TI 2016-096/I
Tinbergen Institute Discussion Paper

Do Women give up Competing more easily? Evidence from the Lab and the Dutch Math Olympiad

Thomas Buser
Huaiping Yuan

Faculty of Economics and Business, University of Amsterdam, and Tinbergen Institute, The Netherlands.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Do women give up competing more easily? Evidence from the lab and the Dutch Math Olympiad*

Thomas Buser and Huaiping Yuan†

November 10, 2016

Abstract

We conduct three lab experiments and use field data from the Dutch Math Olympiad to study how the gender gap in willingness to compete evolves in response to experience. The main result is that women are more likely than men to stop competing if they lose. In the Dutch Math Olympiad, this means that girls who do not make the top 1000, and therefore do not advance to the next round, are less likely to compete again one year later while there is no effect on boys. In an additional experiment, we show that men are more likely than women to start and keep competing after receiving positive feedback. In a third experiment, we show that the gender difference in the reaction to losing is not present when winning and losing are random rather than the outcome of competition. The fact that women are more likely to give up competing after a setback may help to explain why fewer women make it to the top in business and academia. **JEL:** C91, D03, J01, J16

Keywords: willingness to compete, gender, feedback, career decisions, laboratory experiment

*We would like to thank seminar and workshop participants in Amsterdam, Leuven, Maastricht, San Francisco, São Paulo, Stanford, Tel Aviv, Vienna, and Zürich for useful comments. Thomas Buser gratefully acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO) through a personal Veni grant and we would like to thank CREED for letting us use their lab. An early version of this paper was circulated under the title **"How does the gender difference in willingness to compete evolve with experience?"**.

†Buser: University of Amsterdam and Tinbergen Institute. t.buser@uva.nl. School of Economics, Roetersstraat 11, 1018WB Amsterdam, The Netherlands. Yuan: University of Amsterdam and Tinbergen Institute.

1 Introduction

People who want to advance in their careers are regularly confronted with situations that are in essence winner-takes-all competitions, including admission procedures to high-ranked universities, job applications, promotion tournaments, and competition between entrepreneurs. Whether an individual actively seeks out such situations or tries to avoid them will partially determine what kind of career they end up in and how far they advance up the hierarchy. Starting with Niederle and Vesterlund (2007), a large literature in experimental economics shows that women are less willing than men to enter competitions. This gender difference has attracted a lot of attention as a potential explanation for gender differences in career choices and labour market outcomes (Bertrand, 2011).

In most experimental studies, the decision to compete or not is one-shot. But in a career context, these decisions are made in a dynamic setting. To make it to the top, one has to be willing to compete repeatedly and be willing to keep competing after a setback. We conduct three incentivised lab experiments and use field data from the Dutch Math Olympiad to empirically investigate how gender differences in willingness to compete evolve over time. In particular, we ask whether men and women react differently to losing and winning.

In our lab experiments, participants perform the same task over a number of rounds. In each round, they decide whether they want to receive a piece-rate payment or enter a winner-takes-all competition against a randomly selected opponent. Participants who enter the competition and outperform their opponent receive twice the piece rate, those who enter the competition and lose receive no payment in that round. After each round, participants who choose to compete learn whether they won or lost against their opponent. By comparing the likelihood of choosing competition again in subsequent rounds between participants who win in round 1 and participants who lose in round 1, we can determine the effect of losing and whether there is a gender difference in the reaction to losing. In two further experiments, we explore the effect of giving feedback to those who do not compete and the role of risk.

The Dutch Math Olympiad is a yearly math competition which draws participants mainly from years four and five of the six-year academic track of secondary school. In each year, the top 1000 performers advance to the second round of the competition. Those who participate in year four can participate again in year five regardless of their performance. This generates a sharp regression discontinuity design where we compare participants who score just above and just below the top-1000 cutoff to determine the causal effect of losing on the likelihood of participating again a year later, and the gender difference therein. Our sample of all fourth-year academic-track students who participate in 2010-14 contains more than 10,000 observations.

The seminal study on gender differences in willingness to compete is Niederle and Vesterlund (2007). In their design, participants in a lab experiment choose between a piece-rate payment and a winner-takes-all competition against three randomly chosen opponents. Although there is

no gender difference in performance, they find that twice as many men than women choose the competition. While many men are overconfident and enter the competition with low chances of winning, many high-performing women hurt their expected earnings by not competing. This result has been replicated numerous times (see Croson and Gneezy, 2009, and Niederle and Vesterlund, 2011, for surveys).[1]

A growing literature confirms the external relevance of competition decisions made in the lab for predicting career choices. Buser, Niederle, and Oosterbeek (2014) show that competing in an experiment predicts the study choices of Dutch high-school students and explains a significant portion of the gender gap therein. Zhang (2012) shows that it predicts the choice of entering a highly competitive university entrance exam in China. Reuben, Sapienza, and Zingales (2015) show that it predicts the starting salary and industry choice of graduating MBA students. Reuben, Wiswall, and Zafar (2015) show that it predicts future salary expectations of undergraduate students. Buser, Plug, and Geijtenbeek (2015) show that it correlates with monthly earnings in a diverse sample of the Dutch population. Berge et al. (2015) show that it predicts the investment choices of Tanzanian entrepreneurs. Flory, Leibbrandt, and List (2015) and Samek (2015) run field experiments recruiting people for real jobs and find that compensation schemes which depend on relative performance deter women from applying relative to men.

An important open question is how the gender gap in competitiveness evolves over time in response to experience. If the gap is mainly due to wrong beliefs, it should shrink in response to experience as long as people rationally update their beliefs.[2] However, a number of experimental studies suggest that men and women may react differently to winning and losing in competitions in ways which may actually enhance the gender gap. In Buser (2016), participants compete in a task and are then informed of their score and whether they won or lost. Relative to winners, male losers subsequently pick a higher individual performance target in the same task whereas female losers perform worse. Gill and Prowse (2014) similarly find that women are more prone than men to reducing their performance after losing a competition.[3] Möbius et al. (2011) and

[1] A separate literature started by Gneezy, Niederle, and Rustichini (2003) shows that men's performance tends to increase more strongly in response to competitive pressure than women's but a large number of studies on tournament entry indirectly show that this is not the case for the task used in the present experimental design (Niederle and Vesterlund, 2011). Iriberri and Rey-Biel (2016) show that this effect is present in the context of a mathematics contest for students aged between 10 and 16: girls underperform relative to their math grades and this underperformance is worse in later stages of the contest where competitive pressure is higher.

[2] Starting with Niederle and Vesterlund (2007), many experiments show that women are less confident than men. Wozniak, Harbaugh, and Mayr (2014) show that when people receive detailed relative performance feedback prior to choosing the payment scheme, the gender gap in competitiveness is substantially reduced.

[3] A number of papers use data from professional sports to analyse gender differences in the effect of competition outcomes. These are subject to the caveats that mixed-gender competition is rare, samples are highly selective, and identification of gender differences is complicated by the fact that the competitive settings and incentives are rarely equal for male and female athletes even within the same sport. The most relevant is Wozniak (2012), who finds that the effect of prior tournament outcomes on the propensity of entering further tournaments is similar for male and female tennis players (he does find a subtle gender difference, though, in how long the effect of prior experience lasts, with men being more affected than women by outcomes further in the past). Also looking at tennis players, De Paola and Scoppa (2015) find that female, but not male, players perform worse after losing the first set. Legge

Buser, Gerhards, and van der Weele (2016) find that men update their beliefs about their own abilities more strongly in response to both positive and negative feedback which could perpetuate initial gender differences in confidence for high performers.

Our results show that it is far from clear that the difference in willingness to compete between equally performing men and women will disappear with experience. In our lab experiments, we find a large gender difference in the reaction to losing. Women who compete and lose are much more likely to stop competing than men who compete and lose. This leads to a large and persistent gender gap for those individuals who are initially willing to compete. Using our field data from the Dutch Math Olympiad, we show that this gender difference in the reaction to losing carries over to a relevant field setting. Not making the top 1000 hardly has an effect on boys, but girls who do not make the top 1000 are significantly less likely to participate again one year later.

In an additional experiment, we ask what happens if those who do not compete receive equivalent feedback. We find that men who initially do not compete are more likely than women to start and keep competing over the following rounds after receiving positive feedback. In a third experiment, we show that the gender difference in the reaction to losing does not occur in an environment that is characterised by the same risks but which lacks the competitive dimension. Women are no more likely than men to switch from a risky to a safe payment scheme after losing. Finally, we show that gender differences in confidence, belief updating and performance cannot explain the gender difference in the reaction to losing either. This hints at a direct effect of losing on the preference for competition of women who are initially willing to compete. We conclude by discussing what our results imply for our understanding of gender differences in career choices and labour market outcomes, in particular the low number of women in top positions in science and business.

The rest of the paper is structured as follows. Section 2 explains the experimental design. Section 3 presents the experimental results. Section 4 describes the results from the field study. Section 5 explores potential mechanisms behind our findings. Section 6 discusses and concludes.

2 Experimental design

The experimental design is based on Niederle and Vesterlund (2007). Participants earn money for their performance in a real effort task which consists in adding up sets of five two-digit numbers. We ran three experiments which we will refer to as the "main", "feedback" and "risk" experiments. We will start by describing the design of the main experiment and then lay out the differences in the other two experiments.

The main experiment consists of six paid rounds. Participants have three minutes per round

and Schmid (2013) find that skiers who narrowly miss a place on the podium underperform in the next race, this effect being somewhat stronger for female skiers, and Rosenqvist (2016) finds that golfers who narrowly miss the cut underperform in the next tournament but does not find a gender difference.

to solve as many addition problems as they can. Wrong answers are not penalised. In each round, participants are paired with a new randomly chosen, anonymous opponent and have to choose how they would like to be paid for their performance. They can choose between piece-rate and competitive remuneration. If they choose the piece rate, they receive one point per correct answer regardless of the performance of their opponent. If they choose competition, they receive two points per correct answer if they score higher than their opponent and nothing if they do not (in case of a tie, winning or losing is randomly determined). Opponents are randomly chosen amongst all other participants regardless of their choice. One point is worth 25 Euro cents and all rounds are paid.

After each round, participants who choose competition receive feedback on their absolute and relative performance. That is, they learn their score and whether they won or lost against their opponent (the feedback reads "You scored X correct answers. You scored higher (lower) than your opponent. You therefore won (lost) against your opponent."). Before the start of each round, we elicit an incentivised measure of participants' beliefs about their own performance. Participants are asked to predict their rank compared to all other participants in their session in the upcoming round. If their guess is within plus-minus one of the true rank, they receive a bonus of four points.

Before the start of the incentivised part of the experiment, there is a three-minute practice round after which the participants learn their score but receive no relative feedback. At the end of the last round, before they see a screen which summarises their earnings, participants fill in a short questionnaire asking for their gender and age and eliciting their perception about their own willingness to take risk and their own competitiveness.

To elicit risk attitudes, we asked "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" (Dohmen et al., 2011). The answer is on a scale from 0 ("unwilling to take risks") to 10 ("fully prepared to take risk"). To elicit competitive attitudes, we ask an analogous question, "How competitive do you consider yourself to be? Please choose a value on the scale below, where the value 0 means 'not competitive at all' and the value 10 means 'very competitive"'.

To summarise, in each of the six rounds the timeline is as follows. Participants

1. predict their rank in the upcoming round compared to all other participants present in the lab;
2. choose between competitive and piece-rate payment schemes;
3. perform in the task;
4. learn their score and, if they chose competition, whether or not they beat their opponent

In the feedback experiment, we ask what happens if those who choose the piece rate receive feedback which is equivalent to the feedback received by those who compete. In this experiment, participants learn whether they outperformed their opponent even if they choose the piece rate

(for participants who choose the piece rate, the message reads "You scored X correct answers. You scored higher (did not score higher) than your opponent. You therefore would have won (lost) against your opponent."). The rest of the design is identical to the main experiment except that there are four rounds of four minutes each.

The aim of the risk experiment is to determine whether we find the same patterns in a setting which is identical to the main experiment but in which winning and losing are entirely due to luck. Participants perform the adding-up-numbers task over six rounds of three minutes. In each round, they decide between a safe piece rate of 1 point per correct answer and random remuneration whereby they receive 2 points per correct answer with probability x and nothing with probability *1-x* (x is constant across rounds and randomly fixed at 0.3, 0.4, 0.5, 0.6 or 0.7 with equal probability).

In the main experiment, a total of 188 individuals participated in 7 sessions with 21 to 31 participants each. The sessions were run in November 2015. Participants earned 21.2 Euros on average including a 7-Euro show-up fee. In the feedback experiment, a total of 184 individuals participated in 7 sessions with 21 to 31 participants each. The sessions were run in January 2015. Participants earned 20.8 Euros on average including a 7-Euro show-up fee. In the risk experiment, a total of 188 individuals participated in 7 sessions with 22 to 31 participants each. Participants earned 18.4 Euros on average including a 7-Euro show-up fee. The sessions were run in June 2016. Students could only participate in one of the experiments. All sessions were were roughly gender-balanced. All sessions were run at the CREED lab at the University of Amsterdam and made use of z-Tree (Fischbacher, 2007).

3 Experimental results

In this section, we will describe and analyse the data from the main and feedback experiments. We will first describe how the willingness to compete of men and women evolves over the rounds. We will then determine whether for those who initially compete, there is a gender difference in the effect of losing in the first round on subsequent willingness to compete again. We will use the data from the risk experiment, as well as the belief elicitation data from the main and feedback experiments, to analyse potential mechanisms behind our findings in Section 5.

3.1 Evolution of the gender gap in willingness to compete over time

Table 1 shows descriptive statistics of average choices in the two experiments over all rounds by gender. In both experiment, men scored slightly higher compared to women, although this difference is not statistically significant. Men are significantly more confident and more likely

Table 1: Descriptive statistics

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Main experiment				Feedback experiment			
	Scale	All	Male	Female	P-val	All	Male	Female	P-val
Score	0-26	7.30	7.54	7.07	0.243	10.29	10.65	9.88	0.141
		(2.99)	(3.20)	(2.76)		(3.92)	(4.23)	(3.50)	
Confidence	0-1	0.59	0.63	0.55	0.012	0.59	0.64	0.53	0.001
		(0.25)	(0.24)	(0.26)		(0.26)	(0.25)	(0.26	
Choosing competition	binary	0.47	0.53	0.41	0.046	0.47	0.53	0.40	0.026
		(0.50)	(0.50)	(0.49)		(0.50)	(0.50)	(0.49)	
Earnings	0-13	2.37	2.56	2.18	0.109	3.44	3.82	3.02	0.024
		(2.02)	(2.21)	(1.80)		(2.94)	(3.19)	(2.58)	
Risk taking	0-10	5.5	6.0	5.0	0.006	5.4	5.6	5.1	0.282
		(2.45)	(2.33)	(2.47)		(2.76)	(2.68)	(2.84)	
Competitiveness	0-10	6.7	7.1	6.3	0.034	6.8	7.2	6.4	0.035
		(2.39)	(2.20)	(2.51)		(2.56)	(2.52)	(2.54)	
N		188	93	95		184	97	87	

Note: The table shows averages over all rounds. Confidence is the probability of winning as predicted by the participants themselves at the start of each round. Earnings are per-round earnings in Euros. Risk-taking and competitiveness are self-rated questionnaire measures. Columns 4 and 8 show p-values from t-tests of the gender difference.

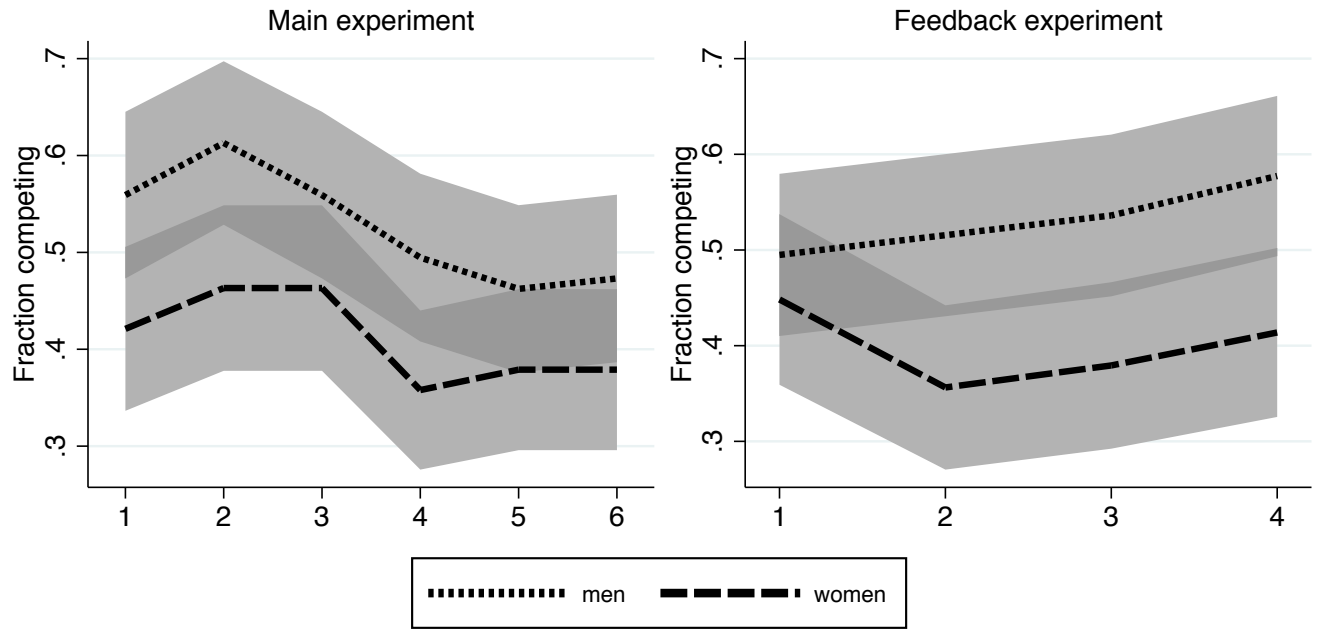
to choose competition over piece rate both in the main and in the feedback experiment.[4] The combined earnings of men are slightly higher than the earnings of women in both experiments. In the questionnaires, men judge themselves to be more risk-taking and more competitive.

Figure 1 shows the proportion of participants who choose the competition in each round separately by gender. In the main experiment, men are more likely than women to choose competition over piece rate in all rounds. In round 1, men are 14 percentage points more likely to compete than women (p=0.06, chi-squared test). This gender gap increases slightly to 15 percentage points in round 2 (p=0.04), shrinks to 10 percentage points in round 3 (p=0.19) and increases again to 14 percentage points in round 4 (p=0.06). In the last round, men are still 9 percentage points more likely to compete (p=0.19).

The gender difference in the likelihood of choosing to compete in the feedback experiment, where non-competers learn whether they would have won or lost had they chosen competition, is initially only 5 percentage points (p=0.53, chi-squared test). But this grows to around 16 percentage points in round 2 (p=0.03) and stays roughly constant in rounds 3 and 4 (p=0.03 in each round). A potential reason for the small initial gender gap is that, contrary to the main experiment (and contrary to the standard design in the literature on gender differences in competitiveness), it is no longer possible to avoid relative feedback by staying away from the competition. Women have

[4]Confidence corresponds to the probability of winning as predicted by the participants themselves and is based on the rank predictions of the participants. It is calculated as (predicted # of participants with lower score)/(# of participants-1).

Figure 1: Willingness to compete by gender and round



Note: Shaded areas represent 90-percent confidence intervals.

been found to be more feedback averse than men on average (Möbius et al., 2011) which might encourage choosing the piece rate if this allows to avoid relative feedback.

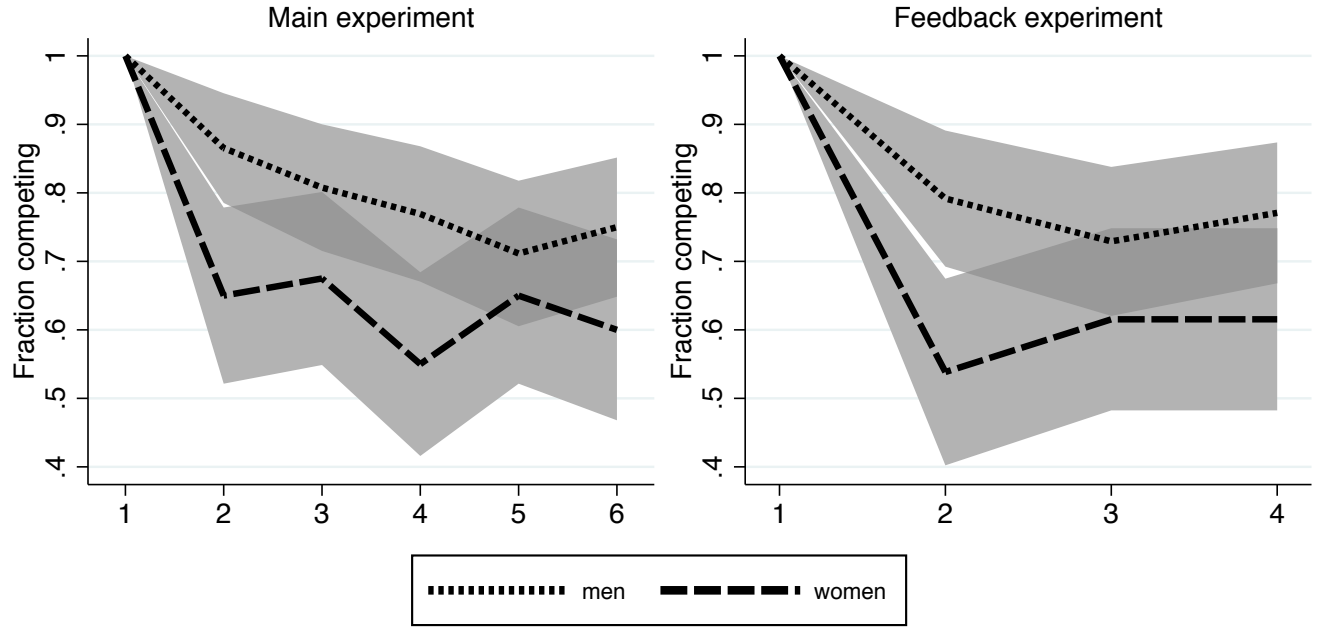
Figure 2 shows the likelihood of competing over all rounds for men and women who choose competition in round 1. Strikingly even for these people, who are initially willing to compete, a large gender difference appears in round 2 of the main experiment. 87 percent of men who compete in round 1 still compete in round 2 whereas for women this number is only 65 percent (p=0.02). By round 4, 77 percent of men and 55 percent of women who initially choose competition still compete (p=0.03). In the last round, men in this group are still 15 percentage points more likely to compete (p=0.13).

The pattern in the feedback experiment looks remarkably similar. A large and statistically significant gender difference appears in round 2. In percentage terms, 79 percent of men who compete in round 1 still compete in round 2, whereas for women this is only 54 percent (p=0.01). By round 4, 77 percent of men and 62 percent of women who initially choose competition still compete (p=0.12).

Figure 3 shows the likelihood of competing for men and women who choose the piece rate in round 1. In the main experiment, a similar proportion of men and women in this group switch to competition in round 2 (29 percent and 33 percent, p=0.72). Over the following rounds, women are slightly but not significantly more likely to choose competition than men (the difference ranges from 3 to 10 percentage points).

In the feedback experiment, despite the fact that those who outperform their opponent in round

Figure 2: Willingness to compete by gender and round (participants who choose competition in round 1)



Note: Shaded areas represent 90-percent confidence intervals. The sample consists of those participants who choose competition in round 1. Main experiment: N=92 (40 women and 52 men). Feedback experiment: N=87 (39 women and 48 men).

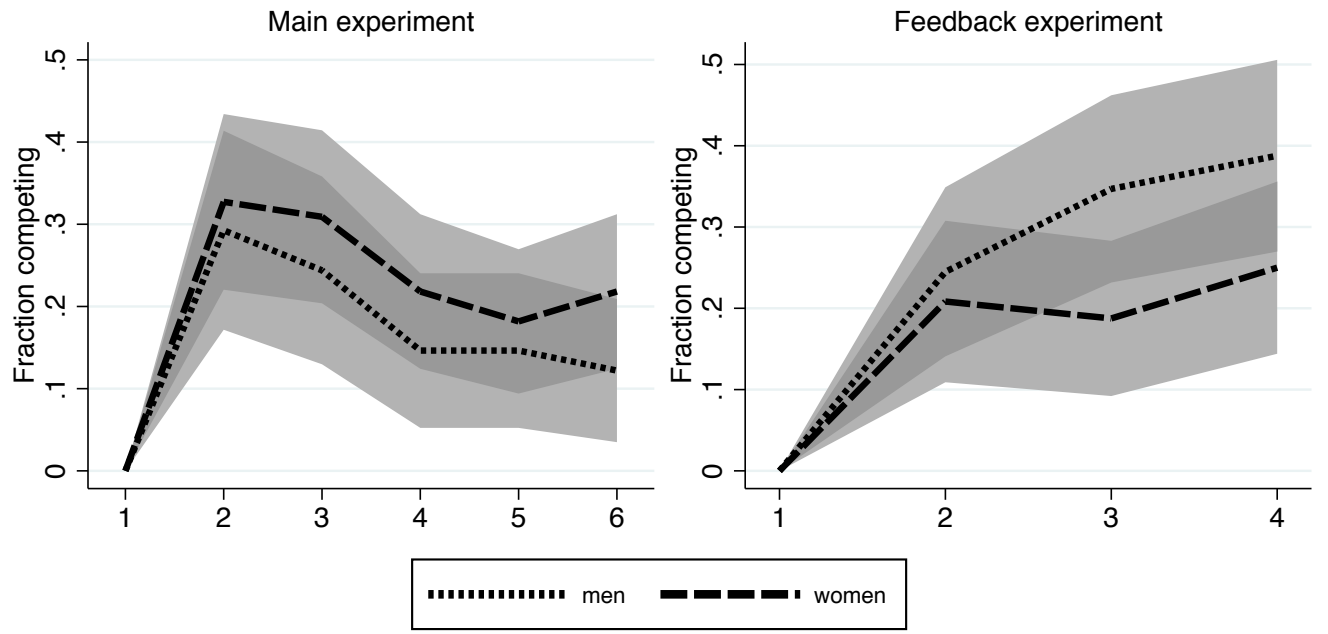
1 receive positive feedback, a smaller proportion than in the main experiment decide to enter in round 2. Furthermore, the pattern we observe from round 3 onwards looks very different. First, a gender gap appears in round 3 with 35 percent of men and 19 percent of women choosing to compete (p=0.08). Second, whereas in the main experiment competition entry steadily declines for this group after its peak in round 2, here it keeps increasing in rounds 3 and 4. A likely explanation for these differences is that, lacking feedback, more low-performing people choose to enter competition after round 1 in the main experiment, whereas in the feedback experiment only those who receive positive feedback choose to enter. This is confirmed by the data. The mean chance of winning (calculated based on round 1 scores) of those who choose to enter in round 2 is 61 percent in the feedback experiment compared to 48 percent in the main experiment.[5]

In Table 2, we repeat the analysis in Figures 1 to 3 using OLS regressions controlling for absolute performance and chance of winning.[6] We use absolute performance and chance of winning in round 1 as controls for all rounds. Because the choice of remuneration scheme might itself affect

[5]Furthermore in the main experiment, a mere 30 percent of those who start with choosing the piece rate but have a higher than 50 percent chance of winning choose competition by round 4 (32 percent by round 6). With feedback for non-competers, the corresponding figure is 74 percent. The biggest difference between the two experiments appears for men who initially choose the piece rate. Whereas in the the main experiment, 15 percent of these men enter competition in round 4, this is now 39 percent (for women, the corresponding numbers are 22 and 25 percent).

[6]Chance of winning is calculated as (# of participants with lower score)/(# of participants-1) and therefore corresponds to rank normalised by the number of subjects in a session.

Figure 3: Willingness to compete by gender and round (participants who choose piece rate in round 1)



Note: Shaded areas represent 90-percent confidence intervals. The sample consists of those participants who choose piece rate in round 1. Main experiment: N=96 (55 women and 41 men). Feedback experiment: N=97 (48 women and 49 men).

performance levels (for example, the same participant might try harder in those rounds where she chooses to compete) and because performance may be affected by earlier competition outcomes, controlling for later-round scores could bias the estimate of the gender coefficient.[7] We explore the importance of changes in relative performance over the rounds (and gender differences therein) as a mechanism behind our results in Section 5.

Columns 1 to 6 of Panel A show regressions of the remuneration choice in each of rounds 1 to 6 of the main experiment on gender. The regression results show that the gender difference in choosing the competition is robust to controlling for performance, meaning that gender differences in initial performance cannot explain gender differences in willingness to compete over the rounds. The gender difference is 14 percentage points initially and fluctuates between 7 and 15 percentage points over rounds 2 to 6.

Panel A of Table 2 also shows separate regressions for the gender gap in rounds 2 to 6 for those who choose competition in round 1 and those who choose piece rate in round 1. The regressions confirm that a significant gender gap appears from round 2 onwards for those who initially compete.

[7]When using round 1 performance as a control, the problem of the remuneration scheme possibly affecting performance still applies in the case of the regressions which use the whole sample (which contains both people who compete in round 1 and people who choose the piece rate in round 1). However, the problem is completely avoided in case of the regressions using the subsamples of those who compete in round 1 and those who choose the piece rate in round 1.

In round 2, a woman who chooses competition in round 1 is 21 percentage points less likely to compete again than a man with the same round 1 choice, same absolute performance, same chance of winning, and same competition outcome in round 1. The gender difference fluctuates between 12 and 22 percentage points over the following rounds and is still 15 percentage points in the last round. The pooled regression in column 7 shows that the average woman who chooses to compete in round 1 is 15 percentage points less likely to compete in each of the following 5 rounds compared to the average man who competes in round 1. For those who initially choose the piece rate, the gender difference is close to zero in all subsequent rounds controlling for absolute and relative performance .

Panel B repeats the analysis using data from the feedback experiment. Columns 1 to 4 show regression results for rounds 1 to 4. Column 5 repeats the analysis for round 4 while additionally controlling for the competition outcomes in rounds 2 and 3. The coefficient therefore shows the difference in tournament entry in round 4 between men and women who have the same initial performance and rank and who received the same feedback in rounds 1 to 3. The estimate is more precise but the magnitude is similar, further confirming that gender differences in performance cannot explain the emerging gender gap in willingness to compete.

For those who initially choose competition, the pooled regression in column 6 shows that the average woman who competes in round 1 is 18 percentage points less likely to compete over each of rounds 2 to 4 than the average man who competes in round 1, conditional on absolute and relative performance and the competition outcome in round 1. For those who choose the piece rate in round 1, a gender gap of around 11 percentage points emerges in rounds 3 and 4 but this is only statistically significant at the ten-percent level in round 4 when additionally controlling for the received feedback in rounds 2 and 3 (column 5).

3.2 Gender difference in the effect of losing on subsequent willingness to compete

We will now investigate whether the larger proportion of women who drop out of competition and the resulting emergence of a gender gap for those who initially compete is due to a gender difference in the reaction to competition outcomes. In particular, we will answer the question whether women who initially choose to compete are more likely than their male counterparts to stop competing following a loss.

Figure 4 shows how often on average participants choose to compete in rounds 2 to 6, split by their gender, their choice of remuneration scheme in round 1, and whether they won or lost in round 1. The left-most panel shows competition rates over rounds 2 to 6 of the main experiment for men and women who choose to compete in round 1. Participants of both genders who win the competition in round 1 mostly keep competing over the following five rounds (4.6 out of 5 times for men and 4.2 out of 5 times times for women; p=0.17, t-test). However, the reaction to losing

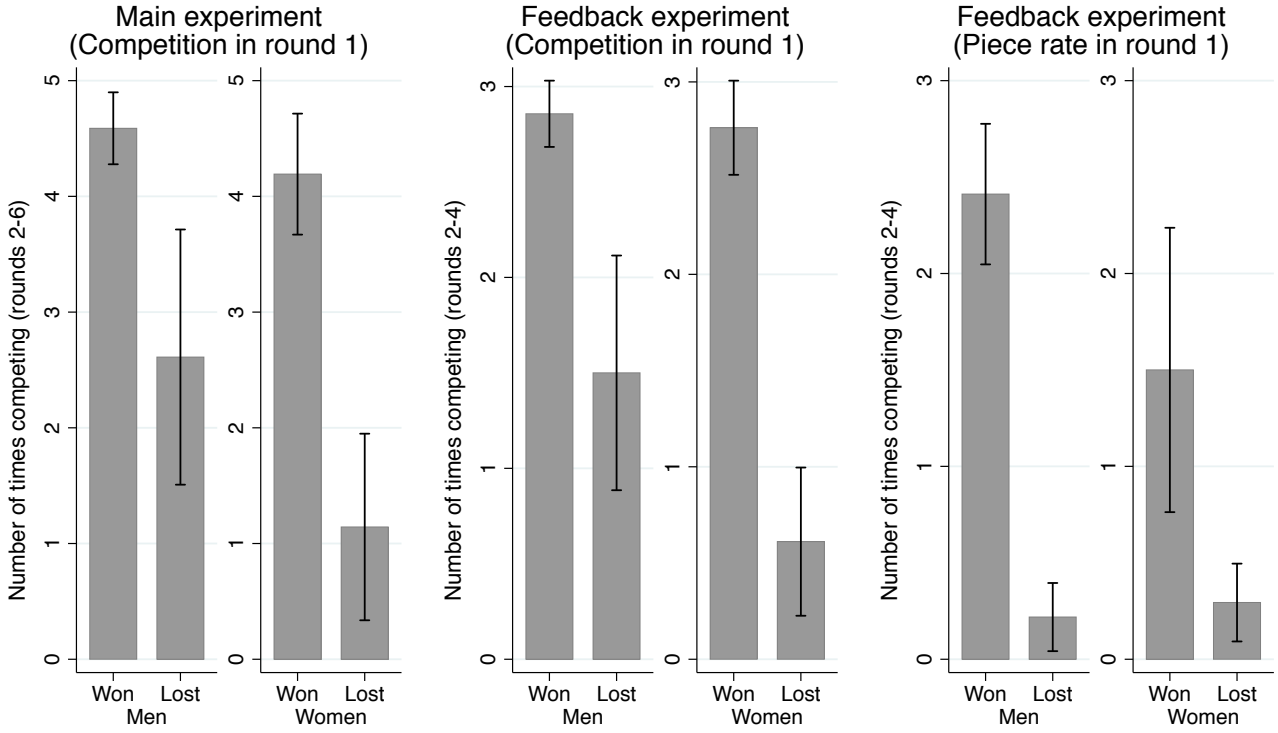
Table 2: Gender gap in competition entry in each round

		Panel A: Main experiment						
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Pooled (2-6)
All	Female	-0.135*	-0.146**	-0.091	-0.125*	-0.068	-0.085	-0.103**
		(0.071)	(0.066)	(0.067)	(0.068)	(0.063)	(0.064)	(0.051)
	N	188	188	188	188	188	188	940
Comp. in round 1	Female		-0.214***	-0.117	-0.215**	-0.043	-0.152*	-0.148**
			(0.065)	(0.081)	(0.095)	(0.086)	(0.085)	(0.059)
	N		92	92	92	92	92	460
PR in round 1	Female		-0.023	0.014	0.041	-0.021	0.030	0.008
			(0.100)	(0.093)	(0.081)	(0.079)	(0.083)	(0.058)
	N		96	96	96	96	96	480
	R1 score	√	√	√	√	√	√	√
	R1 rank	√	√	√	√	√	√	√
	R1 outcome		(√)	(√)	(√)	(√)	(√)	(√)

		Panel B: Feedback experiment					
		(1)	(2)	(3)	(4)	(5)	(6)
		Round 1	Round 2	Round 3	Round 4		Pooled (2-4)
All	Gender	0.012	-0.101**	-0.110*	-0.115*	-0.124**	-0.109**
		(0.069)	(0.049)	(0.059)	(0.060)	(0.049)	(0.048)
	N	184	184	184	184	184	552
Comp. in round 1	Female		-0.255***	-0.113	-0.185**	-0.173**	-0.184***
			(0.068)	(0.090)	(0.083)	(0.070)	(0.059)
	N		87	87	87	87	261
PR in round 1	Female		0.010	-0.119	-0.101	-0.129*	-0.086
			(0.065)	(0.083)	(0.085)	(0.071)	(0.061)
	N		97	97	97	97	291
	R1 score	√	√	√	√	√	√
	R1 rank	√	√	√	√	√	√
	R1 outcome		√	√	√	√	√
	Outcomes R2-3					√	

Note: The table shows coefficients from OLS regressions of a gender dummy on a binary indicator for choosing the competition. Each line is a separate regression. R1 score, R1 rank and R1 outcome means score, normalised within-session rank and the competition outcome in round 1. In Panel A, only the regressions for those who compete in round 1, and therefore receive feedback, control for the round 1 outcome. All regressions control for session fixed effects. Robust standard errors in parentheses. Standard errors for the pooled regressions are clustered at the individual level.

Figure 4: Average number of times competition is chosen after round 1 by gender, initial choice and competition outcome in round 1



Note: The bars show the average number of times that participants chose to compete over rounds 2 to 6 (2 to 4 in case of the feedback experiment). "Lost" denotes participants who lost in round 1 and "won" denotes participants who won in round 1. Error bars represent 95% confidence intervals.

in round 1 differs strongly between men and women. Men who initially compete and lose compete a further 2.6 out of 5 times whereas for women this is only 1.1 times (p=0.04).

The centre panel repeats this analysis for the feedback experiment. The pattern is remarkably similar to the main experiment. Both men and women essentially keep competing if they win in round 1 (2.9 out of 3 times for men and 2.8 out of 3 times for women; p=0.50, t-test). But women who lose are again much less likely to compete over the following rounds than men who lose (1.5 out of 3 times for men and 0.6 out of 3 times for women; p=0.02).

Giving feedback to non-competers in the feedback experiment allows us to also test whether there is a gender difference in the reaction to positive feedback for those who are initially not willing to compete. The right-most panel of Figure 4 shows the average number of times participants choose competition over rounds 2 to 4 for those who choose piece rate in round 1, separately by gender and the feedback received at the end of round 1. There is no gender difference for those who perform worse than their opponent in round 1: both men and women essentially keep choosing the piece rate (men compete 0.2 out of 3 times and women 0.3 out of 3 times; p=0.57). But men who outperform their opponent compete more often over the following rounds than women who outperform their opponent (2.4 out of 3 times vs 1.5 out of 3 times; p=0.02).

In Table 3, we use OLS regressions to confirm the statistical significance and robustness of the gender difference in the reaction to losing. Here, we regress a binary indicator for choosing to compete on a gender dummy, a binary indicator for having lost in round 1, and the interaction of the two. Each individual decision in rounds 2 to 6 is a separate observation and standard errors are clustered at the individual level.

The regressions control for score fixed effects and within-session rank (likelihood of winning) in round 1. It is important to note that conditional on score fixed effects, winning or losing in round 1 only depends on the score of the randomly allocated opponent and can consequently be interpreted as a random treatment. The regressions therefore give us causal estimates of the effect of losing relative to winning in round 1 on the likelihood of competing in the subsequent rounds.

The regression in column 1 of Table 3 uses the sample of participants in the main experiment who choose to compete in round 1. The results show that the gender difference in the reaction to losing is robust to controlling for absolute and relative performance. Whereas male losers are 24 percentage points less likely to enter competition in rounds 2 to 6 compared to winners, this effect is more than twice as large for women at 59 percentage points.

It is interesting to check whether the gender difference in the reaction to losing applies also to high-performing individuals who have a positive expected return from competing. If this is the case, losing early on is more costly for high-performing women than for high-performing men. In column 2, we restrict the sample to those participants who have a higher than 50 percent chance of winning.[8] High-performing men hardly react to losing in round 1 but high-performing women who

[8]Chance of winning is calculated within session based on first-round performance as (# of participants with lower score)/(# of participants-1).

Table 3: Effect of competition outcome in round 1 on subsequent competition entry

	(1)	(2)	(3)	(4)		(5)	(6)
	Main experiment		Feedback experiment			Feedback experiment (PR)	
	All	Top	All	Top		All	Top
Female	-0.065	-0.132**	-0.028	-0.028	Female	0.028	0.213
	(0.056)	(0.053)	(0.073)	(0.073)		(0.049)	(0.161)
Round 1 loser	-0.240*	-0.149	-0.284**	-0.171	Round 1 winner	0.564***	0.645***
	(0.126)	(0.156)	(0.124)	(0.145)		(0.092)	(0.109)
Female x loser	-0.349**	-0.409**	-0.273**	-0.332*	Female x winner	-0.287**	-0.415**
	(0.144)	(0.186)	(0.129)	(0.194)		(0.115)	(0.196)
Score fixed effects	√	√	√	√		√	√
Round 1 rank	√	√	√	√		√	√
Observations	460	275	261	177		291	96
Individuals	92	55	87	59		97	32

Note: The table shows coefficients from OLS regressions of a gender dummy, a dummy for having lost the round 1 competition and the interaction of the two on a binary indicator for choosing the competition in rounds 2 to 6 (2 to 4 in case of the feedback experiment). Each choice is a separate observation and standard errors are clustered at the individual level. The sample in columns 1 to 4 consists of participants who choose competition in round 1. The sample in columns 5 and 6 consists of participants who choose piece rate in round 1. The columns marked "Top" restrict the sample to participants who have a higher than 50 percent chance of winning based on their round 1 performance. Score fixed effects and round 1 rank mean score and normalised within-session rank in round 1. Clustered standard errors in parentheses.

lose are more than 55 percentage points less likely to choose competition again over the subsequent rounds compared to high-performing women who win.

In columns 3 and 4, we repeat this analysis for the feedback experiment. The regression results in column 3 confirm the gender difference in the reaction to losing for those who initially compete. While male losers are 29 percentage points less likely to choose competition in round 2 to 4 compared to male winners, this effect is 57 percentage points for women. In column 4, we show that, again, the magnitude of the gender difference in the reaction to losing is even higher for high performers for whom not competing is costly.

The regression results in columns 5 and 6 show that the gender difference in the effect of positive feedback on future willingness to compete for those who are initially not willing to compete is significant conditional on performance. Whereas men who win are 57 percentage points more likely to compete over the following three rounds than men who lose, this effect is only 27 percentage points for women.

The analysis in this section focused on the reaction to the competition outcome in round 1. The reason for this focus is that in later rounds, the group of participants who choose to compete has heterogeneous prior experience. In particular, because women are more likely to drop out of competition than men, the women amongst the competers in later rounds are more selected on having won previously than are the men.

4 Results from the Dutch Math Olympiad

In this section, we will use data from the Dutch Math Olympiad to determine whether the gender difference in the reaction to losing carries over to the field. We will start by describing how the Math Olympiad works and why it provides a credible setting to identify the gender difference in the effect of losing a competition on subsequent willingness to compete.

4.1 Background, data and identification strategy

The Dutch Math Olympiad is an annual national mathematics competition organised for secondary school students. Its ultimate aim is the selection of a national team to represent the Netherlands in the International Mathematics Olympiad.[9] All Dutch secondary school students in the pre-university track up to grade five are allowed to participate in the first round of the Olympiad.[10] The first round of the competition consists of a two-hour test which is administered in the participants' own schools and graded by their teachers.[11]

Only the top 1000 performers in round one advance to the second round. Because the threshold score for advancing is flexibly determined to approximately select the top 1000 participants, the setting resembles a competition in which the top 1000 performers win and the rest loses. Participants who are in the fourth year of secondary school or lower are free to participate again the year after regardless of their score. We take the binary decision of whether to compete again one year later as our outcome measure.

We have anonymised data for all participants in the 2010-2014 Olympiads including their score and whether they participated again the year after. The vast majority of participants are from years four and five of the pre-university track of secondary school. We restrict our sample to fourth-year pre-university track students who are allowed to participate again one year later when they are in their fifth year. Because winning and losing depend on a strict cutoff, the data present us with a sharp regression discontinuity design. Comparing the subsequent participation choices of individuals just below and just above the cutoff makes it possible to estimate the causal effect of losing vs winning on the likelihood of participating again one year later. In particular, we are interested in the gender difference in this effect.

There are two commonly used approaches to estimating a regression discontinuity design: local

[9]The International Mathematics Olympiad is the most recognised international maths competition for pre-university students, with around a hundred countries participating. Details on the Dutch Math Olympiad can be found at www.wiskundeolympiade.nl.

[10]Secondary education in the Netherlands consists of three different tracks: vocational, general and pre-university. The six-year pre-university track is the most academically demanding, where the highest performing students end up. Because the national final takes place during the following school year, sixth-year students are not allowed to participate. Students from the general track of secondary school are also allowed to participate in the Olympiad but because the general track consists of only five years, students are only allowed to participate up to year four.

[11]There is no discretion in grading as the test consists of multiple choice questions and questions with a clear numerical answer and points are only awarded for a correct answer.

linear and polynomial (Lee and Lemieux, 2010). Local linear regression means restricting the sample to observations that are close to the cutoff and controlling linearly for the forcing variable.[12] The polynomial approach consists of using a higher bandwidth (that is, observations further away from the cutoff) and controlling for the forcing variable using a polynomial. In practice, this usually means using a quadratic control (Gelman and Imbens, 2014, advise against the use of higher order polynomials and show that confidence intervals based on them have poor properties). The local linear approach has the advantage of staying close to the spirit of regression discontinuity by only using observations close to the cutoff. However, in cases where a polynomial is a good approximation of the underlying data, using the polynomial approach is more efficient. The usual recommendation is to use a variety of specifications and bandwidths (Lee and Lemieux, 2010; Imbens and Lemieux, 2008; Athey and Imbens, 2016).

Our design is somewhat unusual as our main interest is not in the discontinuity itself (that is, the effect of losing on the likelihood of participating again a year later) but in the gender difference in the discontinuity (that is, the gender difference in the effect of losing on the likelihood of participating again a year later). In practice, we estimate the following equation:

$$Y_i = \alpha + \delta_1 T_i + \delta_2 F_i + \delta_3 T_i \cdot F_i + f(s) + T_i \cdot f(s_i) + \varepsilon_i$$

where Y is a binary indicator for participating again a year later, T is a binary indicator for scoring above the cutoff (that is, making the top 1000), F is a binary indicator for being female, and $f(s)$ is a polynomial function of the amount of points scored s. The parameter of interest is δ_3 which estimates the gender difference in the reaction to losing. We interact the polynomial $f(s)$ with the cutoff indicator T to allow for different slopes left and right of the cutoff.

4.2 Regression discontinuity results

Our sample consists of 10,004 individuals, which is the universe of fourth-year pre-university track students who participated in the Olympiad from 2010 to 2014. Of these, 62.4 percent are male, reflecting the lower willingness to compete of girls in mathematical areas (see for example Buser, Niederle, and Oosterbeek, 2014). 16.0 percent of male participants and 10.2 percent of female participants score above the cutoff (p=0.00; chi-squared test).[13] Conditional on scoring above the cutoff, the likelihood of actually participating in the second round is virtually identical for male and female participants (90.2 percent vs 89.1 percent; p=0.56). 50.7 percent of all male participants and 44.8 percent of all female participants (p=0.00) participate again the year after.

In Figure 5, we present results using both the local linear and polynomial approaches. Panel A shows the data close to the cutoff (plus-minus 4 points) and a linear regression line which is

[12] In our case, the forcing variable is the first-round score.

[13] Figure A1 in the appendix shows the full distribution of points for boys and girls where we normalised the scores such that a score of zero or above means passing the threshold.

Table 4: Regression discontinuity results

	(1)	(2)	(3)	(4)	(5)	(6)
Range (in points)	-/+ 4	-/+ 4	-/+ 4	-/+ 7	-/+ 10	-/+ 20
Polynomial	constant	linear	quadratic	quadratic	quadratic	quadratic
Female	0.025	0.026	0.024	0.023	0.029	0.025
	(0.034)	(0.036)	(0.036)	(0.026)	(0.025)	(0.024)
Lost	-0.098*	-0.006	0.057*	-0.026	0.001	-0.010
	(0.042)	(0.060)	(0.026)	(0.058)	(0.052)	(0.044)
Female x lost	-0.098**	-0.096**	-0.095**	-0.057*	-0.064**	-0.069***
	(0.037)	(0.039)	(0.039)	(0.031)	(0.027)	(0.025)
N	2489	2489	2489	4730	7459	9967

Note: The table shows coefficients from regressions of a binary indicator for participating again a year later on a female dummy, a dummy for not having made the top 1000 ("lost") and the interaction of the two. Range means the sample selection in terms of points left and right of the cutoff. Clustered standard errors in parentheses.

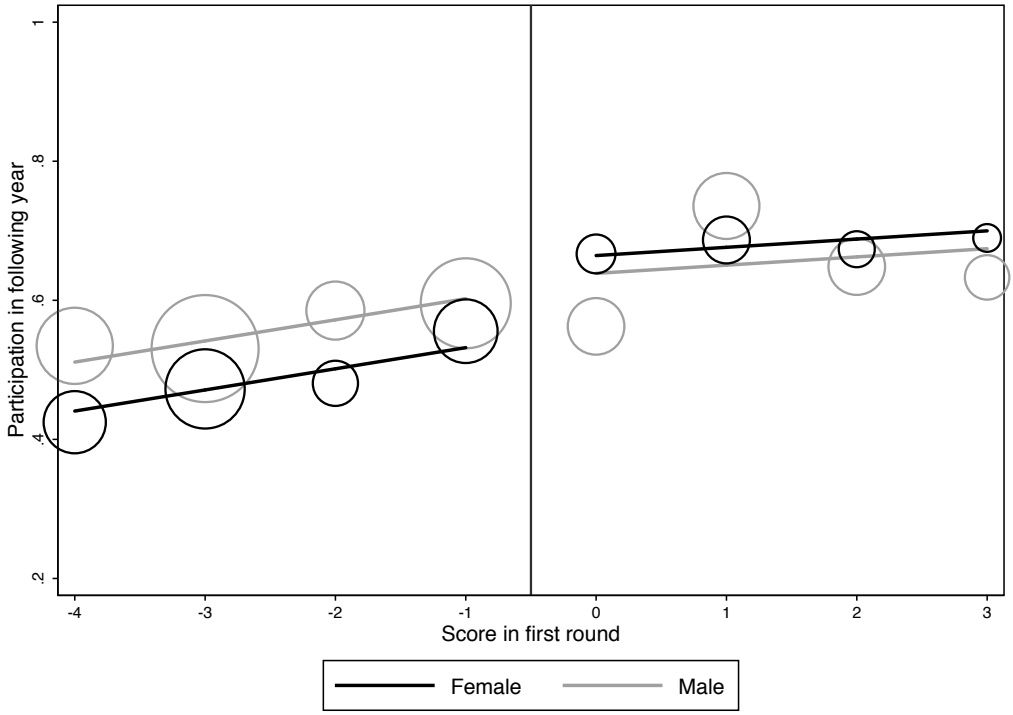
estimated separately below and above the cutoff. The first-round scores are normalised such that a score of zero or above means passing the threshold. The size of the markers is proportional to the number of observations. Panel B shows a wider bandwidth (plus-minus 10 points) and a quadratic approximation. In both cases it is evident that there is a sizeable drop at the cutoff in the likelihood of participating again for girls but not for boys.

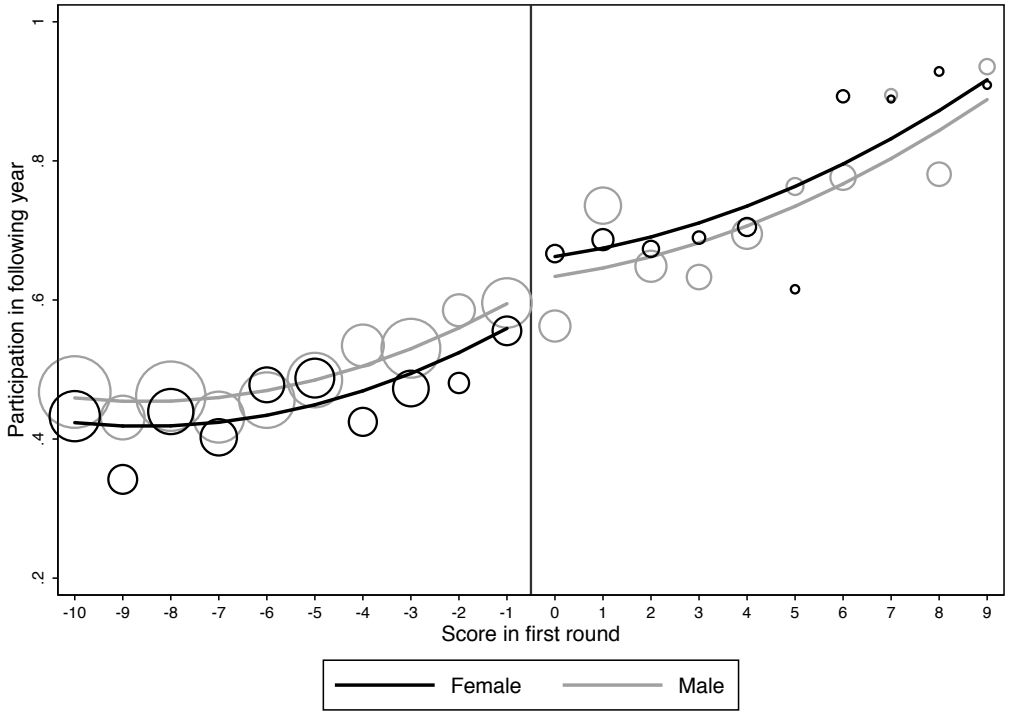
Table 4 shows regression results using bandwidths of 4, 7, 10 and 20 points around the cutoff, which roughly divides the sample into quartiles in terms of the number of observations. The regression in column 1 uses the local constant approach which does not control for the forcing variable. The coefficient on the interaction of the gender and losing dummies therefore simply compares the gender difference in the likelihood of participating again to the left of the cutoff (losers) to the gender difference to the right of the cutoff (winners).[14] The coefficient shows that girls are significantly more affected by losing than boys. The gender difference in the reaction to losing is equal to 10 percentage points and statistically significant. In columns 2 and 3, we keep the bandwidth constant and add linear and quadratic controls of the forcing variable. The estimate of the interaction coefficient hardly changes, staying close to 10 percentage points. In columns 4 to 6, we gradually increase the bandwidth while using a quadratic control for the forcing variable. The estimates for the gender difference in the reaction to losing are somewhat smaller at 6 to 7 percentage points but still statistically and economically significant. In the appendix, we present a full table with all bandwidth-polynomial combinations. The estimate of the gender difference in the reaction to losing is remarkably robust and always statistically significant.

These results show that the gender difference in the reaction to losing, which we observed in the lab, carries over to the field. The Math Olympiad provides a highly relevant setting as past research has shown that willingness to compete in numerical tasks predicts specialising in a science,

[14] In all regressions, we follow the recommendation of Lee and Lemieux (2010) for regression discontinuity with a discrete assignment variable and cluster the standard errors at the score level.

Figure 5: Regression discontinuity





technology, engineering or mathematics (STEM) related field (Buser, Niederle, and Oosterbeek, 2014). Girls being faster to give up on competing in math after a setback could therefore be a partial explanation for the low number of women who opt for a career in STEM fields. The regression discontinuity results also demonstrate that the effect of losing persists in the long term. A full year after their first participation, girls who lose are still significantly less likely to participate again.

5 The role of gender differences in risk preferences, confidence, belief updating, and performance

The results from the main and feedback experiments and from the Math Olympiad show that a gender gap in willingness to compete appears with experience even for individuals who initially choose to compete. At the origin of this gender gap is a strong gender difference in the reaction to losing and winning. The experimental economics literature on gender differences in competitiveness suggests a number of potential mechanisms behind the gender difference in the reaction to competition outcomes, namely gender differences in risk preferences, initial confidence, belief updating, and the evolution of relative performance over time. Our experimental data allow us to explore all four of these.

Risk

Women are often found to be more risk averse than men (Croson and Gneezy, 2009; Charness and Gneezy, 2012) and recent papers by Gillen, Snowberg, and Yariv (2015) and van Veldhuizen (2016) demonstrate that gender differences in risk preferences play an important role in explaining the gender gap in tournament entry in a static setting. In particular, van Veldhuizen (2016) lets participants choose between piece-rate remuneration and a risky option where winning and losing is determined by a lottery. He finds a large and significant gender gap in choosing the risky option.

In this section, we will use the data from our risk experiment to explore whether changes in risk preferences can explain the dynamic patterns in willingness to compete observed in the main and feedback experiments. The setting of the risk experiment is identical but winning and losing are determined by a random draw. This allows us to investigate whether the main result – women are more likely than men to stop competing after losing – occurs also in an environment that is characterised by the same risks but which lacks the competitive dimension.[15]

[15]Over all rounds of the risk experiment, men performed slightly but not significantly better than women and earned slightly but not significantly more. Again, men judged themselves to be more risk-seeking and competitive. However, men were not significantly more likely to choose the risky option compared to women over all six rounds. Scores: 7.4 (men) vs 7.0 points (women); p=0.21, t-test. Earnings: 2.0 vs 1.8; p=0.31. Risk taking: 5.6 vs 4.5; p=0.00. Competitiveness: 7.4 vs 6.8; p=0.06. Choosing the risky option: 0.38 vs 0.34; p=0.48.

Table 5: Gender gap in choosing the risky option in each round

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Pooled (2-6)
All	Female	-0.207***	-0.027	-0.116*	0.041	-0.107*	0.014	-0.039
		(0.058)	(0.064)	(0.063)	(0.064)	(0.061)	(0.055)	(0.041)
	N	188	188	188	188	188	188	940
Risky in round 1	Female		-0.049	-0.113	-0.226*	-0.180	-0.168	-0.147*
			(0.126)	(0.114)	(0.122)	(0.124)	(0.119)	(0.083)
	N		62	62	62	62	62	310
Safe in round 1	Female		0.122*	0.025	0.277***	0.085	0.218***	0.145***
			(0.071)	(0.071)	(0.063)	(0.070)	(0.050)	(0.034)
	N		126	126	126	126	126	630
	R1 score	√	√	√	√	√	√	√
	Prob. FE	√	√	√	√	√	√	√
	R1 outcome		(√)	(√)	(√)	(√)	(√)	(√)

Note: The table shows coefficients from OLS regressions of a gender dummy on a binary indicator for choosing the risky option. Each line is a separate regression. R1 score and R1 outcome means score and lottery outcome in round 1. Only the regressions for those who choose the risky option in round 1, and therefore receive feedback, control for the round 1 outcome. All regressions control for session fixed effects. Probability fixed effects means dummy variables indicating the probability of winning the lottery (randomly fixed at 0.3, 0.4, 0.5, 0.6 or 0.7). Robust standard errors in parentheses. Standard errors from the pooled regressions are clustered at the individual level.

Table 5 reproduces the regressions in Table 2 using the data from the risk experiment, showing the gender difference in the likelihood of choosing the risky option in each round conditional on performance and exogenous probability of winning. Whereas men are 21 percentage points more likely to choose the risky option in round 1, this gender gap is smaller and not consistently significant over the remaining rounds. Looking separately at participants who choose the risky and safe options in round 1 reveals differences with the patterns observed in the main and feedback experiments. For those who choose the risky option in round 1, no significant gap appears in round 2. However, a gap starts appearing from round 3 onwards and women who start out choosing the risky option are significantly less likely than men to do so again over the next five rounds. The pattern for those who choose the safe option in round 1 also looks different from the patterns in the other experiments. From round 2 onwards, women in this group are substantially more likely to choose the risky option than are men in this group.

In Table 6, we repeat the analysis in Table 3 and use OLS regressions to estimate the gender difference in the reaction to losing conditional on probability of winning and performance. Here, we regress a binary indicator for choosing the risky option in rounds 2 to 6 on a gender dummy, a binary indicator for having lost in round 1, and the interaction of the two. The regressions therefore give us the gender difference in the reaction to losing, conditional on the exogenous probability of winning and performance. The coefficient on the gender interaction is very close to

Table 6: Choosing the risky option in rounds 2 to 6

	(1)	(2)
	All observations:	Chance of winning ≥ 0.5:
Female	-0.194*	-0.220*
	(0.112)	(0.123)
Round 1 loser	-0.098	-0.122
	(0.085)	(0.090)
Female x loser	0.011	0.024
	(0.167)	(0.163)
Round 1 score	√	√
Probability FE	√	√
Observations	310	275
Individuals	62	55

Note: The table shows coefficients from OLS regressions of a gender dummy, a dummy for having lost the round 1 random draw and the interaction of the two on a binary indicator for choosing the risky option in rounds 2 to 6. Each choice is a separate observation and standard errors are clustered at the individual level. The sample consists of participants who choose the risky option in round 1. Probability fixed effects means dummy variables indicating the probability of winning the lottery (randomly fixed at 0.3, 0.4, 0.5, 0.6 or 0.7). Clustered standard errors in parentheses.

zero. Women are no more likely than men to stop choosing the risky option if they lose in round 1. This demonstrates that the gender difference in the reaction to competition outcomes observed in the previous two experiments does not obtain in an environment that is characterised by the same risk structure but lacks the competitive elements (competing and winning or losing against someone else's performance). This indicates that gender differences in the dynamic evolution of risk preferences are unlikely to be a mechanism behind the gender difference in the reaction to losing in a competition.

Beliefs and performance

We can use the belief-elicitation and performance data from the main and feedback experiments to check whether gender differences in initial confidence, belief updating, or the evolution of relative performance are mechanisms behind the gender difference in the reaction to competition outcomes. Initial confidence is measured by the participants' predicted rank at the start of round 1, belief updating is measured by the change in predicted rank over time, and relative performance is measured by the change of actual rank over time. These mechanisms are explored in Table 7 using OLS regressions. Column 1 replicates the analyses in columns 1, 3 and 5 of Table 3. We then control for each potential mechanism separately in columns 2 to 4. If a mechanism has explanatory power for the gender difference in the reaction to competition outcomes, the magnitude of the gender-interaction coefficient should shrink.

Most studies on gender differences in competitiveness find that, conditional on performance, women tend to be less confident than men (Niederle and Vesterlund, 2011), which is also the case

in our sample.[16] A loss could therefore more easily push women's confidence below the threshold at which they are no longer willing to compete. This is explored in column 2 by controlling for confidence at the start of round 1 and its interaction with a dummy for having lost(won) the first-round competition.[17] The gender difference in the reaction to losing for those who initially compete is hardly affected. For the main experiment, the coefficient actually increases in magnitude. However, the gender difference in the reaction to positive feedback shrinks by 39 percent for those who initially choose the piece rate in the feedback experiment.

If women update their beliefs more strongly in response to losing than men, this could likewise explain why women's willingness to compete drops more strongly after a loss than men's willingness to compete. This is explored in the regressions in column 3 which control for confidence at the start of each round.[18] For those who initially compete, the magnitude of the coefficient on the gender interaction actually increases slightly, meaning that men update their beliefs more strongly in response to losing than women.[19] For those who initially pick the piece rate, the gender difference in the reaction to positive feedback drops by 36 percent.

Past studies have shown that women might perform worse after losing in a competition compared to men in the same situation (Buser, 2016; Gill and Prowse, 2014). In column 4, we control for normalised rank (that is, likelihood of winning) in each round.[20] The interaction coefficient does not change much in any of the regressions and the gender difference in the effect of losing actually increases in both datasets.

Finally, we combine all three mechanisms in column 5. It is clear that gender differences in confidence, belief updating, and performance changes cannot explain the gender difference in the reaction to losing for those who initially choose to compete. Conditional on initial beliefs, belief updating, and the evolution of relative performance, the magnitude of the gender-interaction coefficient actually increases in both datasets. On the other hand, these mechanisms explain an important part of the gender difference in the reaction to positive feedback for those who do not compete in round 1 of the feedback experiment, mainly due to the explanatory power of gender

[16]Regressing initial confidence in the main experiment (predicted probability of winning in round 1) on gender and performance in round 1, the coefficient on the female dummy is -0.081 (0.032) meaning that conditional on performance the average woman thinks that it is around 8 percentage points less likely that she would win a competition against a randomly selected opponent than the average man.

[17]The likelihood of entering competition is not necessarily linear in confidence. In particular, a risk neutral person would compete if they think their chance of winning is above 50 percent and would choose the piece rate otherwise. To make the regressions more flexible, we also include a dummy indicating that the predicted chance of winning is above 50 percent and include full interactions of this dummy with the continuous confidence variable and the competition outcome dummy.

[18]Again, to make the regressions more flexible we also include a dummy indicating that the predicted chance of winning is above 50 percent and include full interactions of this dummy with the continuous confidence variable and the competition outcome dummy.

[19]This fits with the results of Möbius et al. (2011) and Buser, Gerhards, and van der Weele (2016) who find that men update beliefs about their own abilities more strongly than women in response to feedback.

[20]Again, to make the regressions more flexible, we also include a dummy indicating that the chance of winning is above 50 percent and include full interactions of this dummy with the continuous rank variable and the competition outcome dummy.

Table 7: Competition entry after round 1: the role of confidence, belief updating and performance in explaining the gender difference in the reaction to competition outcomes

	(1)	(2)	(3)	(4)	(5)
	Competition in round 1 (main experiment)				
Female x round 1 loser	-0.349**	-0.386***	-0.354***	-0.373**	-0.424***
	(0.144)	(0.134)	(0.128)	(0.142)	(0.129)
N	460	460	460	460	460
	Competition in round 1 (feedback experiment)				
Female x round 1 loser	-0.273**	-0.251*	-0.289**	-0.300**	-0.305**
	(0.129)	(0.143)	(0.111)	(0.130)	(0.124)
N	261	261	261	261	261
	Piece rate in round 1 (feedback experiment)				
Female x round 1 winner	-0.287**	-0.176	-0.183*	-0.253**	-0.114
	(0.115)	(0.117)	(0.108)	(0.107)	(0.101)
N	291	291	291	291	291
R1 conf x Win R1		√			√
Confidence			√		√
Rank				√	√
R1 score fixed effects	√	√	√	√	√
R1 rank	√	√	√	√	√

Note: The table shows coefficients from OLS regressions of a gender dummy, a dummy for having won the round 1 competition and the interaction of the two on a binary indicator for choosing the competition in the subsequent rounds. R1 score means score in round 1, R1 rank means normalised within-session rank in round 1, R1 conf means probability of winning in round 1 as predicted by the participant, confidence means confidence at the start of each round and rank means rank in each round. Standard errors in parentheses. Standard errors are clustered at the individual level.

differences in initial beliefs and belief updating.

We conclude that gender differences in confidence, belief updating, performance changes and changes in risk preferences cannot explain the gender difference in the reaction to losing. Rather, it seems that losing in a competition has a direct negative impact on the preference for competition of women who are initially willing to compete.

6 Conclusions

Differences in willingness to compete between men and women of equal talent are well documented in the experimental economics literature. Moreover, gender differences in willingness to compete have been shown to partially explain gender differences in career choices and labour market earnings. This paper asks what happens to the gender gap in competitiveness over time as people accumulate experience. The evidence we present shows that it is far from certain that the type of noisy feedback people typically accumulate over their careers serves to eliminate differences in willingness to compete between equally well-performing men and women.

In particular, we find that men are not only more likely to compete in the first place but they are also more likely to keep competing after losing. This leads to a significant gender difference in willingness to compete even among those individuals who start out competing. Using field data from the Dutch Math Olympiad, we not only show that the gender difference in the effect of losing on subsequent willingness to compete carries over to the field but also that the effect persist in the long term. One year after participating in the Math Olympiad, female losers are significantly less likely to enter again, an effect that is much weaker for male losers.

If these findings translate to career settings, they could be relevant for explaining gender differences in career choices and outcomes. In particular, the gender difference in the reaction to losing amongst those who are initially willing to compete is a plausible explanation for the so-called "leaking pipeline". In European universities, women make up 59 percent of graduates, 46 percent of PhDs, 37 percent of associate professors and only 20 percent of full professors. In science and engineering women make up 35 percent of PhD graduates, 23 percent of associate professors, and just 11 percent of full professors (European Commission, 2013). Consistent with an explanation based on gender differences in the reaction to competition outcomes, Fischer (2015) finds that women, but not men, are more likely to drop out of a STEM major if they are grouped with higher ability peers. The patterns in the world of business resemble those in academia. The higher up the corporate ladder one goes, the smaller the share of women becomes (Bertrand and Hallock, 2001). Bertrand, Goldin, and Katz (2010), using a sample of MBA graduate from a top business school, show that although male and female graduates earn similar salaries right after graduation, men earn almost twice as much only a decade later.

Considering the fierce competition for journal space, research funding and academic positions, persisting in the face of failure is arguably crucial for being successful in academia. Similarly, it is

difficult to imagine somebody making it to the top of the corporate hierarchy without losing out to competitors at some point on the way. Our experimental results show that the gender difference in the reaction to competition outcomes is at least as strong for high performers who are initially willing to compete. These are exactly the people we would expect to become future full professors and executives. This suggests that bad luck early on in a career could be especially costly for high-performing women and introduces an important element of path dependence in the careers of women.

The gender difference in the reaction to losing, which we observe in two separate experiments and in the field, does not occur in an additional experiment which replicates the risk structure of the other experiments but lacks their competitive dimension. This strongly implies that gender differences in risk preferences cannot explain the gender difference in the reaction to losing. We also collect detailed data on participants' beliefs about their own relative performance and use it to exclude gender differences in confidence and belief updating as a significant mechanism behind our findings. This suggests that there is a direct effect of losing on the preference for competition and that this effect is stronger for the average woman than the average man. The psychology literature on gender differences in the reaction to success and failure gives some hints for potential psychological mechanisms behind this gender effect. According to this literature, men tend to attribute success to internal factors (such as talent) and failure to external factors (such as effort or lack thereof) while women tend to do the opposite (Dweck et al., 1978; Ryckman and Peckham, 1987). An interesting direction for future research would be to investigate whether the propensity to give up competing after a loss predicts individual career choices and outcomes.

References

Athey, Susan and Guido Imbens. 2016. "The State of Applied Econometrics-Causality and Policy Evaluation." *arXiv preprint arXiv:1607.00699* .

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, Armando Jose Garcia Pires, and Bertil Tungodden. 2015. "Competitive in the lab, successful in the field?" *Journal of Economic Behavior & Organization* 118:303–317.

Bertrand, Marianne. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics*, vol. 4B, edited by Orley Ashenfelter and David Card. Elsevier, 1543 – 1590.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2 (3):228–55.

Bertrand, Marianne and Kevin F. Hallock. 2001. "The Gender Gap in Top Corporate Jobs." *Industrial and Labor Relations Review* 55 (1):3–21.

Buser, Thomas. 2016. "The impact of losing in a competition on the willingness to seek further challenges." *Management Science* forthcoming.

Buser, Thomas, Leonie Gerhards, and Joël van der Weele. 2016. "Measuring Responsiveness to Feedback as a Personal Trait." *Tinbergen Institute Discussion Paper 16-043/I* .

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, competitiveness and career choices." *Quarterly Journal of Economics* 129 (3):1409–1447.

Buser, Thomas, Erik Plug, and Lydia Geijtenbeek. 2015. "Do Gays Shy Away from Competition? Do Lesbians Compete Too Much?" *IZA Discussion Paper No. 9382* .

Charness, Gary and Uri Gneezy. 2012. "Strong evidence for gender differences in risk taking." *Journal of Economic Behavior & Organization* 83 (1):50–58.

Croson, Rachel and Uri Gneezy. 2009. "Gender differences in preferences." *Journal of Economic Literature* 47 (2):448–74.

De Paola, Maria and Vincenzo Scoppa. 2015. "Gender Differences in Reaction to Psychological Pressure: Evidence from Tennis Players." *IZA Discussion Paper No 9315* .

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jurgen Schupp, and Gert G. Wagner. 2011. "Individual risk attitudes: Measurement, determinants, and behavioral consequences." *Journal of the European Economic Association* 9 (3):522–550.

Dweck, Carol S, William Davidson, Sharon Nelson, and Bradley Enna. 1978. "Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis." *Developmental psychology* 14 (3):268–276.

European Commission. 2013. "She Figures 2012. Gender in Research and innovation."

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2):171–178.

Fischer, Stefanie. 2015. "The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence." *Working paper* .

Flory, Jeffrey A., Andreas Leibbrandt, and John A. List. 2015. "Do competitive workplaces deter female workers? A large-scale natural field experiment on job-entry decisions." *Review of Economic Studies* 82 (1):122–155.

Gelman, Andrew and Guido Imbens. 2014. "Why high-order polynomials should not be used in regression discontinuity designs." *NBER Working Paper 20405* .

Gill, David and Victoria Prowse. 2014. "Gender Differences and Dynamics in Competition: The Role of Luck." *Quantitative Economics* 5 (2):351–376.

Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2015. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *NBER Working Paper 21517* .

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance In Competitive Environments: Gender Differences." *The Quarterly Journal of Economics* 118 (3):1049–1074.

Imbens, Guido W and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142 (2):615–635.

Iriberri, Nagore and Pedro Rey-Biel. 2016. "Competitive Pressure Widens the Gender Gap in Performance." .

Lee, David S and Thomas Lemieux. 2010. "Regression discontinuity designs in economics." *Journal of Economic Literature* 48 (2):281–355.

Legge, Stefan and Lukas Schmid. 2013. "Rankings, Random Successes, and Individual Performance." *Working paper* .

Möbius, Markus, Muriel Niederle, Paul Niehaus, and Tanya Rosenblat. 2011. "Managing Self-Confidence: Theory and Experimental Evidence." *Working paper* .

Niederle, Muriel and Lise Vesterlund. 2007. "Do women shy away from competition? Do men compete too much?" *The Quarterly Journal of Economics* 122 (3):1067–1101.

———. 2011. "Gender and competition." *Annual Review of Economics* 3 (1):601–630.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2015. "Taste for competition and the gender gap among young business professionals." *Working paper. Columbia University.* .

Reuben, Ernesto, Matthew Wiswall, and Basit Zafar. 2015. "Preferences and Biases in Educational Choices and Labor Market Expectations: Shrinking the Black Box of Gender." *Economic Journal* forthcoming.

Rosenqvist, Olof. 2016. "Is there a gender difference in the ability to deal with failures? Evidence from professional golf tournaments." *Working paper* .

Ryckman, David B and Percy D Peckham. 1987. "Gender differences in attributions for success and failure." *The Journal of Early Adolescence* 7 (1):47–63.

Samek, Anya Savikhin. 2015. "Gender Differences in Job Entry Decisions: A University-Wide Field Experiment." *Working Paper* .

van Veldhuizen, Roel. 2016. "Gender Differences in Tournament Choices: Risk Preferences, Overconfidence or Competitiveness?" *Working paper* .

Wozniak, David. 2012. "Gender differences in a market with relative performance feedback: Professional tennis players." *Journal of Economic Behavior & Organization* 83 (1):158–171.

Wozniak, David, William T. Harbaugh, and Ulrich Mayr. 2014. "The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices." *Journal of Labor Economics* 32 (1):161–198.

Zhang, Y. Jane. 2012. "Can experimental economics explain competitive behavior outside the lab?" *Working paper* .

Online appendix: Additional RD analyses

Figure A1: Distribution of first-round scores by gender

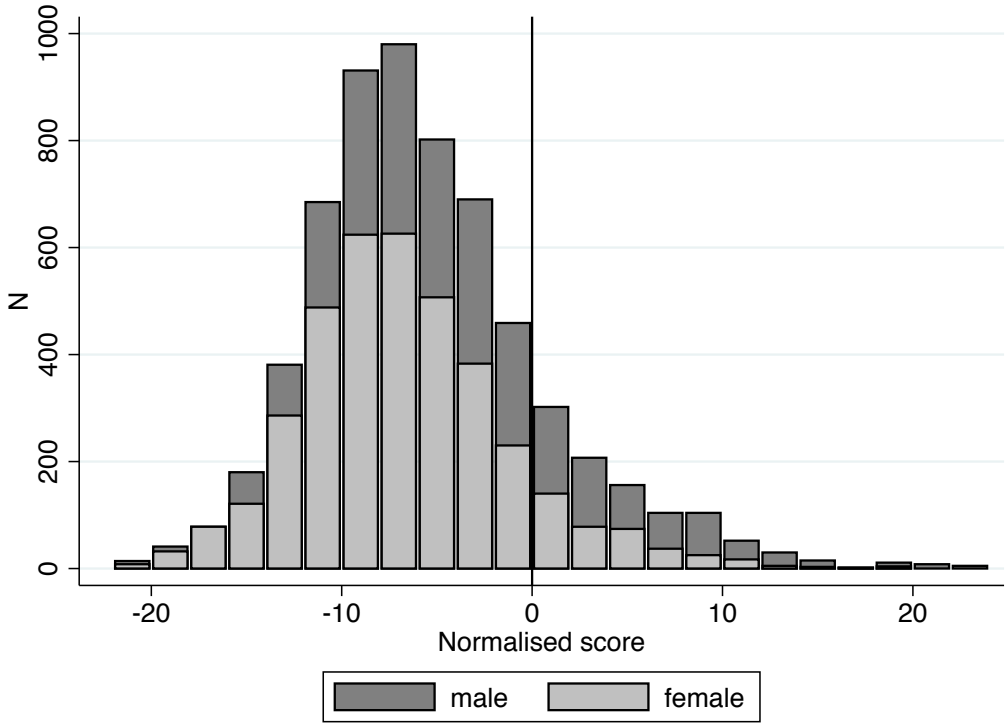


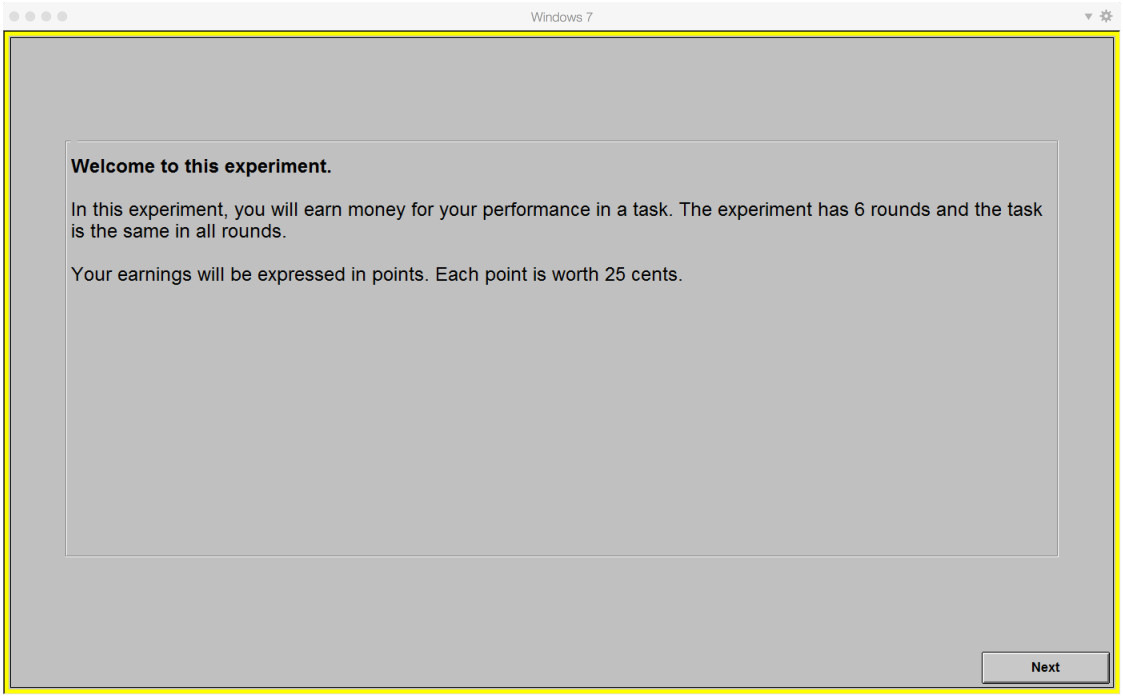
Table A1: Full regression discontinuity results

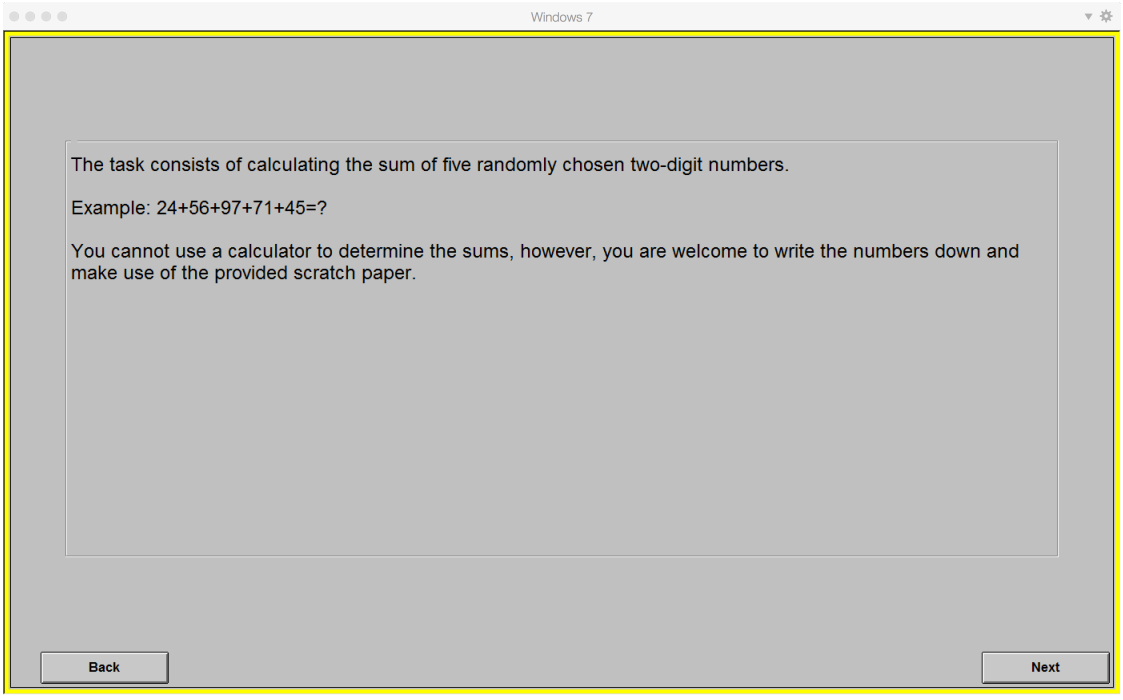
	(1)	(2)	(3)	(4)
Range (in points)	-/+ 4	-/+ 7	-/+ 10	-/+ 20
		Constant:		
Female	0.025	0.020	0.018	0.013
	(0.034)	(0.025)	(0.022)	(0.022)
Lost	-0.098*	-0.173***	-0.214***	-0.245***
	(0.042)	(0.036)	(0.032)	(0.029)
Female x lost	-0.098**	-0.059*	-0.057**	-0.063**
	(0.037)	(0.030)	(0.025)	(0.024)
		First-order polynomial:		
Female	0.026	0.023	0.028	0.025
	(0.036)	(0.026)	(0.025)	(0.024)
Lost	-0.006	-0.019	-0.046	-0.080**
	(0.060)	(0.046)	(0.044)	(0.036)
Female x lost	-0.096**	-0.056*	-0.064**	-0.069***
	(0.039)	(0.031)	(0.027)	(0.025)
		Second-order polynomial:		
Female	0.024	0.023	0.029	0.025
	(0.036)	(0.026)	(0.025)	(0.024)
Lost	0.057*	-0.026	0.001	-0.010
	(0.026)	(0.058)	(0.052)	(0.044)
Female x lost	-0.095**	-0.057*	-0.064**	-0.069***
	(0.039)	(0.031)	(0.027)	(0.025)
N	2489	4730	7459	9967

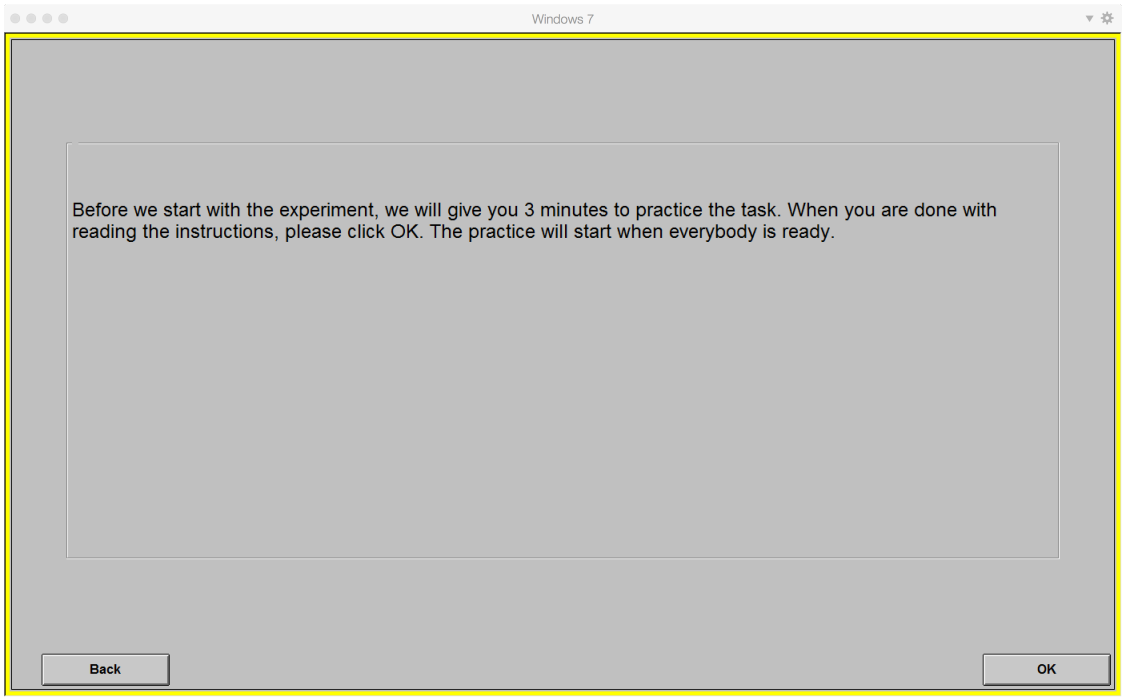
Note: The table shows coefficients from regressions of a binary indicator for participating again a year later on a female dummy, a dummy for not having made the top 1000 ("lost") and the interaction of the two. Range means the sample selection in terms of points left and right of the cutoff. Clustered standard errors in parentheses.

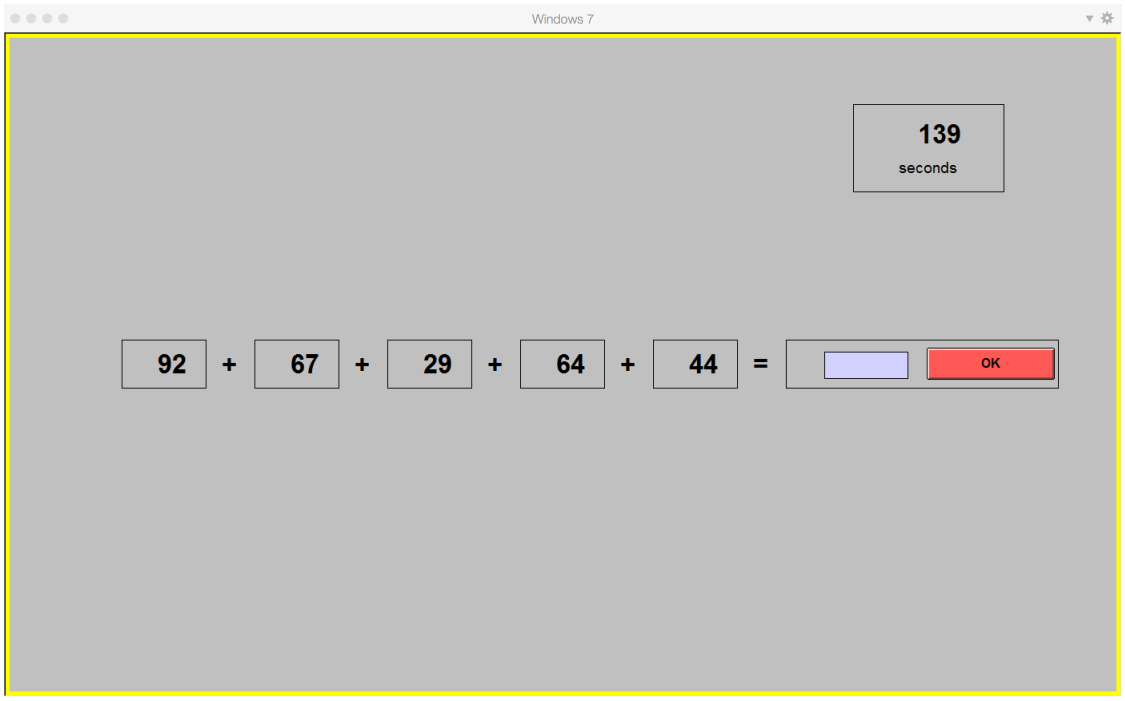
Online appendix: Experimental screenshots

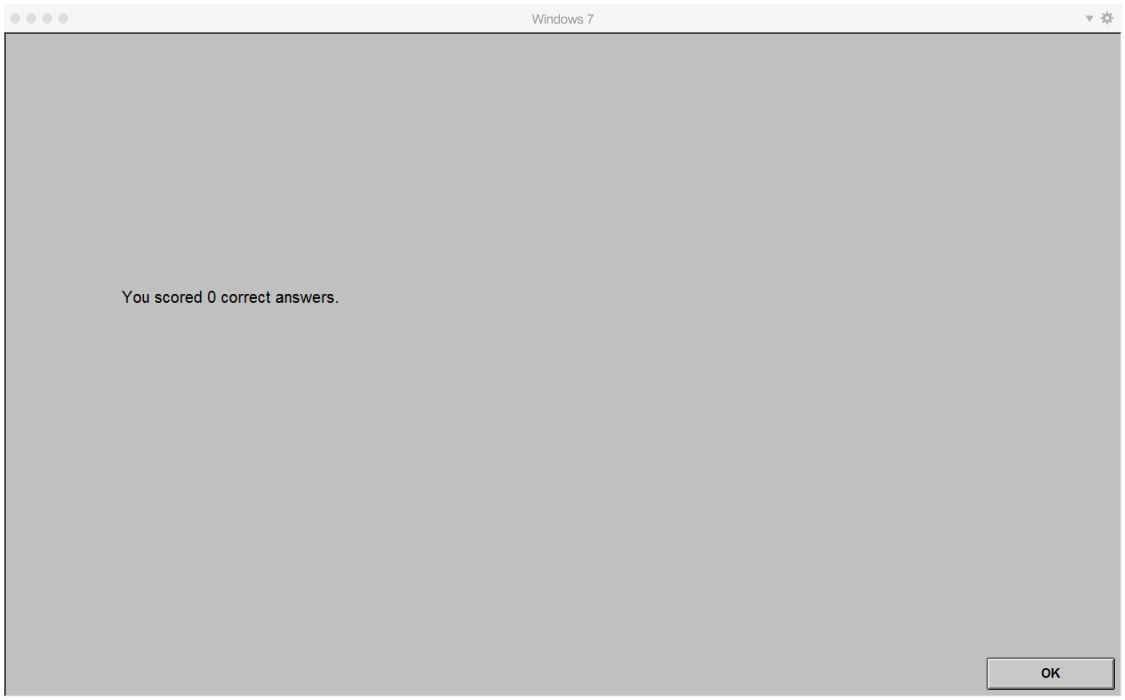
Main experiment

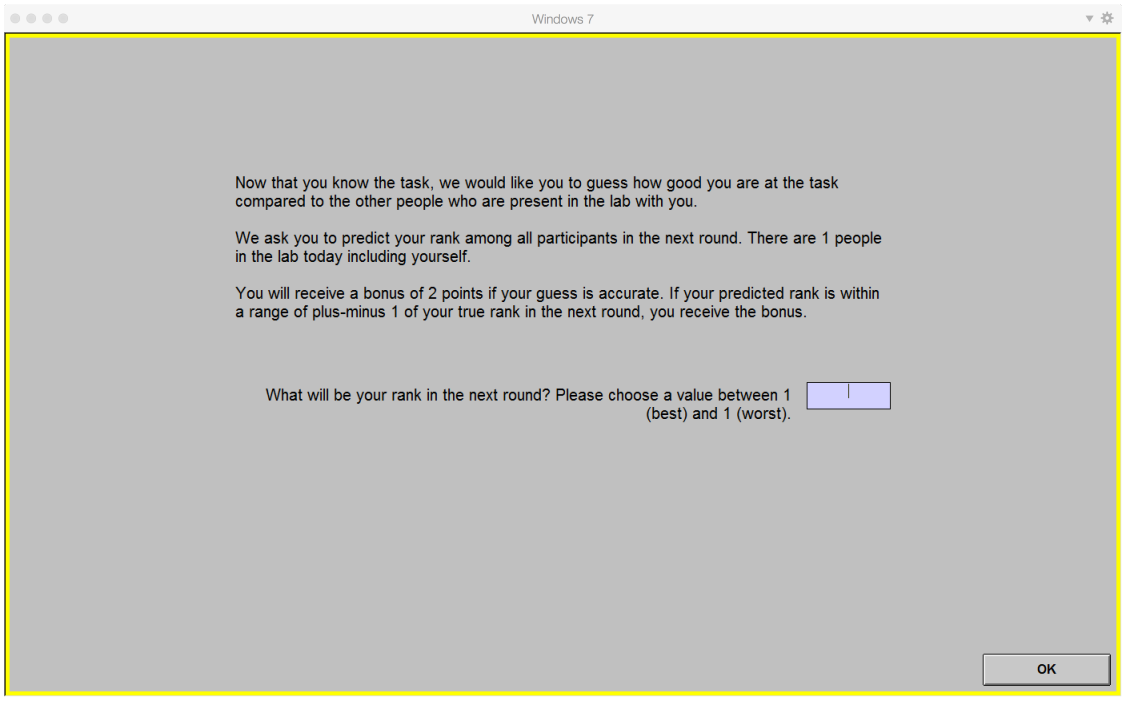


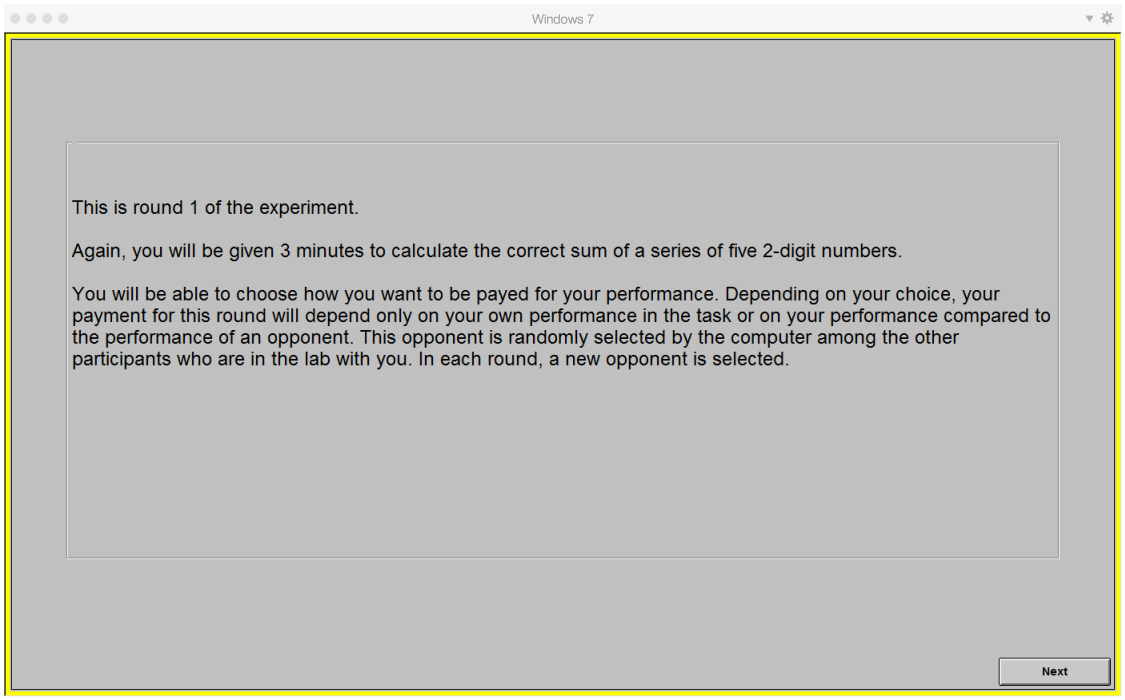


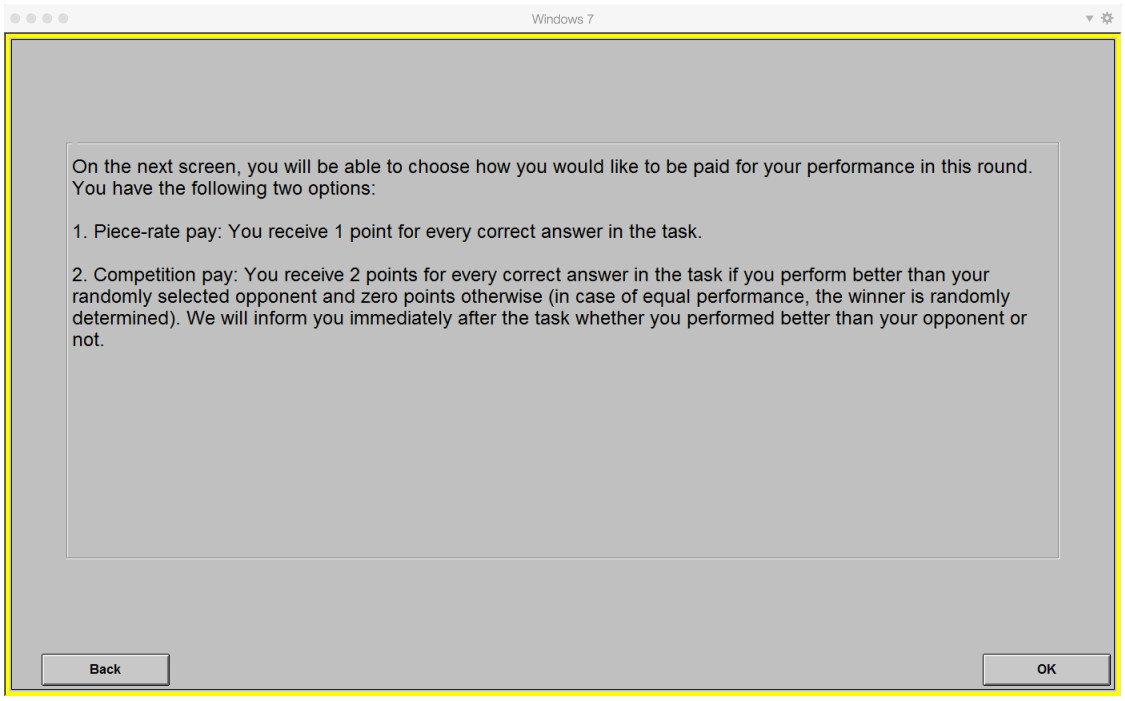

Windows 7
Before we start with the experiment, we will give you 3 minutes to practice the task. When you are done with reading the instructions, please click OK. The practice will start when everybody is ready.
Back
OK

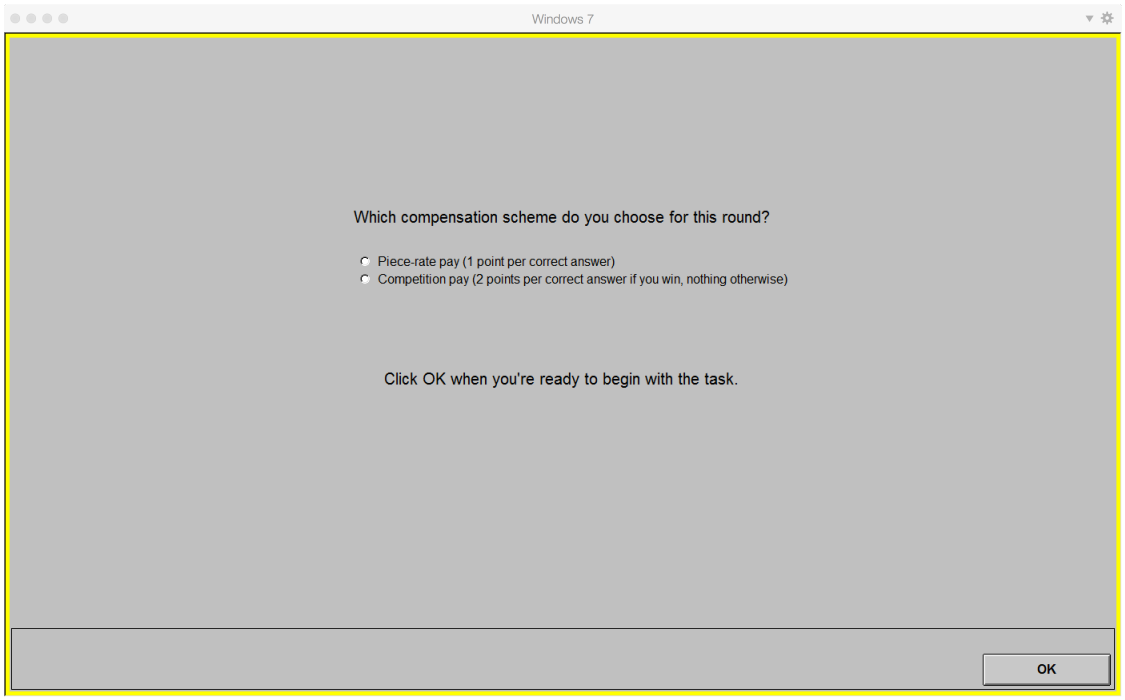

Windows 7
139
seconds
92
+
67
+
29
+
64
+
44
=
OK

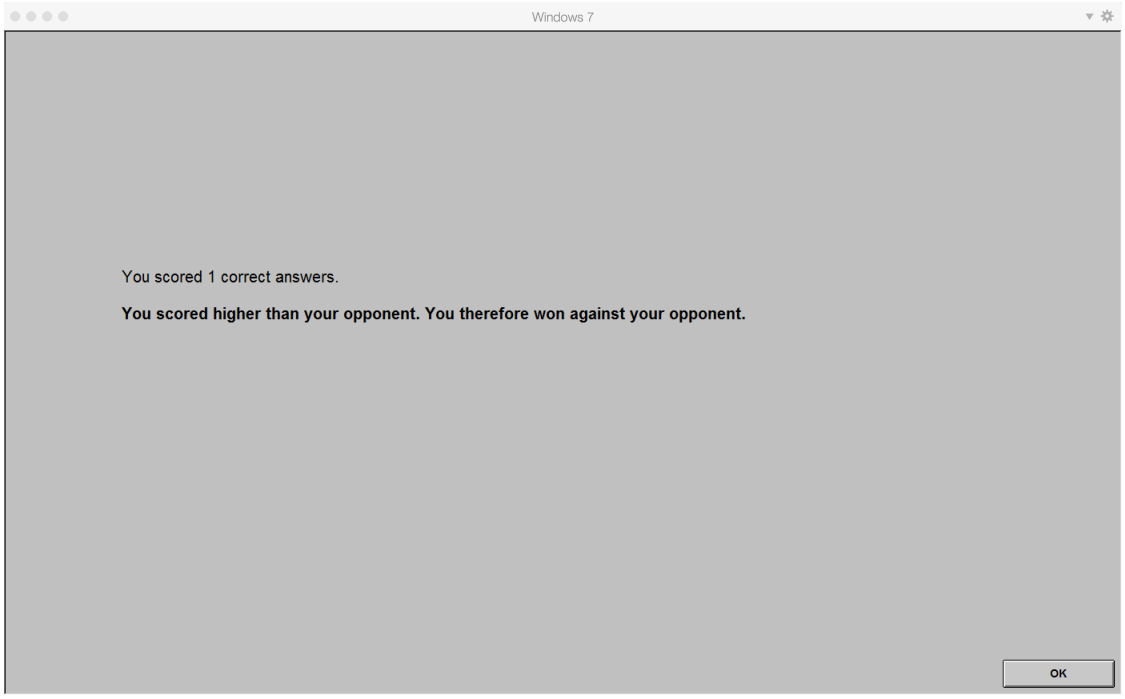

Windows 7
You scored 0 correct answers.
OK

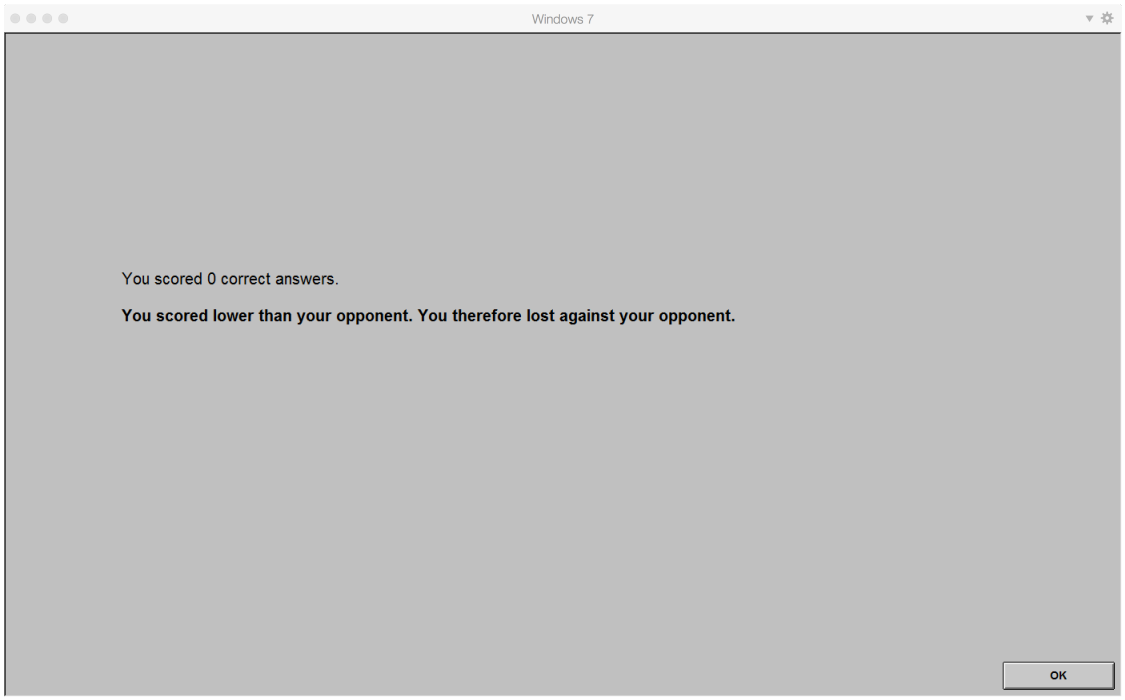

Windows 7
Now that you know the task, we would like you to guess how good you are at the task compared to the other people who are present in the lab with you.
We ask you to predict your rank among all participants in the next round. There are 1 people in the lab today including yourself.
You will receive a bonus of 2 points if your guess is accurate. If your predicted rank is within a range of plus-minus 1 of your true rank in the next round, you receive the bonus.
What will be your rank in the next round? Please choose a value between 1 (best) and 1 (worst).
OK

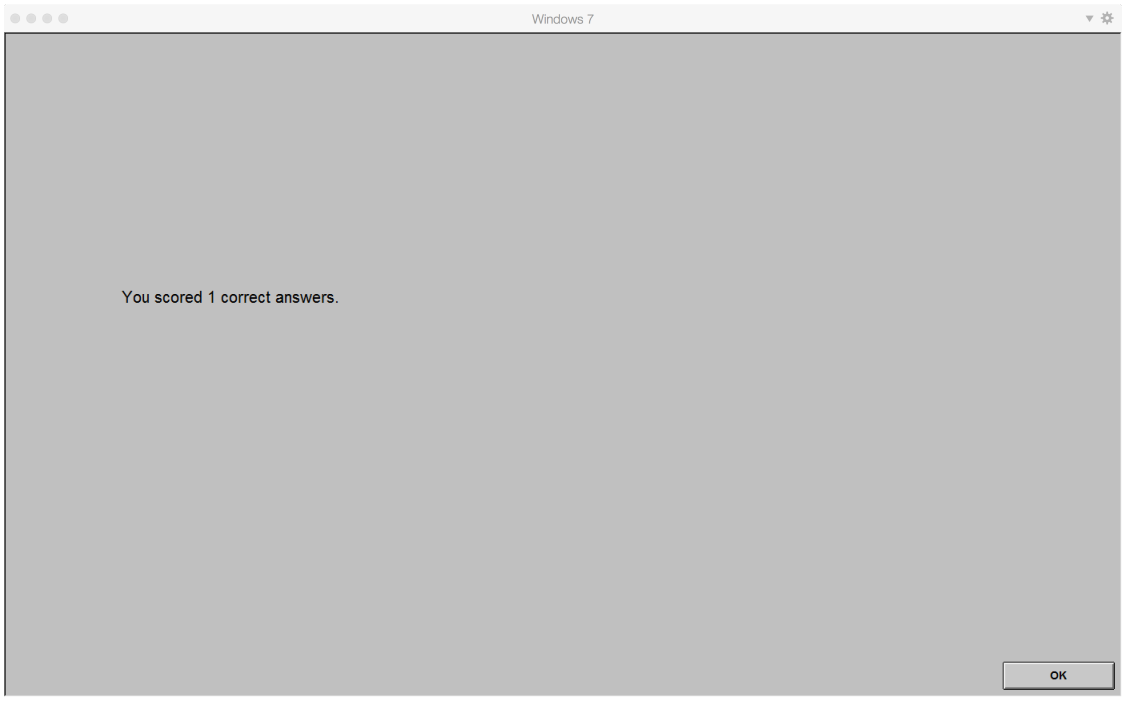

Windows 7
This is round 1 of the experiment.
Again, you will be given 3 minutes to calculate the correct sum of a series of five 2-digit numbers.
You will be able to choose how you want to be payed for your performance. Depending on your choice, your payment for this round will depend only on your own performance in the task or on your performance compared to the performance of an opponent. This opponent is randomly selected by the computer among the other participants who are in the lab with you. In each round, a new opponent is selected.
Next

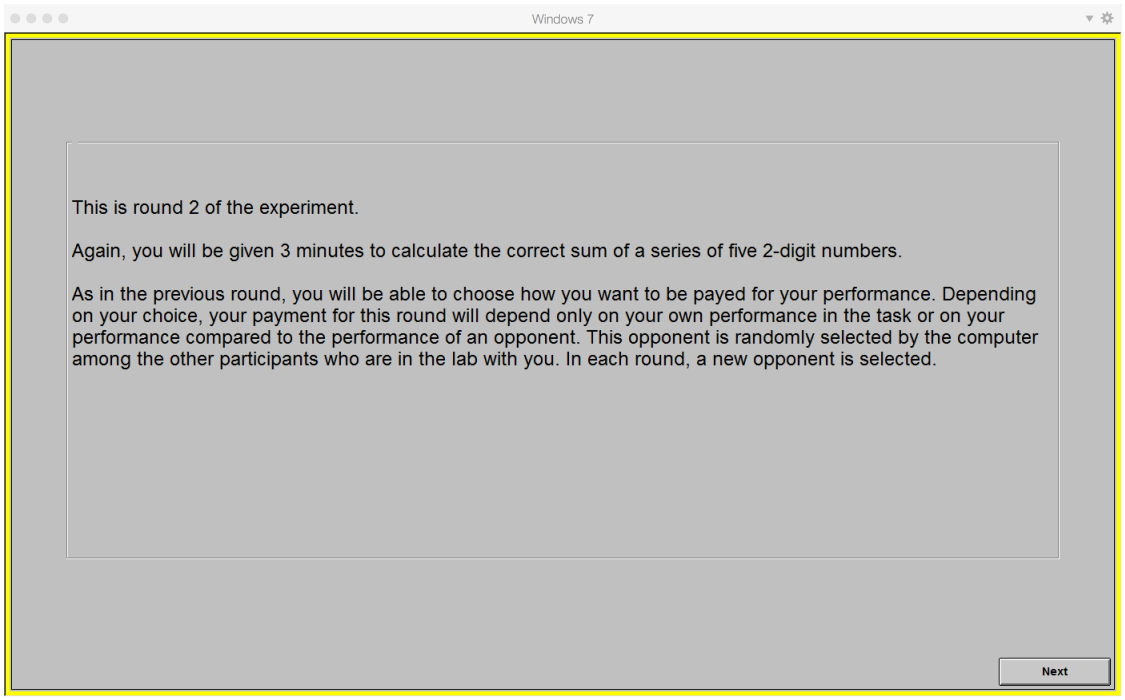

Windows 7
On the next screen, you will be able to choose how you would like to be paid for your performance in this round. You have the following two options:
1. Piece-rate pay: You receive 1 point for every correct answer in the task.
2. Competition pay: You receive 2 points for every correct answer in the task if you perform better than your randomly selected opponent and zero points otherwise (in case of equal performance, the winner is randomly determined). We will inform you immediately after the task whether you performed better than your opponent or not.
Back
OK

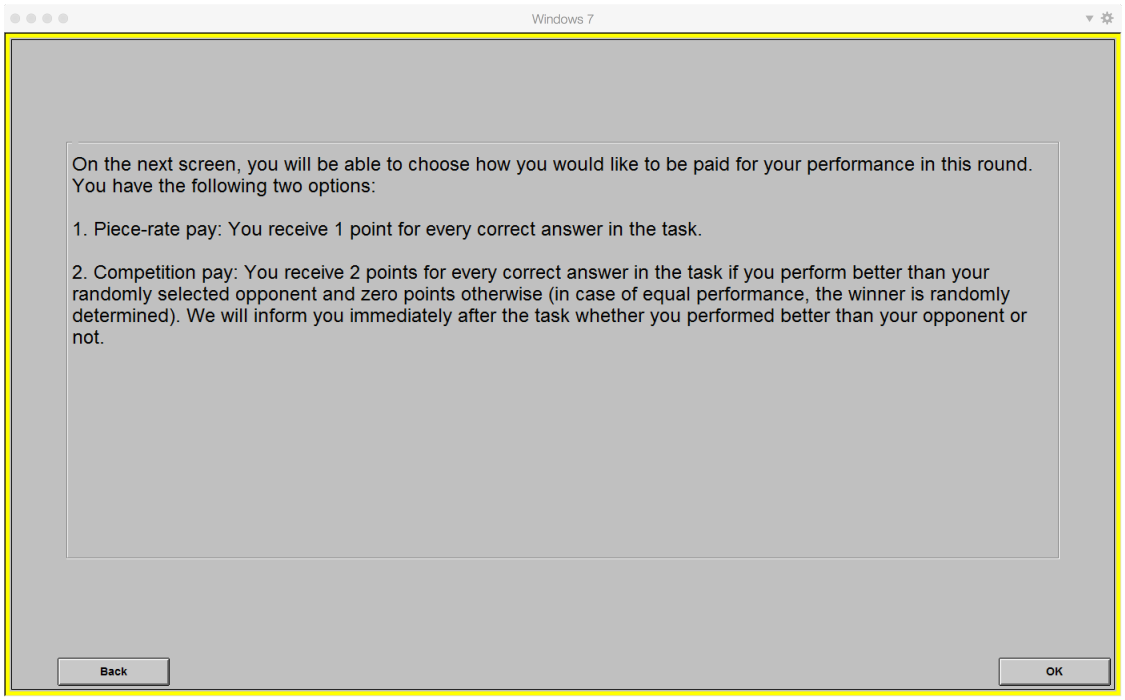

Windows 7
Which compensation scheme do you choose for this round?
Piece-rate pay (1 point per correct answer)
Competition pay (2 points per correct answer if you win, nothing otherwise)
Click OK when you're ready to begin with the task.
OK

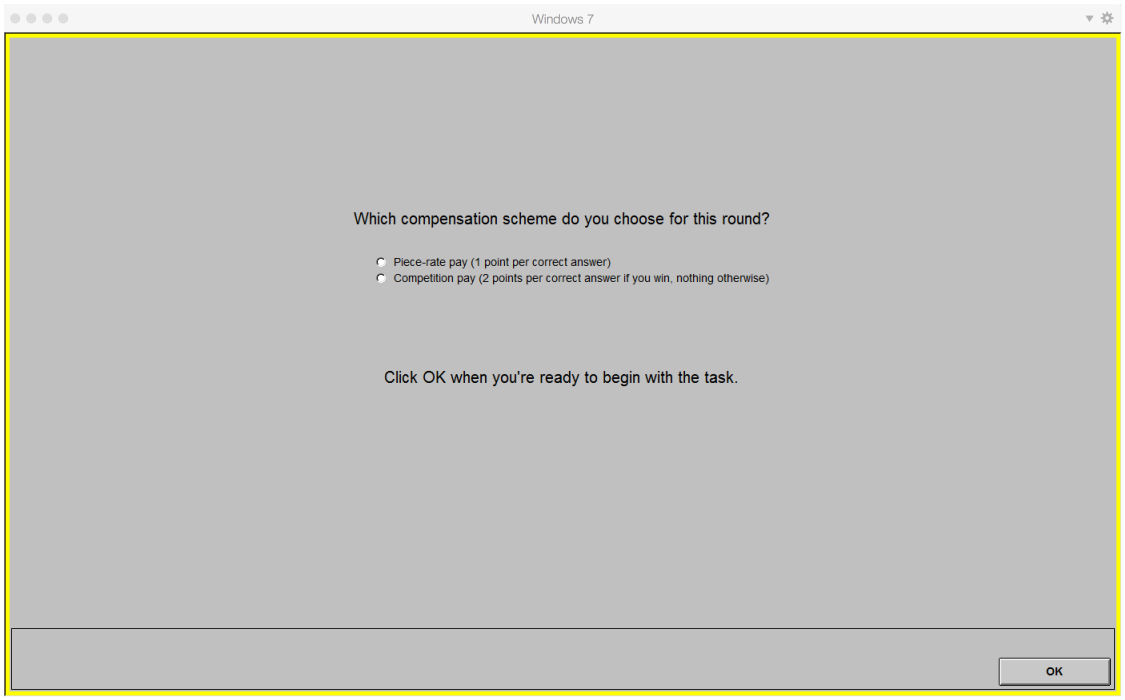

Windows 7
You scored 1 correct answers.
You scored higher than your opponent. You therefore won against your opponent.
OK

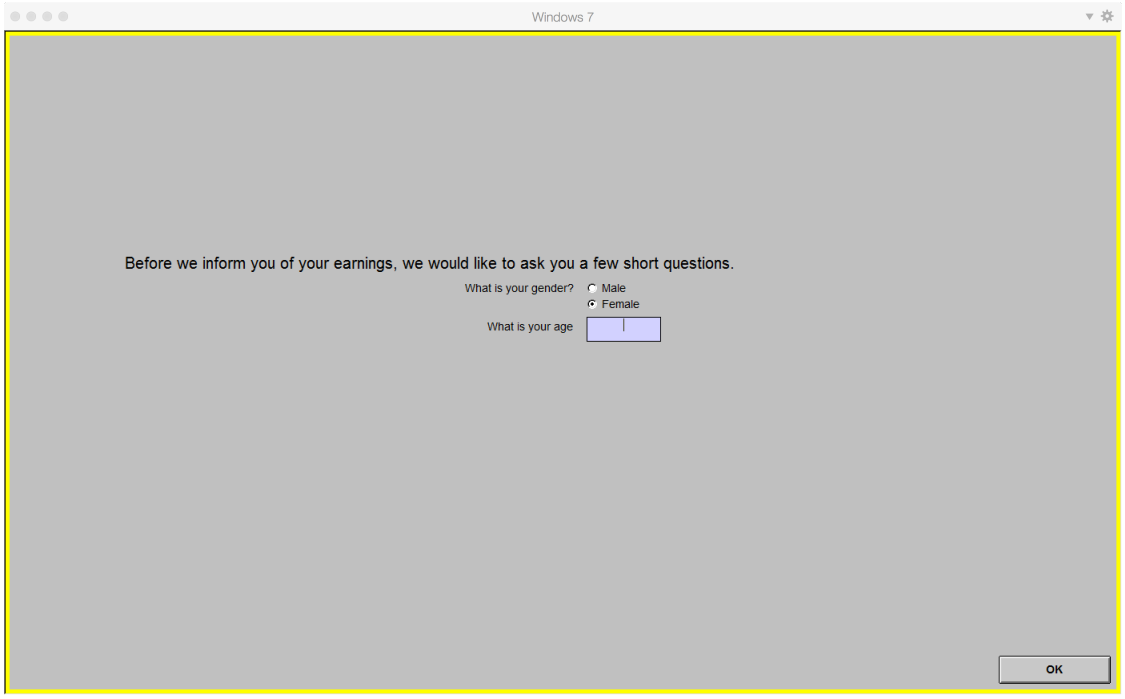

Windows 7
You scored 0 correct answers.
You scored lower than your opponent. You therefore lost against your opponent.
OK

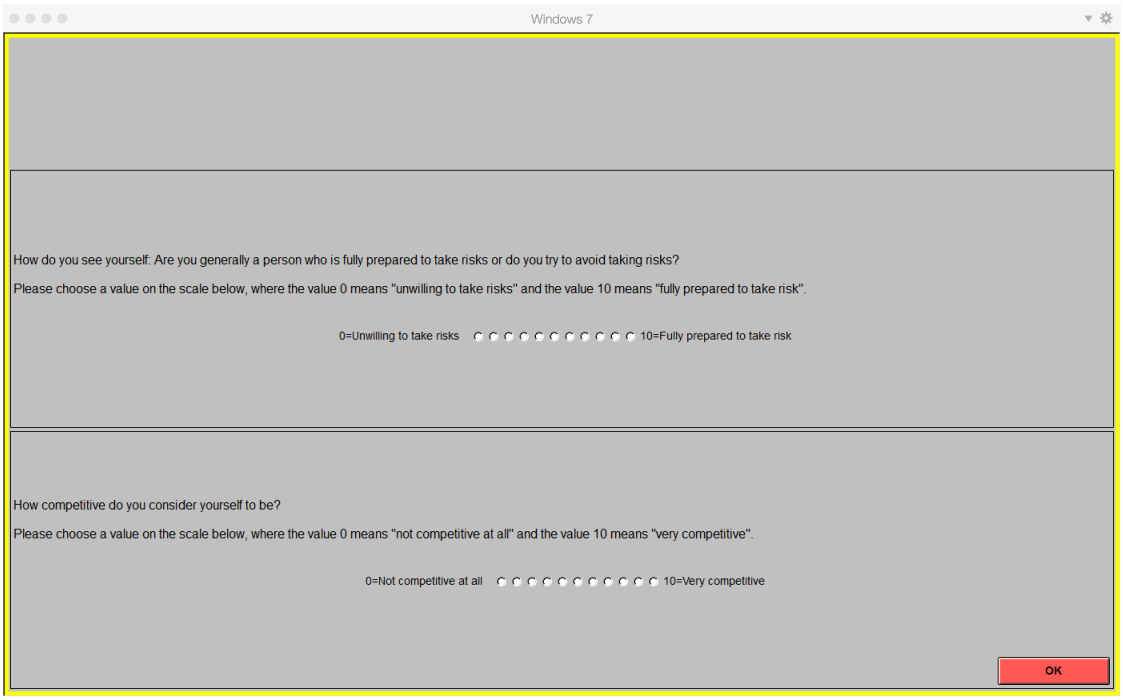

Windows 7
You scored 1 correct answers.
OK


Windows 7
This is round 2 of the experiment.
Again, you will be given 3 minutes to calculate the correct sum of a series of five 2-digit numbers.
As in the previous round, you will be able to choose how you want to be payed for your performance. Depending on your choice, your payment for this round will depend only on your own performance in the task or on your performance compared to the performance of an opponent. This opponent is randomly selected by the computer among the other participants who are in the lab with you. In each round, a new opponent is selected.
Next

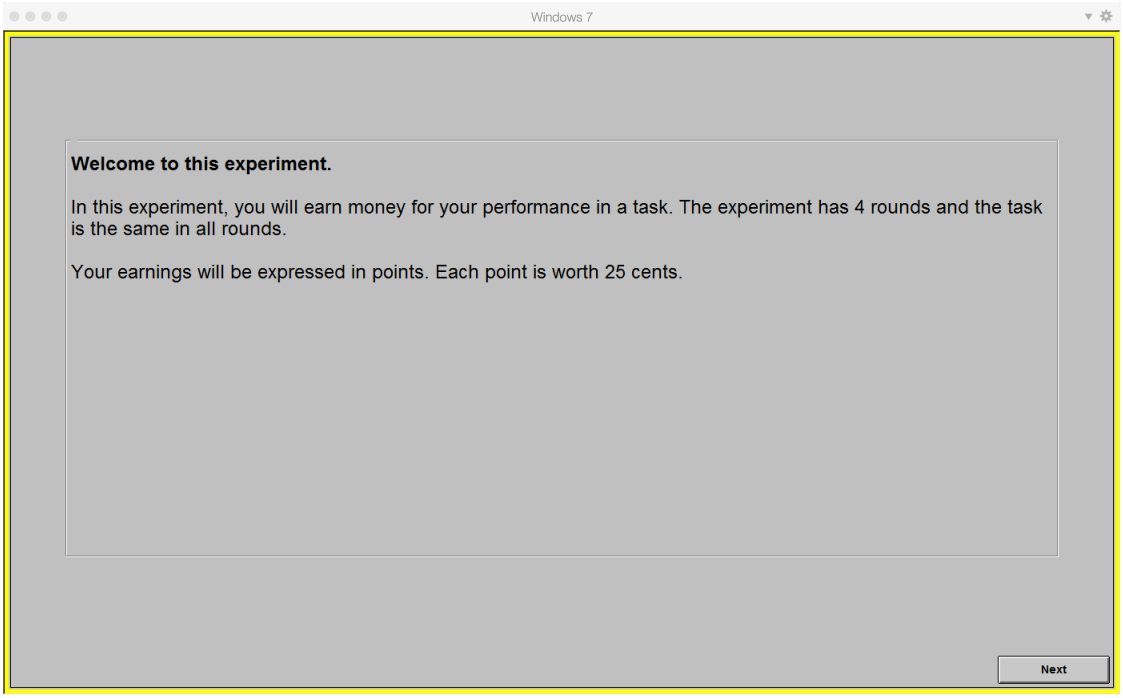

Windows 7
On the next screen, you will be able to choose how you would like to be paid for your performance in this round. You have the following two options:
1. Piece-rate pay: You receive 1 point for every correct answer in the task.
2. Competition pay: You receive 2 points for every correct answer in the task if you perform better than your randomly selected opponent and zero points otherwise (in case of equal performance, the winner is randomly determined). We will inform you immediately after the task whether you performed better than your opponent or not.
Back
OK

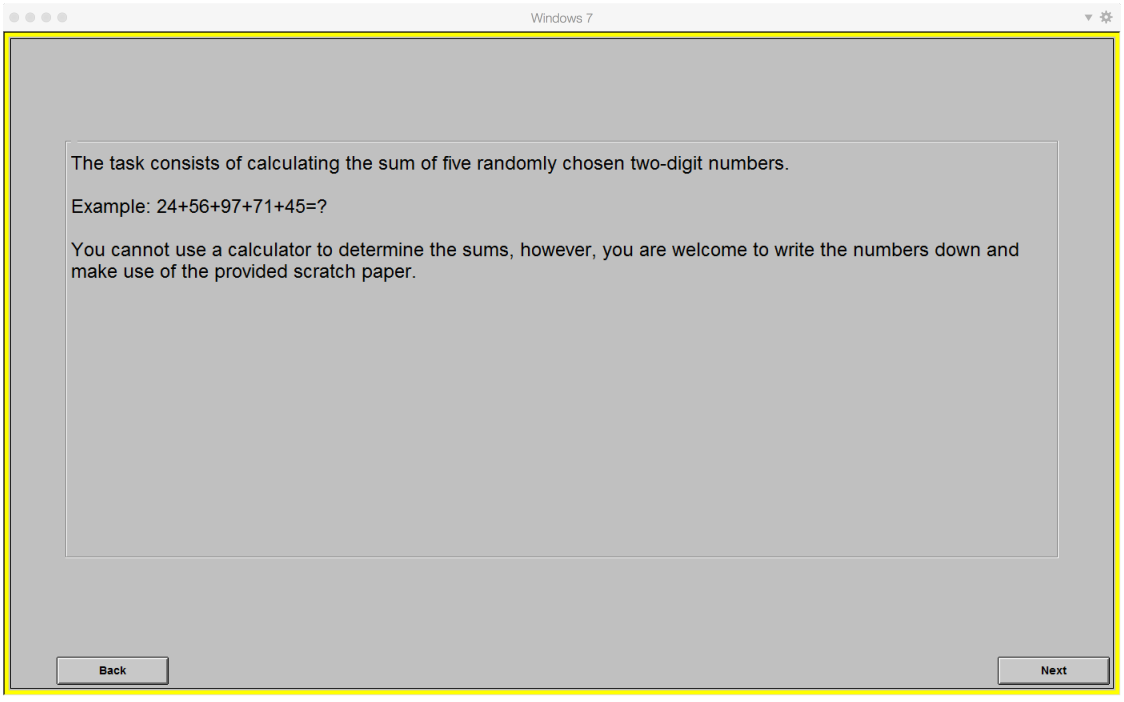

Windows 7
Which compensation scheme do you choose for this round?
Piece-rate pay (1 point per correct answer)
Competition pay (2 points per correct answer if you win, nothing otherwise)
Click OK when you're ready to begin with the task.
OK

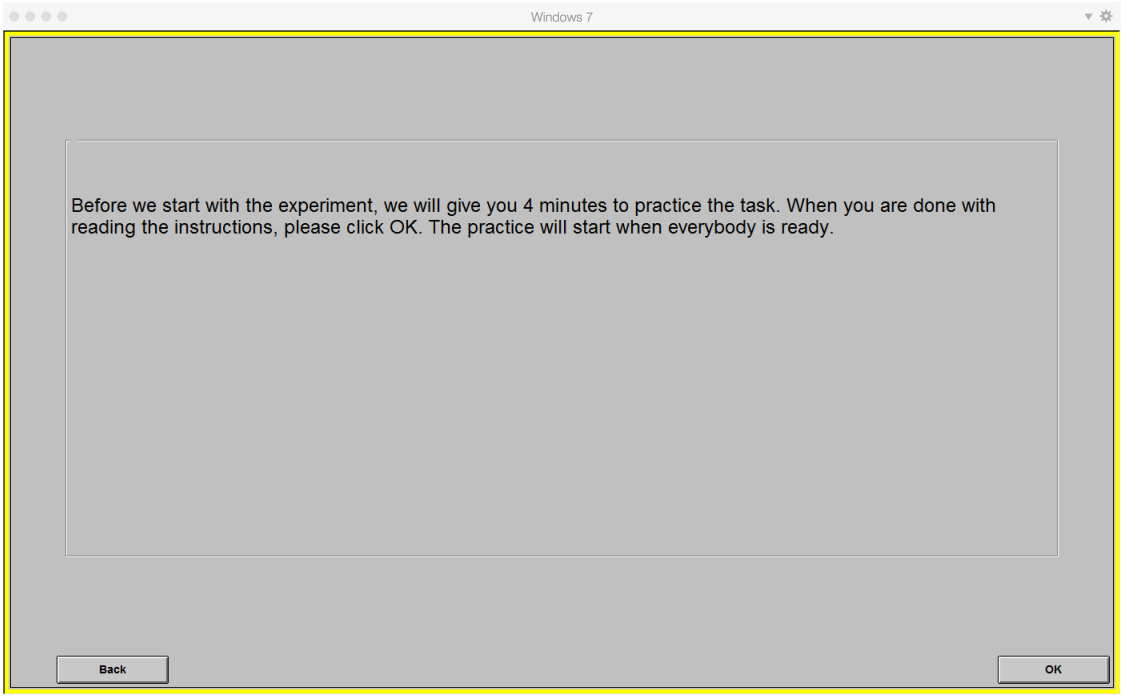

Windows 7
Before we inform you of your earnings, we would like to ask you a few short questions.
What is your gender?
Male
Female
What is your age
OK

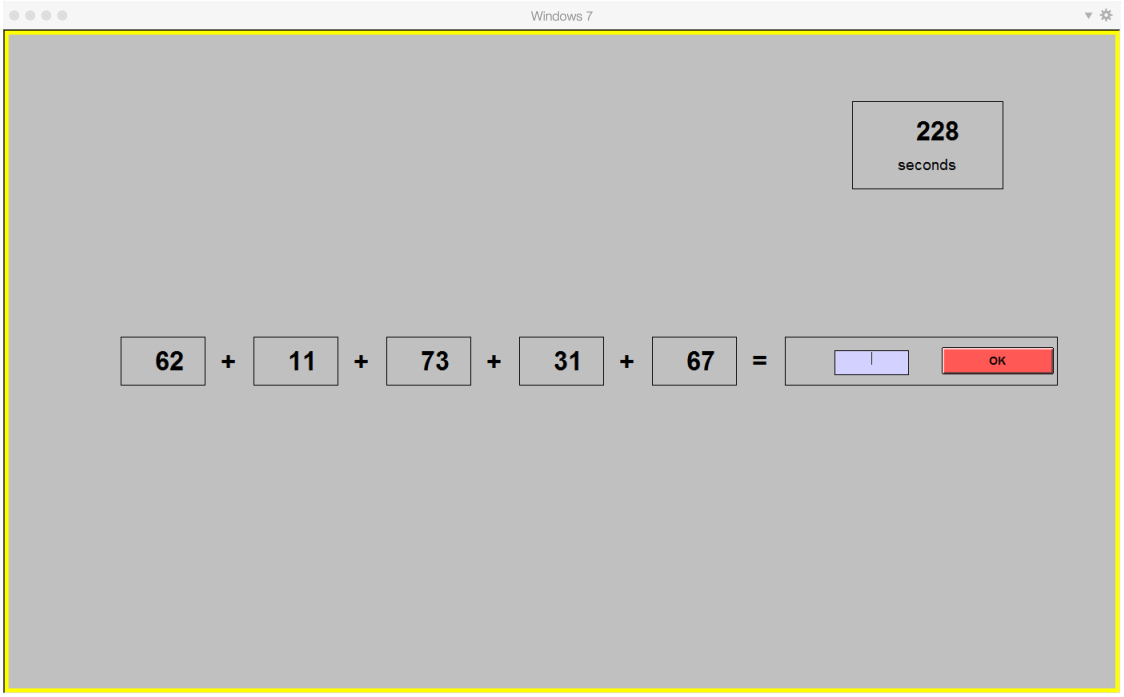

Windows 7
How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?
Please choose a value on the scale below, where the value 0 means "unwilling to take risks" and the value 10 means "fully prepared to take risk".
0=Unwilling to take risks
10=Fully prepared to take risk
How competitive do you consider yourself to be?
Please choose a value on the scale below, where the value 0 means "not competitive at all" and the value 10 means "very competitive".
0=Not competitive at all
10=Very competitive
OK

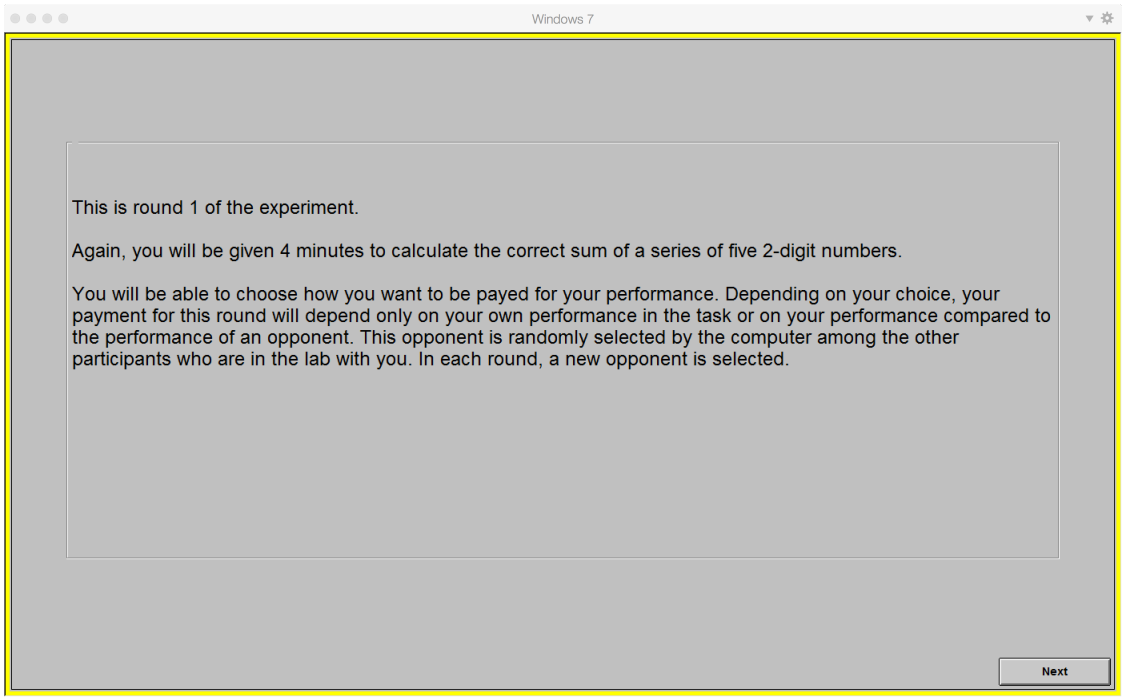
Feedback experiment

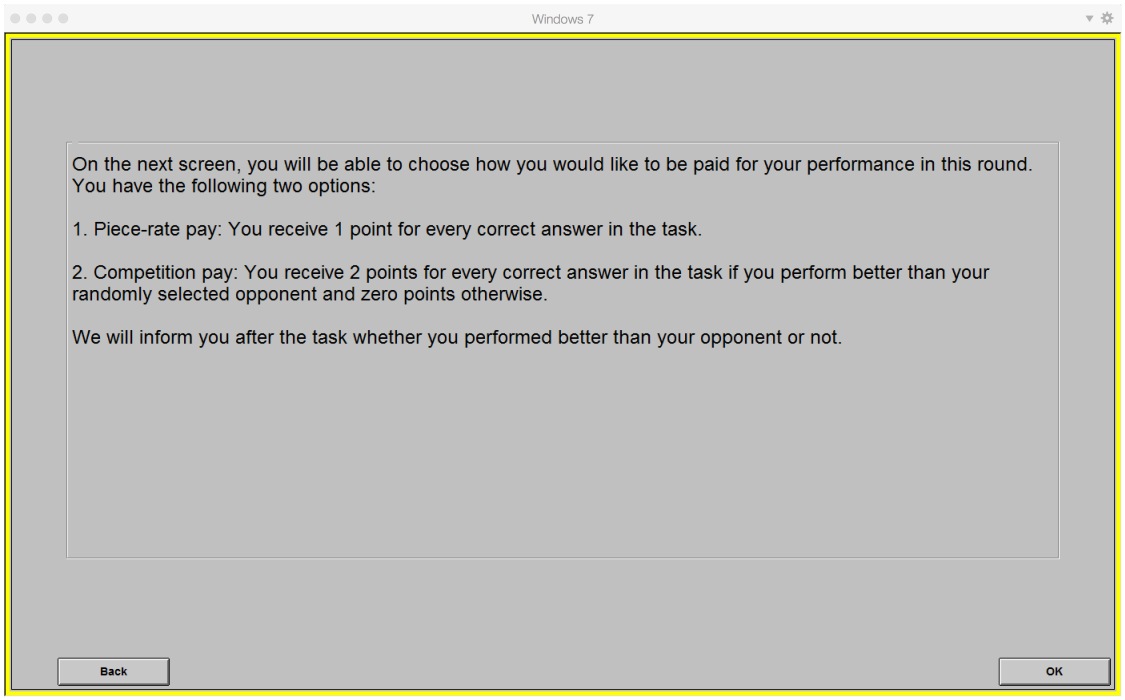


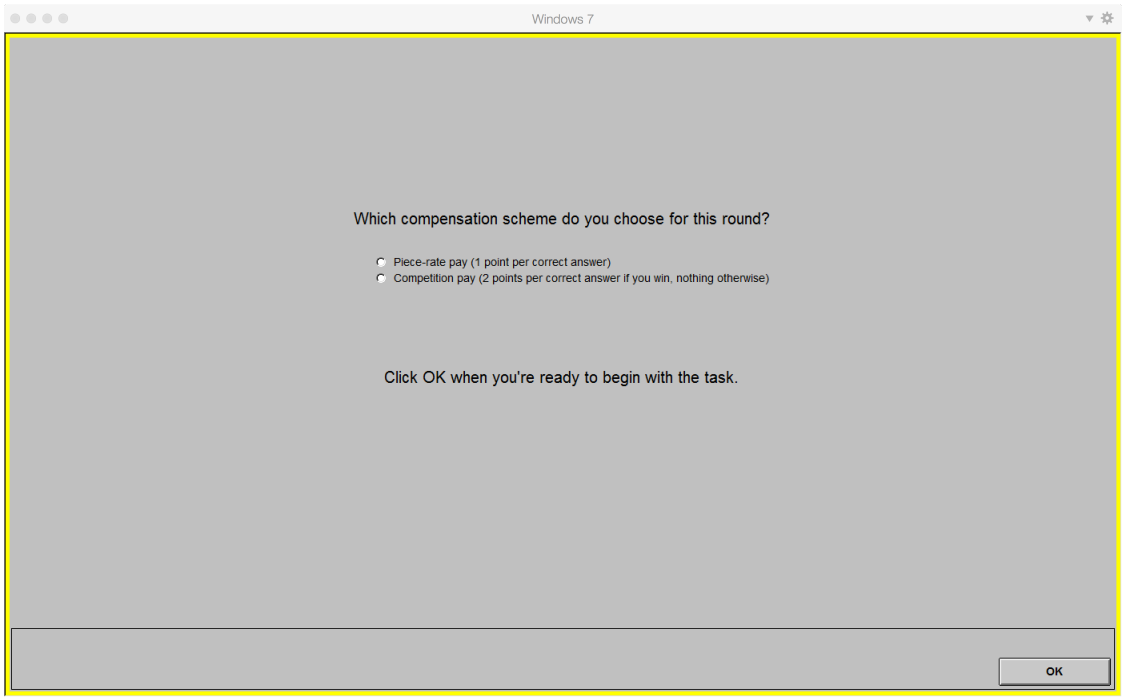


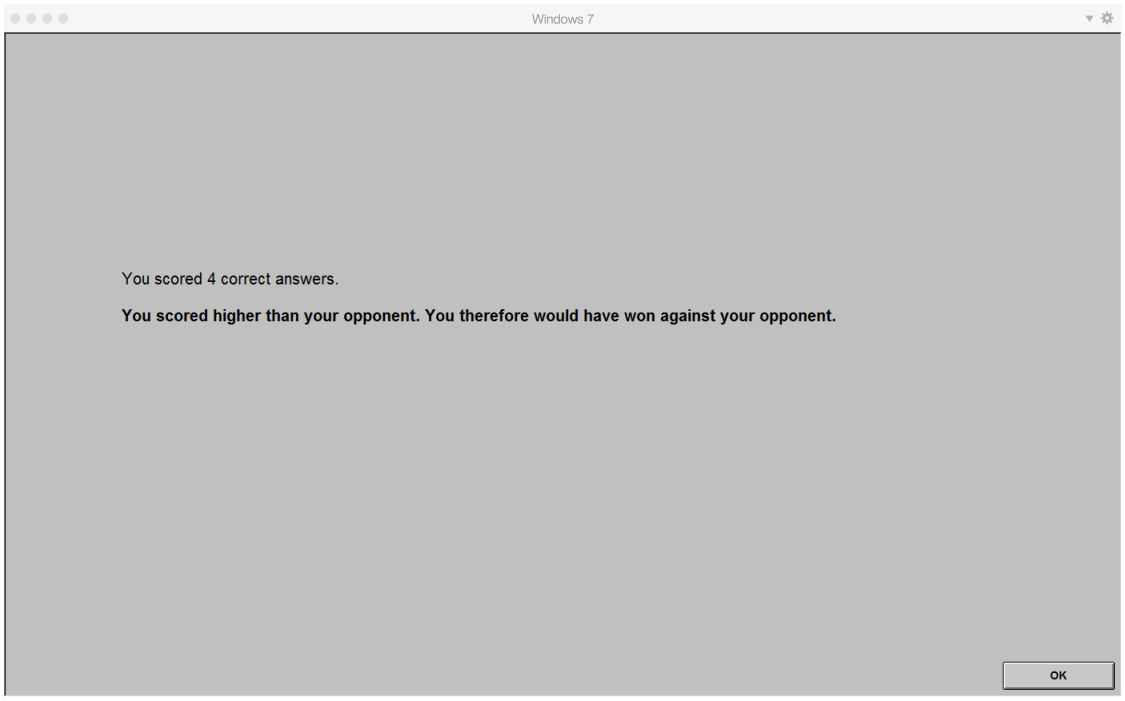

Windows 7
Before we start with the experiment, we will give you 4 minutes to practice the task. When you are done with reading the instructions, please click OK. The practice will start when everybody is ready.
Back
OK

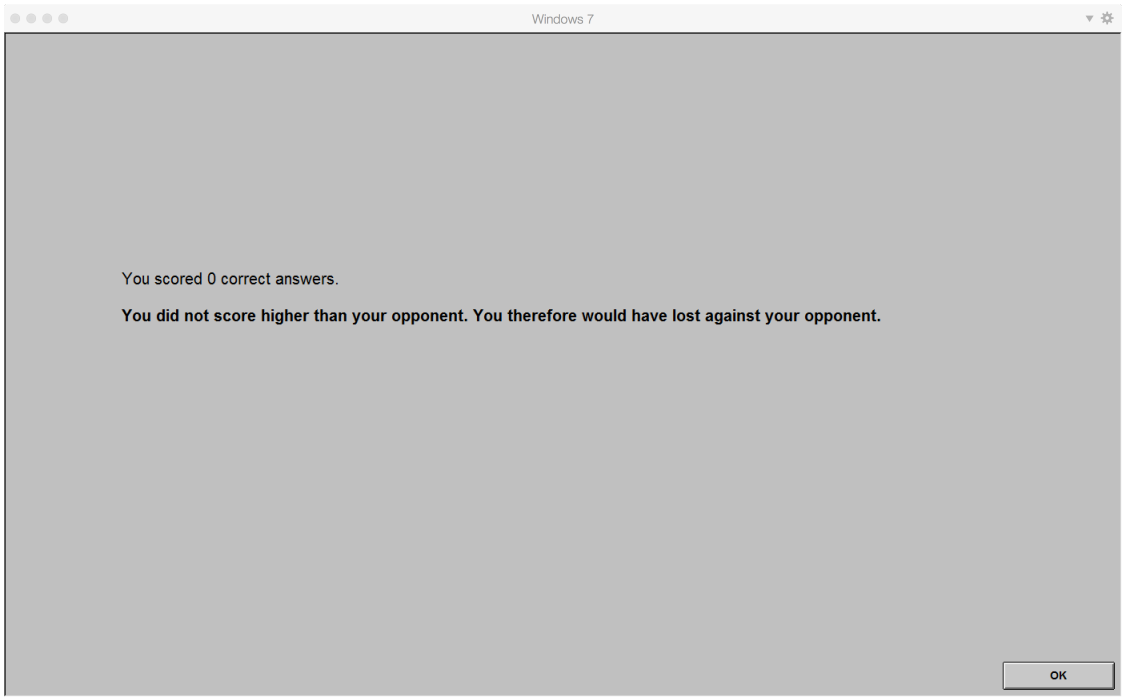

Windows 7
228
seconds
62
+
11
+
73
+
31
+
67
=
OK


Windows 7
This is round 1 of the experiment.
Again, you will be given 4 minutes to calculate the correct sum of a series of five 2-digit numbers.
You will be able to choose how you want to be payed for your performance. Depending on your choice, your payment for this round will depend only on your own performance in the task or on your performance compared to the performance of an opponent. This opponent is randomly selected by the computer among the other participants who are in the lab with you. In each round, a new opponent is selected.
Next

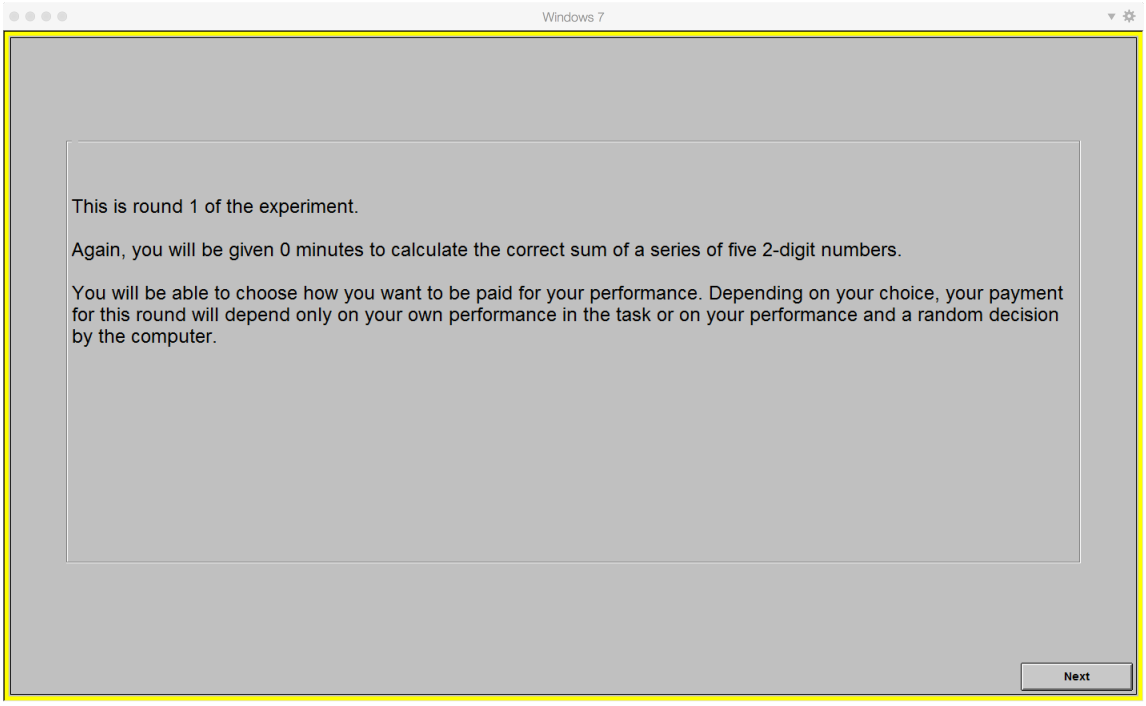

Windows 7
On the next screen, you will be able to choose how you would like to be paid for your performance in this round. You have the following two options:
1. Piece-rate pay: You receive 1 point for every correct answer in the task.
2. Competition pay: You receive 2 points for every correct answer in the task if you perform better than your randomly selected opponent and zero points otherwise.
We will inform you after the task whether you performed better than your opponent or not.
Back
OK

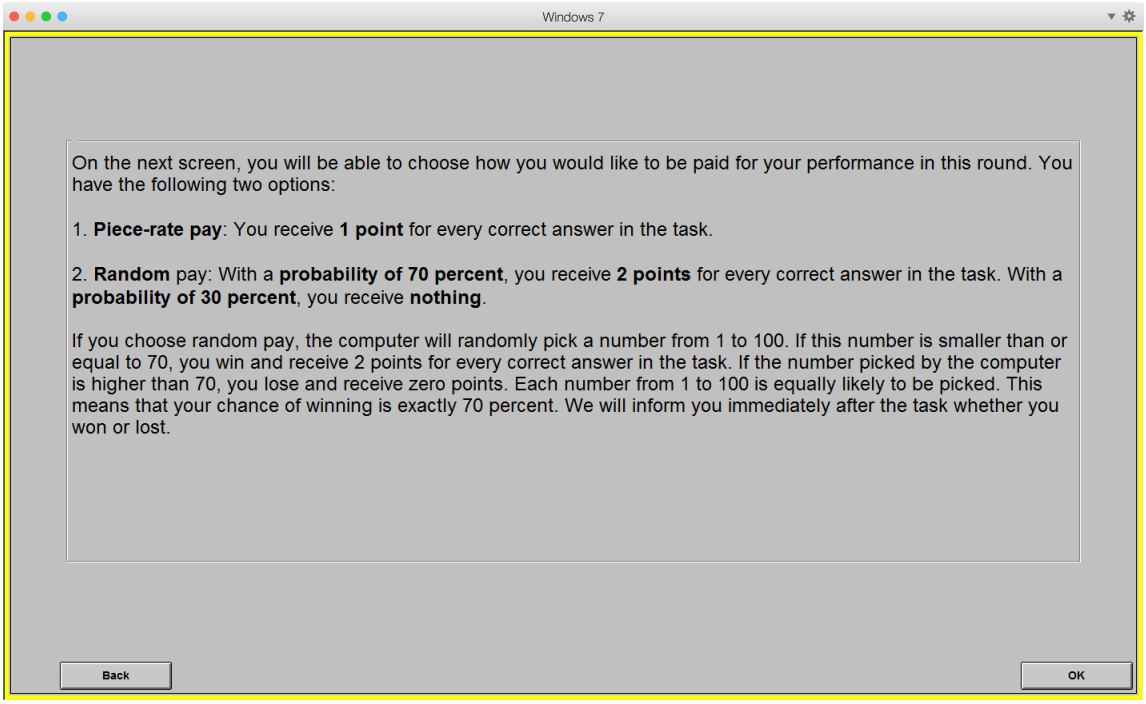

Windows 7
Which compensation scheme do you choose for this round?
Piece-rate pay (1 point per correct answer)
Competition pay (2 points per correct answer if you win, nothing otherwise)
Click OK when you're ready to begin with the task.
OK

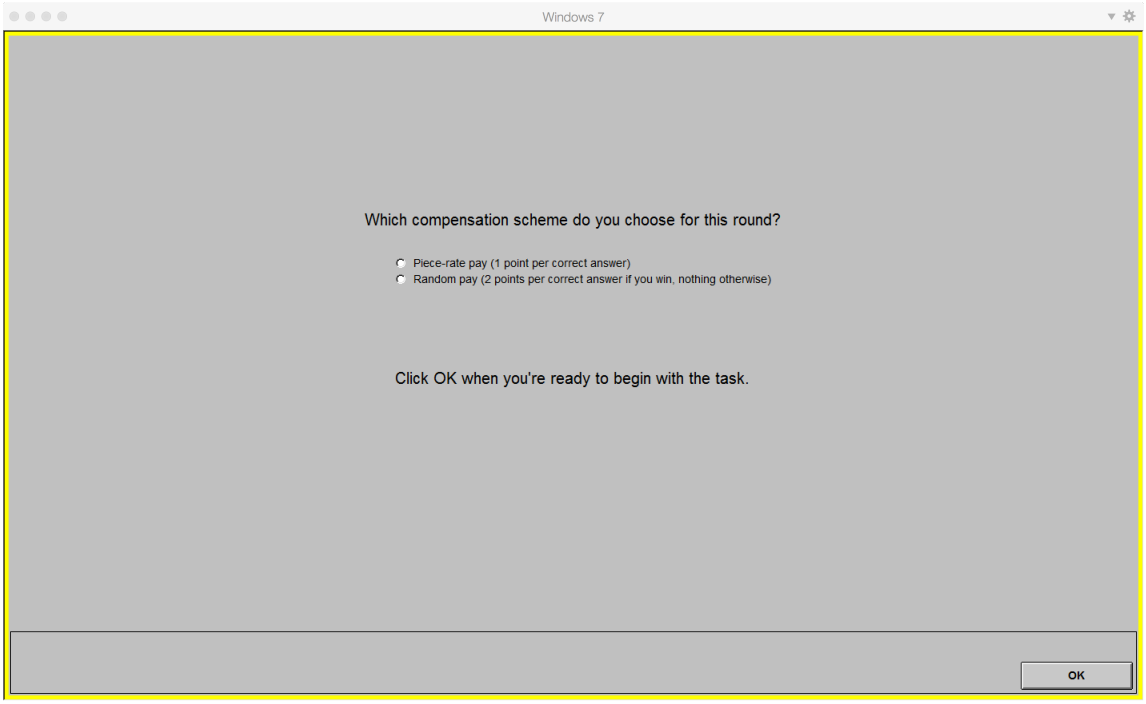

Windows 7
You scored 4 correct answers.
You scored higher than your opponent. You therefore would have won against your opponent.
OK

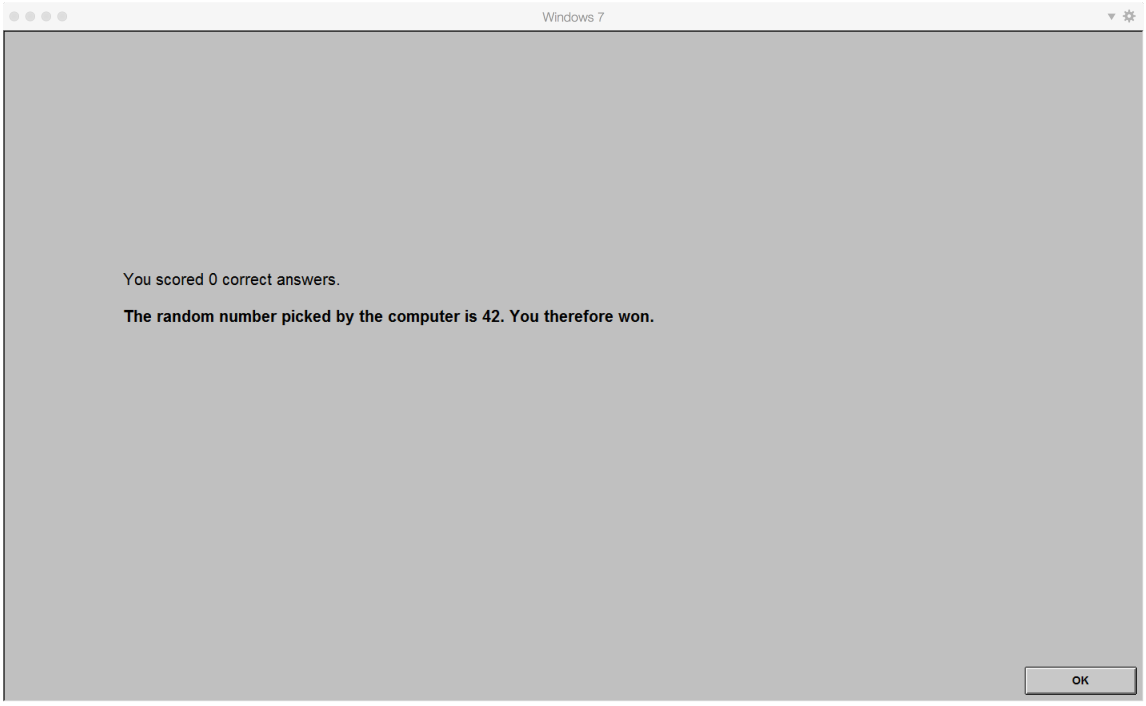

Windows 7
You scored 0 correct answers.
You did not score higher than your opponent. You therefore would have lost against your opponent.
OK

Risk experiment






Windows 7
Which compensation scheme do you choose for this round?
Piece-rate pay (1 point per correct answer)
Random pay (2 points per correct answer if you win, nothing otherwise)
Click OK when you're ready to begin with the task.
OK


Windows 7
You scored 0 correct answers.
The random number picked by the computer is 42. You therefore won.
OK