally able men to enter a tournament (competition) where they could earn money.[2] As a result, women left the experiment having earned less money on average than the participating men, and less money than they would have had they been willing to compete more often. This finding has since been replicated hundreds of times all over the world.[3] The experimental design that is generally used in this area of research is outlined in Box 2.

More recent work has focused on the predictive power of such laboratory assessments of competitiveness over career choices and labor market outcomes. In general, there is a strong and significant correlation between behavior inside and outside of the laboratory, respec-

1 Francince Blau, Marianne Ferber, and Anne Winkler, "The Economics of Women, Men, and Work, 7th edition," Prentice Hall (2013). For evidence on Germany, see for example Elke Holst and Katharina Wrohlich, "Top decision-making bodies in large companies: gender quota shows initial impact on supervisory boards; executive board remains a male bastion," *DIW Economic Bulletin* 1+2 (2017), 3–16, and Ann-Christin Hausmann and Corinna Kleinert, "Berufliche Segregation auf dem Arbeitsmarkt: Männer und Frauendomänen kaum verändert," *IAB-Kurzbericht* no. 9 (2014).

2 Muriel Niederle and Lise Vesterlund, "Do Women Shy Away From Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* vol. 122 no. 3 (2007), 1067–1101.

3 Muriel Niederle and Lise Vesterlund, "Gender and Competition," *Annual Review of Economics* 3 (1) (2011), 601–603. Muriel Niederle, "Gender," Alvin Roth and John Kagel (eds): *Handbook of Experimental Economics* vol. 2 (2016), Princeton University Press.

Box 1

Experimental Economics

The study of modern experimental economics dates back to the 1930s. Since then, this research field has experienced a rapid growth that has not yet leveled off. The core idea behind conducting an experiment instead of relying on observational data is that causality can be established, which is rarely the case in the harder-to-control environment beyond the laboratory.

Economic experiments can investigate questions from all areas of economics, and the focus can be on individual and/or group decision-making.

Economic experiments are typically conducted in laboratories with each participant seated at a computer in separate cubicles. Participants are matched anonymously: they do not know whom they are playing against nor how much money anyone is earning.

Though most experiments are still conducted in the lab, field experiments of various types have become more common. A field experiment can be set up so that participants are not aware that they are part of an experiment, or it can take the form of a "lab-in-the-field," meaning that the experiment is conducted in the participants' natural environment. Online experiments, for which participants use their own computers or smartphones, fall into the latter category.

Many other academic disciplines use experiments on human behavior as a core method. This is most notable in psychology, which has a long history of this kind of experimentation. Economic experiments have a number of features that are different, however.

First, deception is prohibited in economic experiments—that is, participants are always fully and correctly informed about the consequences, monetary or otherwise, of their choices. For example, participants would never be told that they are matched with another player when in fact they are not.

Second, economic experiments are nearly always incentivized in the sense that participants' payments vary in relation to what they (and, potentially, others that they are matched with) do in the lab. In psychology, it is common that participants are merely rewarded with a flat payment (or a course credit) for showing up, but in economics, there is a fear that participants would not take the game as seriously without monetary incentives.

A somewhat newer line of research in experimental economics investigates what happens when the student populations, which are typically used in these studies, are replaced by other, more representative, samples. The results depend on the exact game being played, but in general, results that have been found to be robust among student populations are also robust among other groups. There are also several studies that investigate the extent to which laboratory behavior can predict behavior beyond the laboratory. Again, the results differ somewhat depending on the exact topic, but most often there is a strong correlation between—for example—competitiveness, altruism, and cooperativeness in the laboratory and similar behaviors outside the lab. This confirms the relevance of the experimental method and ensures us that learning about human behavior through experiments can actually teach us things relevant to the wider setting beyond the lab.[1]

1 Alvin Roth and John Kagel, "Handbook of Experimental Economics," Princeton University Press (1993) and Alvin Roth and John Kagel, "Handbook of Experimental Economics, vol. 2," Princeton University Press (2016).

tively.[4] In one study, for example, students in the Netherlands who, according to the laboratory experiment, were more willing to compete were also significantly more likely to choose the most competitive—and potentially most financially rewarding—track in high school. This, in turn, is a key explanation for why women were underrepresented in that high-school track.[5]

The research on gender differences in the willingness to compete has focused exclusively on the type of competitions described above—that is, competitions against other people (*other-competition*). This type of competition is important, not least because of its prevalence. But there are other ways of creating competitive pressure: in addition to other-competition, the experiments described

4 Thomas Buser, Muriel Niederle, and Hessel Oosterbeek, "Gender, Competitiveness and Career Choices," *Quarterly Journal of Economics,* vol. 129 no. 3 (2014), 1409–1447. Ernesto Reuben, Paula Sapienza, and Luigi Zingales, "Taste for Competition and the Gender Gap Among Young Business Professionals," National Bureau of Economic Research Working Paper 21695 (2015) and Thomas Buser, Noemi Peter, and Stefan Wolter, "Gender, Competitiveness and Study Choices—Evidence from Switzerland," *American Economic Review* P&P, vol. 107 no. 5 (2017), 125–130.

5 Thomas Buser, Noemi Peter, and Stefan Wolter (2014), supra.

Box 2

Laboratory competition experiment

The standard experiment investigating the willingness to compete is designed as follows:[1]

Participants are seated in the laboratory in private cubicles. Each participant has access to a computer and is told that they will complete one task per round in three different rounds, each of which will last for five minutes. The task consists of solving a series of addition problems involving five two-digit numbers, such as 53 + 84 + 31 + 64 + 12, for which the correct answer is 244. Participants are also told that more information about each subsequent round will be provided just before the round starts and that at the end of the experiment, one round will be randomly selected and the participants will receive the earnings from that round. Furthermore, participants will not receive any feedback during the experiment on their performance in the various rounds.

Before the first round, participants are told that they will be completing as many addition tasks as possible and receive a fixed amount, often one U.S. dollar, per correct task executed during the five-minute round. They then perform in Round 1.

Ahead of the Round 2, each participant is informed that he or she has been matched with another participant. The match is anonymous and their identities are not revealed to each other either during or after the experiment. Both participants complete math problems for five minutes; the one with the better performance will receive double the piece rate amount—say, two U.S. dollars—for each correct answer. The other participant will receive nothing for this round. The participants then perform in Round 2.

Participants are then given a choice of which payment scheme will apply to their performance in the third (and final) round: piece-rate or competition. If they choose to be paid on a piece-rate basis, they will receive one U.S. dollar for every math problem they answer correctly in Round 3; if they choose to compete, their performance in Round 3 will be compared to that of their matched partner in Round 2, regardless of what that person has chosen for Round 3.[2] If they perform better than their partner, they will earn two U.S. dollars for each correct answer; if they perform worse, they will earn nothing in Round 3. Participants select a payment scheme for Round 3 and then perform.

At the end of the three rounds, a measure of participants' *confidence* is collected by asking them to guess how well they performed in comparison to their opponents. They are paid an additional bonus if their guess is correct. A measure of participants' *risk attitudes* is usually collected as well: either they self-assess their willingness to take risk on a Likert scale,[3] or they make a series of incentivized choices between lotteries and safe amounts of money.

Thereafter, participants are paid out their earnings from the experiment and informed whether they won or lost in the second (and, if they chose to compete the third) round. They fill out a demographic questionnaire (which, among other things, asks for their gender) and are then free to leave the laboratory.

The outcome variable of interest collected during the experiment is what the participant chooses for Round 3: to receive a piece-rate payment or to enter into competition with an opponent. The finding in the original experiment, which has since been replicated numerous times, is that women are less likely than men to select to compete even when controlling for ability, i.e. performance in Rounds 1 and 2. Moreover, had they elected to compete, female participants who performed above average could have earned significantly more money than men with average performances.[4]

1 Muriel Niederle and Lise Vesterlund, "Do Women Shy Away From Competition? Do Men Compete Too Much?" Quarterly Journal of Economics, vol. 122 no. 3 (2007), 1067–1101.

2 The reason that the comparison is made against the other person's second round performance is two-fold: First, it ensures that the choice of whether or not to compete cannot have negative external effects on the matched partner, i.e. it is possible for both participants to compete in the third round and win. Second, it eliminates the need for participants to attempt to second-guess their partner's choice about whether or not to compete.

3 Such a scale ranges from „not willing to take any risk" to "very willing to take risks".

4 Muriel Niederle and Lise Vesterlund (2011): Gender and Competition. Annual Review of Economics 3 (1), 601–603.

here examine competition against one's own previous performance (*self-competition*).[6] This type of competition embodies notions of self-improvement, progress, and mastery. Our study is the first to consider this issue in the field of economics.

The question of potential gender differences in the willingness to self-compete is important for a multitude of

6 Coren Apicella, Elif Demiral, and Johanna Mollerstrom, "No Gender Difference in Willingness to Compete When Competing Against Self," *American Economic Review* P&P, vol. 107 no. 5 (2017), 136–140.

reasons, with two standing out. First, the ability to continually challenge oneself and improve is vital to success in most areas of the labor market. Determining whether there are gender differences in the willingness to self-compete can help us understand if men and women are on equal or unequal footing when it comes to this area of self-development.

Second, competition is often used to motivate employees to perform better. It has been established that allowing people to compete against others enhances their performance[7]—but since women are less willing than men to enter these other-competitive contexts, competition as a motivational tool may have undesirable side effects when it comes to gender equality. If self-competitions—which exhibit no gender differences—are just as motivating as other-competitions, they may be a good alternative for institutions striving toward gender equality.

Gender gaps differ depending on competition type

The two competitiveness experiments described here were conducted in the fall of 2016 by Johanna Mollerstrom (DIW Berlin and Humboldt University) together with Coren Apicella of the University of Pennsylvania and Elif Demiral of George Mason University (GMU). The first experiment was carried out in the laboratory of the Interdisciplinary Center for Economics Science at GMU with 204 participants. The second experiment was implemented using the online labor market *Amazon MTurk;* 994 participants took part.

Laboratory experiment

The laboratory experiment comprised two treatments: *Other* and *Self*. The *Other*-treatment followed the design usually used in this literature (see Text Box 2). The 204 participants in the experiment performed a series of math tasks in three rounds (each round lasting five minutes), with no feedback given between rounds. In the first round, participants were paid one U.S. dollar for every correctly solved problem (*piece rate*). In the second round, participants were matched in pairs, and the partner with the higher score was paid double the piece rate (two U.S. dollars) for every correctly solved problem, whereas the other participant received nothing. This payment scheme was called the *tournament rate*. Then participants were given a choice about which payment scheme to apply for Round 3: the piece rate (one U.S. dollar for every correctly solved problem) or the tournament rate (two U.S. dollars for every correct answer but only if he or she won the tournament in this round).

The *Self*-treatment was identical with the following important differences. First, the participants were not matched against another player; rather, their scores in the tournament from Round 2 were compared to their own Round 1 scores. Second, if participants chose the tournament rate for Round 3, their score from Round 3 would be compared to their score from Round 2.

Out of the 204 participants, 50.5 percent were female. They earned an average of 17.42 U.S. dollars (per person) for their participation. Sessions lasted 40 minutes on average.[8]

Table 1 shows the percentage of participants who chose the tournament rate in advance of Round 3, broken down by gender and treatment. Table 2 shows the corresponding regression results (ordinary least square regressions where the decision between the piece rate and the tournament rate for Round 3 is the dependent variable). Controlling for their ability to solve the problems, women were significantly less likely than men to compete against others, with 57.7 percent of men choosing to compete compared to only 37.5 percent of women. However, when the participant was competing against his or her own previous score, the gender gap was significantly reduced and no longer statistically significant.

Online Experiment

The online experiment consisted of four treatments. The *Other* and *Self* treatments were identical to those implemented in the laboratory, except that the math problems were replaced with a counting task, Captcha-style to prevent cheating,[9] and the rounds were shortened to 90 seconds. Two additional treatments were also conducted: *Other, Same Gender* and *Other, Same Ability*. In *Other, Same Gender*, participants were matched to someone of the same gender. In *Other, Same Ability*, participants were matched with someone who completed the same number of tasks correctly in Round 1. Participants were informed about these aspects of the matching process in advance of the competitions.

The platform *Amazon MTurk* was used to recruit the 994 people from North America (49.9 percent female) who

7 Uri Gneezy, Muriel Niederle, and Aldo Rustichini, "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics* vol. 118 no. 3 (2003), 1049–1074.

8 The laboratory experiment was programmed in z-Tree. See Urs Fischbacher, "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," *Experimental Economics* vol. 10 no. 2 (2007), 171–178.

9 *Captcha* stands for "Completely Automated Public Turing test to tell Computers and Humans Apart." Captcha-style tasks can't be solved by a calculator or a computer program.

took part in the online experiment.[10] The experiment was conducted in November 2016; on average, participants earned 1.20 U.S. dollars over the course of a session that lasted approximately 12 minutes.

Table 1 shows that the gender gap in the willingness to other-compete amounted to 12 percentage points in the online experiment. This is highly statistically significant as can be seen in the corresponding regression in Table 2. In the *Self*-treatment, the sign of the gap reversed and it was no longer statistically significant.[11]

The two additional versions of the *Other*-treatment were conducted in order to investigate if mimicking certain aspects of self-competition (the knowledge about one's gender and one's own ability) is enough to reduce the gender gap in competitiveness. There was still a gender difference in the *Other, Same Gender* treatment but not in *Other, Same Ability* (see Tables). The latter result indicates that receiving a signal that the matched opponent has an ability akin to one's own is indeed enough to eliminate the gender difference.

Table 1

Percentage Choosing Tournament Rate, by Treatment and Gender
In percent

Panel A: Laboratory Experiment			
Treatment:	**Women**	**Men**	**Total**
Other	37.5 (7.1)	57.7 (6.9)	48.0 (5.0)
Self	41.8 (6.7)	55.1 (7.2)	48.1 (4.9)
Total	39.8 (4.8)	56.4 (5.0)	48.0 (3.5)
Panel B: Online Experiment			
Other	27.8 (4.2)	40.0 (4.3)	34.3 (3.0)
Other, Same Gender	21.9 (3.7)	34.1 (4.2)	28.0 (2.8)
Other, Same Ability	30.6 (4.2)	33.3 (4.3)	32.0 (3.0)
Self	35.7 (4.2)	31.1 (4.3)	33.5 (3.0)
Total	29.0 (2.0)	34.7 (2.1)	31.9 (1.5)

Note: Standard errors in parentheses.

Source: Authors' own calculations.

Confidence and risk attitude play a role in the decision to compete

In both the laboratory and online versions of the experiment, all participants filled out a questionnaire after the main experiment that collected basic demographics and self-reported risk attitudes. Participants were also asked incentivized questions about how they believed their performance progressed over time and how it compared to that of other participants.

Similar to the findings of previous studies, women were found to be less willing to take on risk and less overconfident in the two experiments reported here.[12] This fundamentally explains why women are less willing than men to compete against others.[13]

That there is no gender gap in the willingness to self-compete could be because the roles of risk attitude and confidence differ depending on which form of competition (against others or against oneself) is considered. Women are more risk-averse than men both in the *Self*- and the *Other*-treatments. Whereas men are significantly more confident in their abilities when competing against others than women are, there is no gender difference in confidence when it comes to self-competition.[14]

This indicates that confidence may indeed play a different role in self-competition than it does in other-competition, and that this can partly explain why there is a gender gap in the willingness to other-compete, but not in the willingness to self-compete.[15]

Both forms of competition enhance performance

Competitions are often used to boost performance. If self-competitions are to be used as an alternative to other-competitions in, for example, organizational settings, it

10 *Amazon MTurk* is an online labor market with tens of thousands of active participants. Increasingly, *Amazon MTurk* is being used for experimental studies. See also John Horton, David Rand, and Richard Zeckhauser, "The online laboratory: Conducting experiments in a real labor market," *Experimental Economics* vol. 14 no. 3 (2011), 399–425.

11 Moreover, a difference-in-difference estimation reveals that the gender gaps in the two treatments differ.

12 Rachel Croson and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature* vol. 47 no. 2 (2009), 171–178 and Muriel Niederle and Lise Vesterlund (2007), supra.

13 This is shown in Table 2 by the decrease in magnitude and significance of the coefficient for "female" between columns 2 and 4, respectively.

14 Coren Apicella, Elif Demiral, and Johanna Mollerstrom (2017), supra.

15 A regression analysis confirms that this is the case. This analysis was carried out by regressing a dummy indicating the participants' choice to compete in the third round on confidence on a dummy that takes the value 1 if the treatment is *Other* and 0 if the treatment is *Self*, and the interaction between the two (controlling for Round 1 scores as in all regressions). The results show that confidence has a significantly larger impact on the choice of whether to compete in the *Other*-treatment than in the *Self*-treatment (p=0.068 and p=0.014 in the laboratory and online experiment, respectively).

Table 2

Regression Analyses

Panel A: Lab Experiment				
	(1)	(2)	(3)	(4)
	(Other)	(Other)	(Self)	(Self)
Female	−0.195**	−0.114	−0.132	−0.029
	(0.10)	(0.10)	(0.10)	(0.10)
Confidence		0.246**		−0.013
		(0.11)		(0.10)
Risk		0.039*		0.091***
		(0.02)		(0.02)
Constant	0.177	−0.212	0.503***	−0.008
	(0.14)	(0.22)	(0.16)	(0.20)
N	100	100	104	104
R-square	0.116	0.180	0.019	0.140
Panel B: Online Experiment				
	(5)	(6)	(7)	(8)
	(Other)	(Other)	(Self)	(Self)
Female	−0.126**	−0.090	0.052	0.083
	(0.06)	(0.06)	(0.06)	(0.06)
Confidence		0.246***		0.128**
		(0.06)		(0.06)
Risk		0.045***		0.032**
		(0.01)		(0.01)
Constant	0.297***	−0.114	0.371***	0.120
	(0.07)	(0.10)	(0.08)	(0.12)
N	245	245	248	248
R-square	0.028	0.172	0.006	0.042
Panel C: Online Experiment, ctd				
	(9)	(10)	(11)	(12)
	(Other, Same Gender)	(Other, Same Gender)	(Other, Same Ability)	(Other, Same Ability)
Female	−0.122**	−0.094*	−0.028	0.030
	(0.06)	(0.05)	(0.06)	(0.06)
Confidence		0.269***		0.287***
		(0.06)		(0.05)
Risk		0.027**		0.042***
		(0.01)		(0.01)
Constant	0.349***	0.063	0.307***	−0.117
	(0.07)	(0.09)	(0.07)	(0.11)
N	257	257	244	244
R-square	0.019	0.158	0.002	0.158

Note: The dependent variable is a dummy indicating choice of competition in the third round. Robust standard errors in parentheses. All regressions control for task ability measured as the score in task 1. Risk is a 1-10 self-assessed index of willingness to take risk with 1="Not at all willing to take risks" and 10="Very willing to take risk". Confidence is a dummy that takes on the value 1 for subjects who believed that they improved their performance between the second and the third round ("Self"-treatment) or that they performed better than the person they were matched to in the second round (the three "Other"-treatments).
*Significance: ***p<0.01 **p<0.05 *p<0.1*

Source: Authors' own calculations.

is important to know if both types of competition affect performance in the same way. For this purpose, we can analyze the performance improvement between the first and second rounds in the experiments. In the laboratory experiment, the average score improvement was 23.9 in the *Other* and 18.2 percent in the *Self*-treatment. The corresponding figures for the online experiment were 22.0 and 18.0 percent, respectively. The difference wasn't significant in either case. It can hence be concluded that self-competition does not yield inferior results than other-competition when it comes to enhancing performance.

Conclusions

Taken together, the laboratory and the online experiment lead to the conclusion that, even though women are less willing than men to compete against others, there is no such gender gap when the competition is instead against one's own previous performance. Organizations keen on guaranteeing women an even playing field should thus consider using self-competition (that is, competition against one's past performance) as opposed to competition against others (for instance, against colleagues or peers) as a motivation and performance-enhancing tool. Self-competition can also be a useful instrument when it comes to determining remuneration, that is, when making decisions on promotions, raises, or bonus payments.

One possible concrete example is a sales team. Giving team members the objective "increase your performance (for instance, sales figures) by *x* percent compared to last year's" instead of "you have to sell more than the other teammates" would yield equally good results in terms of performance. In addition, it would ensure that women are not put at a disadvantage compared to men.

Johanna Mollerstrom is Head of the of Competition and Consumers Department at DIW Berlin | jmollerstrom@diw.de

Katharina Wrohlich is Research Associate on the Executive Board at DIW Berlin | kwrohlich@diw.de

JEL: C90, C91, J16, J71

Keywords: gender, competition, discrimination, experiment

DIW Berlin — Deutsches Institut
für Wirtschaftsforschung e. V.
Mohrenstraße 58, 10117 Berlin
T +49 30 897 89 –0
F +49 30 897 89 –200

Publishers
Prof. Dr. Tomaso Duso
Dr. Ferdinand Fichtner
Prof. Marcel Fratzscher, Ph.D.
Prof. Dr. Peter Haan
Prof. Dr. Claudia Kemfert
Prof. Dr. Lukas Menkhoff
Prof. Johanna Mollerstrom, Ph.D.
Prof. Karsten Neuhoff, Ph.D.
Prof. Dr. Jürgen Schupp
Prof. Dr. C. Katharina Spieß
Prof. Dr. Gert G. Wagner

Reviewer
Dr. Frauke H. Peter

Editors in chief
Dr. Gritje Hartmann
Dr. Wolf-Peter Schill

Editorial staff
Renate Bogdanovic
Dr. Franziska Bremus
Prof. Dr. Christian Dreger
Sebastian Kollmann
Markus Reiniger
Mathilde Richter
Miranda Siegel
Dr. Alexander Zerrahn

Layout and Composition
eScriptum GmbH & Co KG, Berlin

Sale and distribution
DIW Berlin
ISSN 2192-7219