(Re)Measuring preferences for competition

LINA LOZANO
New York University Abu Dhabi, Center for Behavioral Institutional Design, e-mail: lina.lozano@nyu.edu

ERNESTO REUBEN
New York University Abu Dhabi, Center for Behavioral Institutional Design, Luxembourg Institute of Socio-Economic Research, e-mail: ereuben@nyu.edu

ABSTRACT
This study investigates individuals' competitive behavior. We present two experiments designed to evaluate the consistency of individuals' choices to compete, obtain direct measures of their preferences for competition, and evaluate the stability of these preferences over time. We find strong evidence that many individuals are willing to forgo a significant portion of their expected earnings to either engage in or avoid competition. Additionally, their choices are consistent with expected utility maximization and are relatively stable over time. Preferences for competition vary with the number of competitors but we do not find significant differences by gender.

This version: February 2025

ACKNOWLEDGEMENTS
We gratefully acknowledge financial support from Tamkeen under the NYU Abu Dhabi Research Institute Award CG005. The studies in this paper were approved by the Maastricht University Ethical Committee (No. ERCIC-076-05-04-2018) and New York University Abu Dhabi Institutional Review Board (No. 065-2018).

جامعة نيويورك أبوظبي
NYU | ABU DHABI

مركز التصميم السلوكي المؤسساتي
CENTER FOR BEHAVIORAL INSTITUTIONAL DESIGN

1. Introduction

In recent decades, economists have increasingly recognized the importance of non-cognitive factors as determinants of economic behavior. For instance, Heckman et al. (2019) conclude that factors such as psychological traits and preferences play a significant role in shaping important life outcomes, including wages and health. One such trait is individuals' preferences for competition. People face competition in many areas, such as the workplace and the educational system. Therefore, it is understandable that several studies have found that competitive behavior in the laboratory predicts important labor market and educational outcomes (e.g., Buser et al., 2014; Berge et al., 2015; Buser et al., 2017b; Reuben et al., 2017; Kamas and Preston, 2018; Reuben et al., 2024; Zhang, 2019; Buser et al., 2022; Dariel and Nikiforakis, 2022; Buser et al., 2024; for a review see Lozano et al., 2023).

Since the seminal paper by Niederle and Vesterlund (2007), many studies have documented individual heterogeneity in the decision to compete (Niederle, 2017). In particular, numerous studies have consistently shown that women are less likely to compete with others than men (for reviews see, Niederle, 2017; Dariel et al., 2017; Markowsky and Beblo, 2022). However, the underlying factors behind competitive behavior remain unclear (Gillen et al., 2019; van Veldhuizen, 2022; Buser et al., 2024). Is competitive behavior driven by individual differences in ability, beliefs about relative performance, and risk attitudes, or by differences in preferences for competition? Are gender differences in competition entry due women being more averse to competition, or are they the result of gender differences in other factors? Given the importance of competition in determining economic outcomes, it is crucial to understand whether decisions to engage in competition—and the gender gap in competitive behavior—are driven by a desire or aversion to perform in competitive environments.

Typically, preferences for competition are measured by asking participants in an experiment to choose how they wish to be paid for performing an addition task. Participants select between two payment schemes: *individual pay* and *competitive pay*. Under individual pay, they earn a fixed amount per correct answer. Under competitive pay, they earn a higher amount per correct answer but only if they achieve the highest performance in a randomly assigned group (for details, see Niederle and Vesterlund, 2007). Preferences for competition are often identified as the residual from regressing the payment choice on measures of other determinants of the decision to compete, such as risk preferences, beliefs, and performance. In this approach, gender differences in preferences for competition are measured by including a gender dummy as an explanatory variable (e.g., Buser et al., 2014, 2017b; Reuben et al., 2017; Buser et al., 2024). This approach has sparked an ongoing debate on whether a statistically significant gender coefficient is evidence of a gender difference in preferences for competition or merely reflects measurement error in the other determinants of competition entry (Gillen et al., 2019; van Veldhuizen, 2022). In fact, Gillen et al. (2019) find that the gender gap in competition disappears once risk preferences, beliefs, and measurement error are accounted for. A finding that can be interpreted as evidence that preferences for competition do not exist. However, this interpretation contrasts with growing evidence showing that laboratory measures of competitive behavior predict distinct labor market and educational outcomes from those predicted by risk preferences and beliefs (Buser et al., 2024; Reuben et al., 2024), supporting the view that preferences for competition constitute a distinct trait.

In this paper, we contribute to this debate by developing a more precise method for measuring preferences for competition and examining these preferences in more detail. Our design allows us to quantify the strength of individuals' preferences for competition while eliminating the confounding effects of risk preferences and more precisely controlling for beliefs. We elicit multiple measures per individual, allowing us to test whether individuals are persistently competition loving or competition averse and evaluate whether their choices are consistent with expected utility maximization, suggesting we are indeed measuring preferences. Additionally, we assess the stability of competition preferences over time to determine whether they represent a persistent trait. Finally, we examine two key features of these preferences: how they vary with the number of competitors and whether they differ between men and women.

To remove the influence of risk preferences on the decision to compete, we modify the payment choice. As in Niederle and Vesterlund (2007), participants who choose competitive pay earn a large amount per correct answer if they are the best performer in their group and a low amount otherwise. Unlike the traditional design, in our experiment, participants who choose individual pay take part in a lottery. If they win the lottery, participants earn the same large amount per correct answer as those who choose to compete and win, and the same low amount otherwise. By ensuring that both payment options have identical outcomes, we prevent risk preferences from influencing participants' choices. This approach removes the need to statistically control for risk preferences and allows us to measure preferences for competition without making assumptions about the correlation between these two traits.

A crucial variable to identify preferences for competition is participants' belief in their relative performance—specifically, their belief that they will be the best performer in their group. We employ two different approaches to deal with beliefs. As a first approach, we focus on accurately measuring participants' beliefs. To achieve this, we incentivize belief elicitation using a binarized scoring rule (Hossain and Okui, 2013), which has the advantage that it is not confounded by varying levels of risk preferences and has been shown to outperform other belief elicitation methods (e.g., see Trautmann and van de Kuilen, 2015). We further reduce noise in belief elicitation by using an interactive graphical interface that automatically calculates the applicable incentives (see, Danz et al., 2022) and by helping participants calculate how their expected percentile ranking in the performance distribution translates into the probability of being the best performer in a randomly-formed group. As a second approach, instead of eliciting beliefs, we directly inform participants of their precise probability of winning the competition in a randomly-formed group within their session, which we calculate based on their past performance.

Another feature of our experimental design is that participants choose between competitive and individual pay using a multiple price list (MPL), where the probability of winning the lottery in individual pay gradually increases. For a given belief of being their group's winner, the point at which participants switch from competitive to individual pay allows us to calculate the precise amount of utility they are willing to forgo to either avoid or engage in competition. For example, competition-neutral participants will switch from competitive pay to individual pay as soon as the probability of winning the lottery exceeds their expected probability of winning the competition. Participants who switch before this point accept a lower expected utility to avoid competition, making them competition averse. Analogously, participants who switch after this point accept a lower expected utility to continue competing, making them competition loving. Additionally, the pattern of choices in the MPL allows us to determine whether a participant's payment-scheme choices are consistent with expected utility maximization.

A common drawback of most studies that measure preferences for competition is that they rely on a single competition-entry decision, which means they do not observe how noisy individuals' decisions are. In our experiment, we elicit choices in five different settings, each involving an MPL with ten different choices between competitive and individual pay. This approach provides us with much more data, allowing us to better assess the role of decision errors. Additionally, in some treatments, we repeated the experiment with the same participants across different weeks, allowing us to assess the stability of our measures of preferences for competition over time.

Finally, while the literature on preferences for competition has grown substantially, there is limited understanding of how these preferences are influenced by the competitiveness of a situation. We explore this question by varying the number of competitors. There are a few studies exploring the effect of group size on effort provision in contests (e.g., Sheremeta, 2011; Lim et al., 2014; Dechenaux et al., 2015; List et al., 2020). We contribute to this line of research by studying the impact of group size on competition entry decisions and, more specifically, on individuals' preferences for competition.

We find strong evidence that most individuals have preferences for competition. Most participants are either persistently competition loving or competition averse across different settings, and their decisions within these settings are consistent with expected utility maximization. For example, between 14% and 31% of participants switch from competitive to individual pay at points that imply they are willing to forgo more than 5% of the utility they would gain from winning the competition to avoid performing under competitive conditions. Conversely, between 29% and 47% of participants are willing to forgo at least 5% of the utility gain of winning to ensure they perform under competition. In monetary terms, these values imply that most participants are prepared to pay the utility equivalent of at least €1.35 to either avoid or ensure they compete, which is a non-negligible amount given that median earnings are around €15.00. Additionally, we find that the test-retest correlation of our measures of preferences for competition is comparable to that of risk preferences, suggesting that these preferences are stable over time.

Our findings also reveal two other intriguing patterns. First, individuals become more competition loving when competing in groups of six compared to groups of three. Second, while men choose to enter competitions more often than women, this difference is not due to their preferences for competition, as they are quite similar across genders. Instead, we find that the commonly reported gender difference in competition entry is better explained by gender differences in risk preferences and beliefs.

2. Literature review

This paper contributes to the literature on measuring preferences for competition. Starting with the seminal paper of Niederle and Vesterlund (2007), there have been many papers studying preferences for competition, especially in the context of gender differences (Niederle and Vesterlund, 2011; Niederle, 2017). Dariel et al. (2017) reviews the papers based on slight variations of the Niederle and Vesterlund (2007) experimental design. In Table A1 in the Appendix A, we list these papers and whether they differ from the original design. We also include a few more recent papers that share design choices with our experiment. We concentrate on three aspects of the experimental design: the way participants choose to compete, the method used to elicit beliefs, and the way risk preferences are elicited.

Most papers use a binary competition-entry choice to measure preferences for competition. However, like us, some papers use an MPL. Petrie and Segal (2017) use an MPL to elicit the prices at which one obtains a gender balance in competition entry. In their experiment, participants choose seven times between individual and competitive pay. Individual pay is always \$0.50 per correct sum, while competitive pay varies from \$0.75 to \$2.25 per correct sum (for the group's winner). Ifcher and Zarghamee (2016) use a similar approach, but they keep competitive pay constant at \$2.00 per correct sum and vary individual pay from \$0.00 to \$2.00 per correct sum. Jung and Vranceanu (2019) also uses an MPL, but instead of varying prices per correct sum, they elicit the lump-sum payment that makes participants indifferent between individual and competitive pay. Saccardo et al. (2018) develop a continuous version of the competition-entry choice. In their design, participants choose the percentage of their compensation derived from individual pay and the percentage derived from competitive pay. All these papers find that the choice to compete is sensitive to changes in incentives. In addition, they report considerable heterogeneity at the point at which individuals switch from individual to competitive pay. In terms of gender differences, these papers find men are more willing to compete than women, but there is substantial overlap. Namely, there is a considerable fraction of competition-loving women and a noticeable fraction of competition-averse men. Compared to our paper, these papers use only one MPL, and their individual pay is not a lottery.[1]

There is also some variation in methods for belief elicitation in the literature on preferences for competition. As seen in Table A1, most papers elicit beliefs by asking participants to guess their rank within their group and reward them with an additional payment if they guess correctly. The drawback of this method is that ranks do not map to a unique probability of winning the competition, which is necessary to calculate the participants' expected earnings

[1] Another related paper is Dohmen and Falk (2011). In their paper, participants choose between a fixed payment (i.e., independent of performance) and a variable payment, which could be a piece rate, a tournament rate, or revenue sharing. In addition to their choice, participants use a hypothetical MPL to indicate their willingness to choose the variable payment.

from competing.[2] For this reason, we opted for directly eliciting the participants' expected probability of winning. A few other papers do so as well. Most of them also use a binarized scoring rule to incentivize participants to provide accurate responses (Berlin and Dargnies, 2016; Petrie and Segal, 2017; Reuben et al., 2017; Fallucchi et al., 2020; van Veldhuizen, 2022), while a couple choose not to incentivize at all (Buser et al., 2017b; Saccardo et al., 2018).[3] Although these papers can calculate the expected earnings of winning the competition, the fact that they use a binary decision to compete means that, unlike us, they cannot estimate the strength of the participants' preferences for competition. The exception is Saccardo et al. (2018), who use a continuous measure of competition entry and a (non-incentivized) measure of the probability of winning. As mentioned above, compared to Saccardo et al. (2018), our design is not confounded by risk preferences, and it allows us to study the consistency with utility maximization and stability over time of competition-entry choices.

To identify preferences for competition, one must also account for the effect of risk preferences. Most papers in this literature do so with an independent risk-preference elicitation task such as the MPL of Holt and Laury (2002), the lottery choice of Eckel and Grossman (2002), the Bomb Risk Elicitation Task of Crosetto and Filippin (2013), or the self-reported question of Dohmen and Falk (2011) (see Table A1). Another approach was introduced by Niederle and Vesterlund (2007). In this approach, participants are given an additional opportunity to be paid again for their past performance and can decide whether they want to be paid according to individual or competitive pay. Since Niederle and Vesterlund (2007) define preferences for competition as a preference to seek or avoid performing in competitive environments, they argue that this additional choice captures the riskiness of the competitive pay but is unaffected by preferences for competition. Differently from these papers, instead of accounting for risk preferences with a different task, we change the reward structure of individual pay to ensure that risk preferences do not play a role in the decision to compete. To our knowledge, Geraldes (2020) and Molnar and Paolacci (2024) are the only other papers that use a lottery as individual pay. However, our experimental design differs in multiple ways. For instance, Geraldes (2020) uses a single binary

[2]Other approaches include eliciting the participants' expected number of correct sums for themselves and/or others (e.g. Dargnies, 2012; Wozniak et al., 2014; Brandts et al., 2015), their self-reported assessment of their relative performance (e.g. Bönte et al., 2017; Buser et al., 2024), and their ranking within the session (e.g. Cárdenas et al., 2012). These approaches also cannot be easily mapped to a probability of winning the tournament.

[3]There is an ongoing discussion on the best way to describe the binarized scoring rule to participants. Danz et al. (2022) propose that simpler descriptions outperform more complex ones. We simplify the binarized scoring rule by giving participants an interactive graphical interface to calculate incentives and compute how a percentile ranking translates into the probability of winning the competition (see Section 3 for details).

competition-entry choice and does not elicit beliefs with a proper scoring rule while Molnar and Paolacci (2024) allows participants to cheat in the real-effort task.

Lastly, two important papers question the existence of a gender gap in preferences for competition. Gillen et al. (2019) point out that identifying gender differences in preferences for competition by regressing the choice to compete on a gender dummy plus measures of risk preferences, beliefs, and performance can result in the overestimation of gender differences in preferences for competition due to measurement error in the control variables. They show that accounting for noise in the measurement of control variables reduces the magnitude and significance of the coefficient of the gender dummy in the Niederle and Vesterlund (2007) design. van Veldhuizen (2022) introduces new treatments that remove the role of competition and overconfidence from the decision to compete. He compares gender gaps in these treatments to that in the Niederle and Vesterlund (2007) design to identify under what conditions the gender gap in the decision to compete disappears. As Gillen et al. (2019), van Veldhuizen (2022) conclude that the gender gap in competition is mainly captured by gender differences in risk preferences and beliefs. Conceptually, our work differs from these papers in that we do not focus on identifying the sources of the gender difference in decisions to compete. Instead, we focus on measuring individuals' preferences for competition directly.

3. Experimental design

Our experiment is based on the experimental design developed by Niederle and Vesterlund (2007) to measure preferences for competition. Similar to their approach, participants in our study are randomly assigned to groups and perform an addition task.[4] The task involves summing four two-digit numbers within four minutes. The integers are randomly drawn from a uniform distribution with support 10 to 99. Participants are provided with scratch paper but are not allowed to use calculators. The participant who correctly solves the most sums in their group is declared the winner, with ties being broken randomly. Our elicitation of preferences for competition is based on the following two tasks.

[4]Unlike Niederle and Vesterlund (2007), our participants cannot physically see others in their group. They are only aware that they are matched with other participants in the session.

3.1. Belief elicitation task

In this task, participants report their belief about being their group's winner. We incentivize belief elicitation using a robust scoring rule (Karni, 2009). Specifically, participants take part in a lottery for a prize of €20. The probability of winning the prize depends on their stated belief and whether they are their group's winner. For a stated belief of being the winner, b_i, participant i has a probability of $1-(1-b_i)^2$ of getting the prize if they are the winner, and a probability of $1-b_i^2$ if they are one of the losers. This belief elicitation method is easy to implement, incentive-compatible for a wide range of risk preferences, and has been shown to outperform other elicitation methods (Gächter and Renner, 2010; Wang, 2011; Hossain and Okui, 2013; Harrison and Phillips, 2014; Trautmann and van de Kuilen, 2015). To facilitate understanding, participants submit their beliefs using an interactive graphical interface that automatically calculates the probabilities of getting the prize and the associated expected earnings for any selected belief. Additionally, we provide participants with easy-to-follow examples illustrating why it is optimal to report truthfully. The instructions for this task, including a screenshot of the graphical interface, are available in Appendix F.

Another feature of our belief elicitation task is that participants can answer the belief elicitation question by indicating their expected percentile ranking in the performance distribution. Upon selecting a percentile, participants are shown the probability of being a winner in a randomly formed group. This feature should help participants who struggle to calculate compound probabilities by allowing them to respond in terms of an easy-to-understand ranking. We also provide participants with a table displaying the probability of being their group's winner for every percentile in the performance distribution (see the instructions in Appendix F).[5]

3.2. Payment-scheme choices

Before performing the addition task, participants choose how they want to be paid per correct sum. Specifically, they choose between *competitive pay* and *individual pay* in five independent decision sets. Each decision set is a multiple price list (MPL) containing ten rows, with each row presenting a choice between competitive pay (left) and individual pay (right). After making their choices, one row from one decision set is randomly selected and implemented. Subsequently,

[5] The probabilities of being a group's winner are calculated assuming that participants in a session are randomly drawn from a continuous performance distribution. We do not find that performance in the first or the second addition tasks varies by session (Kruskal-Wallis tests, $p > 0.118$ for groups of three and $p > 0.137$ for groups of six).

Table 1. Example of a decision set

Notes: Example of a decision set with a high amount $\pi^H = 4$, a low amount $\pi^L = 1$, and a range of probabilities for individual pay between 0.17 to 0.44.

	Competitive pay		**Individual pay**			
Row	Win (€)	Lose (€)	Win (€)	P(Win)	Lose (€)	P(Lose)
1	4	1	4	0.17	1	0.83
2	4	1	4	0.20	1	0.80
3	4	1	4	0.23	1	0.77
4	4	1	4	0.26	1	0.74
5	4	1	4	0.29	1	0.71
6	4	1	4	0.32	1	0.68
7	4	1	4	0.35	1	0.65
8	4	1	4	0.38	1	0.62
9	4	1	4	0.41	1	0.59
10	4	1	4	0.44	1	0.56

participants perform the addition task knowing the decision set and row that was randomly selected.

Under competitive pay, participants earn a high amount π^H per correct sum if they are their group's winner and a low amount π^L per correct sum otherwise. Under individual pay, participants earn π^H per correct sum with some probability p and π^L per correct sum with probability $1-p$. Within a decision set, the probability of earning π^H in individual pay increases as one goes from the first to the tenth row. The values of π^H and π^L are constant within a decision set but vary across decision sets. Table 1 displays an example of a decision set where $\pi^H = 4$, $\pi^L = 1$, and p ranged from 0.17 to 0.44.

3.3. Treatment variations

We use a 2×2 treatment design. The first treatment variation consists of varying the timing of the belief elicitation task. In treatment *Belief-first*, participants first complete the belief elicitation task, followed by the payment-scheme choice, and ending with the addition task. In treatment *Choice-first*, participants first make the payment-scheme choice, followed by the addition task, and ending with the belief elicitation task. This treatment variation serves as a robustness check to see whether our proposed method to elicit preferences for competition is sensitive to the sequence in which beliefs and choices are elicited. Most papers in the literature on preferences for competition elicit beliefs after the decision to compete (as in Niederle and Vesterlund, 2007). A few papers elicit beliefs before this decision (e.g., Mayr et al., 2012;

Almås et al., 2016; Buser et al., 2024; van Veldhuizen, 2022) but to the best of our knowledge, this is the only paper where participants are randomly assigned to *Belief-first* and *Choice-first* treatments.[6] We implement this treatment variation between subjects.

Notably, the timing at which we elicit beliefs also affects how we construct the MPLs of the payment-scheme choice. There are two desirable conditions to accurately measure preferences for competition. First, the change in probabilities from one row to the next in the MPLs should not be too large. Second, the participants' belief of being their group's winner is contained within the range of probabilities in the MPLs. To meet these conditions, the probabilities in the MPLs are based on previous stages. Specifically, in treatment *Belief-first*, the range of probabilities in the MPLs is centered around the participants' elicited belief of being their group's winner. In treatment *Choice-first*, we narrow the range of probabilities by giving participants two additional decision sets in which they choose between competitive and individual pay. In the first additional decision set, the probabilities for individual pay range from 0.05 to 0.95 in steps of 0.10. Based on the number of times a participant chooses competitive pay in that first set, the probabilities for individual pay in the second additional set range from either 0.05 to 0.50, 0.30 to 0.70, or 0.50 to 0.95 in steps of 0.05. The five decision sets used to measure preferences for competition are centered around the probability at which participants switch from competitive pay to individual pay in this second additional decision set. Across decision sets, we randomize the position of the belief or switching probability from two rows above to two rows below the fifth row of the MPL. This way, participants face a different range of probabilities across decision sets.[7] Appendix B provides a more detailed description of the procedure used to determine the probabilities in individual pay and the high and low amounts in each decision set.

The second treatment variation consists of varying the size of the group in which participants compete. We implemented groups of *three* and *six* participants. This treatment variation is meant to study how individuals' preferences for competition are affected by how competitive a situation is since, intuitively, competing against five others is a more competitive situation than competing against two others.[8] This treatment variation is implemented within subjects. In

[6] Some papers measure beliefs both before and after the choice to compete, but they also give participants feedback on their relative performance (e.g., Berlin and Dargnies, 2016). In Appendix A, we provide a list of papers that follow the design of Niederle and Vesterlund (2007) and describe their preferences for competition, beliefs, and risk preferences elicitation methods.

[7] As with many elicitation methods (e.g., van de Kuilen and Wakker, 2011; Toubia et al., 2013; Chapman et al., 2019), participants are not informed how their previous choice affects future decision sets.

[8] We are unaware of a formal definition of how competitive a situation is. One can think of other ways to vary the competitiveness of a situation, such as higher stakes, matching participants to compete with others of similar

other words, all participants complete the belief elicitation task, payment-scheme choice, and addition task twice: once for a group of three and once for a group of six. The order of the group size is counterbalanced across sessions. Moreover, note that participants do not receive feedback on their relative performance until the end of the experiment.

3.4. Experimental procedures

The experiment was conducted in April 2019 at the [blinded for review] University. We ran in 11 sessions with a total of 224 participants (133 women and 91 men). We recruited participants with the online recruitment system ORSEE (Greiner, 2015), and the experiment was programmed and run using z-Tree (Fischbacher, 2007).

All participants signed an informed consent form before participating in the experiment. The experiment consisted of five parts: the two belief elicitation tasks, the two payment-scheme choices with their respective addition tasks, and a final risk elicitation task.[9] Instructions for each part were provided at the beginning of the respective part. At the end of the experiment, one part was randomly selected for payment. Additionally, participants completed an unincentivized practice round of three minutes to familiarize themselves with the addition task and completed a demographics questionnaire that included their gender. On average, participants received €29.08 (including a €10 show-up fee). The instructions for the experiment can be found in Appendix F.

4. Measuring preferences for competition

In this section, we describe the conceptual framework used to measure preferences for competition. We assume that participants' preferences can be represented by a utility function $U(\pi_i, C)$ that depends on the monetary payoffs π and whether a participant is performing under competitive ($C = 1$ or 1) or individual ($C = 0$) pay. More specifically, We assume the following utility function for participant i:

$$U(\pi_i, C) = u_i(\pi_i) + C\theta_i,$$

ability, or implementing affirmative action policies. We opted to vary group size because these other variations might introduce confounding factors or change the degree of competition differently for different participants.

[9] The risk elicitation task allows us to compare choices in decision sets designed to measure preferences for risk and competition. We present this analysis in Appendix E and discuss these comparison later on.

where π_i is the monetary value of performing the task (i's number of correct sums multiplied by either π^H or π^L) and θ_i is the parameter that captures i's non-pecuniary utility of performing under competition. As usual, we assume $u_i'(\pi_i) > 0$ and individual risk preferences are represented by the curvature of the utility function, $u_i''(\pi_i)$. Hence, we assume separability between preferences for competition and risk.

Under these assumptions, i is indifferent between competitive and individual pay if

$$b_i u_i\big(x_i \cdot \pi^H\big) + (1-b_i) u_i\big(x_i \cdot \pi^L\big) + \theta_i = p u_i\big(x_i \cdot \pi^H\big) + (1-p) u_i\big(x_i \cdot \pi^L\big),$$

where b_i is i's mean belief of being their group's winner in competitive pay, p is the probability of obtaining π^H in individual pay, and x_i is i's expected number of correct sums in the addition task.Note that using the same π^H and π^L for both payment schemes ensures that we can solve for θ_i irrespective of i's risk preferences. Specifically, we get

$$\theta_i = (p - b_i)\big[u_i\big(x_i \cdot \pi^H\big) - u_i\big(x_i \cdot \pi^L\big)\big].$$

This feature of our design addresses some of the concerns raised by Gillen et al. (2019) and van Veldhuizen (2022) as our setting does not rely on assuming that risk and competitive preferences are orthogonal.

We measure preferences for competition by eliciting the probability at which i is indifferent between competitive and individual pay, $\hat{p}_i$. Intuitively, if $\hat{p}_i > b_i$, then i is giving up a higher chance of winning a large amount under individual pay for a lower chance of winning the same amount under competitive pay, implying that i must 'like' competing. Conversely, if $\hat{p}_i < b_i$, then i is giving up a higher chance of winning under competitive pay for a lower chance of winning under individual pay, implying that i must 'dislike' competing.

A challenge with estimating θ_i directly is that it requires us to make assumptions about the functional form of $u_i(.)$. Moreover, even if we assume all participants share the same functional form, differing degrees of risk aversion imply that it is not straightforward to compare values of θ_i across participants. For these reasons, we opt for a normalized measure of participants' preferences. More precisely, we set $u_i\big(x_i \cdot \pi^H\big) = 1$ and $u_i\big(x_i \cdot \pi^L\big) = 0$, which gives us the following measure of participant i's preferences for competition in decision set t:

$$\omega_{it} = \hat{p}_{it} - b_i,$$

where $\hat{p}_{it}$ is the midpoint between the probability of the row at which i switched from competitive pay to individual pay and the probability of the preceding row, and b_i is i's reported belief of being their group's winner. Measuring ω_{it} has the advantage that it is simple to elicit and has an intuitive interpretation as the fraction of the utility increase due to winning π^H instead of π^L that a participant is willing to forgo to perform under competitive pay. This allows us to make meaningful comparisons across participants with utility functions that are potentially of different functional forms. An $\omega_{it} > 0$ implies i is competition loving whereas $\omega_{it} < 0$ implies i is competition averse.

5. Results

Descriptive statistics of the participants' performance in the addition task, their beliefs of being their group's winner, and their degree of overconfidence are presented in Table C1 in Appendix C. On average, participants correctly answer 11.4 sums and believe they have a 44.9% chance of winning, indicating they are overconfident as they overestimate their chances of winning by 19.9%.[10] If we look at treatment differences, we find that beliefs of being the group's winner systematically change between group sizes in the expected direction: participants exhibit higher beliefs of winning in groups of three compared to groups of six (53.8% vs. 36.0% on average). Group size does not influence performance in the addition task (11.4 vs. 11.3 sums) or the participants' overconfidence (20.4% vs. 19.4%).[11] Unlike group size, the moment at which beliefs are elicited does not have a noticeable effect. On average, participants in the *Belief-first* and *Choice-first* treatments exhibit similar beliefs of being their group's winner (42.6% vs. 47.0%), performance in the addition task (10.9 vs. 11.7 sums), and overconfidence (19.7% vs. 20.1%).[12] Given these findings, we pool the data for the *Belief-first* and *Choice-first* treatments in our main analysis. However, in Appendix C, we present the results separately for these treatments as a robustness check. The results are not quantitatively different using the ungrouped data.[13]

[10] To measure overconfidence, we simulate each participant's probability of winning by rematching participants 100,000 times into groups of three or six to observe how often each participant wins, given their observed performance in a particular group size. Overconfidence is calculated as the difference between participants' belief of being their group's winner and their estimated probability of winning.

[11] Participants' beliefs of being their group's winner are significantly higher in groups of three than in groups of six in both the *Belief-first* and *Choice-first* treatments (Wilcoxon signed-rank tests, $p < 0.001$ in both cases). By contrast, there are no statistically significant differences in the number of correct sums (Wilcoxon signed-rank tests, $p > 0.928$) or overconfidence (Wilcoxon signed-rank tests, $p > 0.610$).

[12] In both group sizes, there are no statistically significant differences between *Belief-first* and *Choice-first* in the number of correct sums (Mann-Whitney U tests, $p > 0.396$), the belief of being the group's winner (Mann-Whitney U tests, $p > 0.144$), and the degree of overconfidence (Mann-Whitney U tests, $p > 0.657$).

[13] The order in which participants face the two group sizes is counterbalanced. Hence, we also pool both orders for the main analysis. Nevertheless, we show in Appendix C that the paper's main results remain unaffected when considering each order separately.

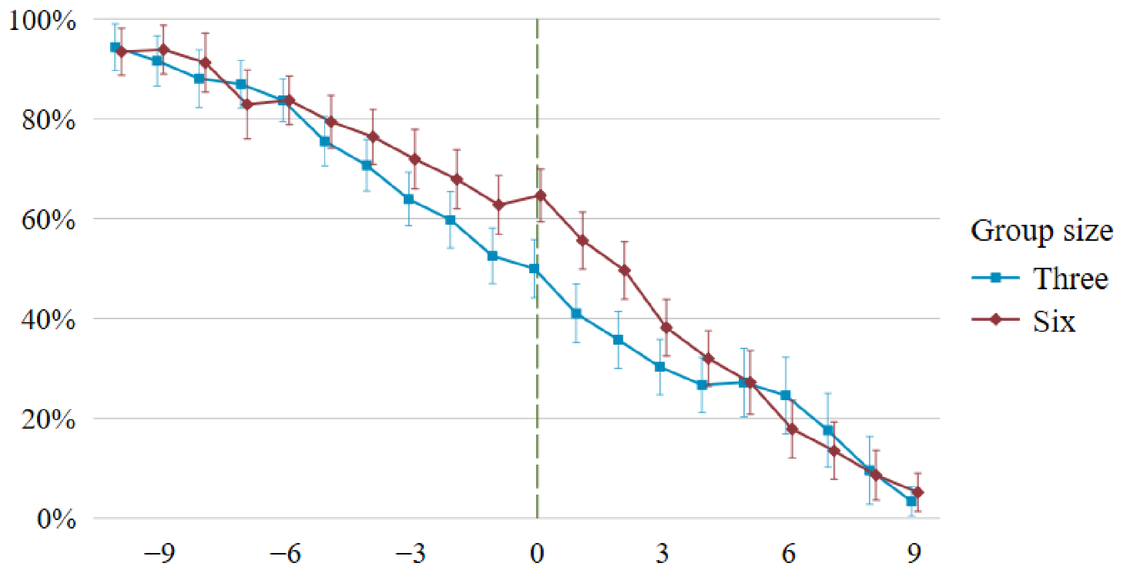


Figure 1. Fraction of competitive pay choices per normalized row in the decision sets

Notes: Rows within decision sets are normalized such that row zero corresponds to the highest row where the probability of earning the high amount in individual pay surpasses the participant's belief of being their group's winner. 95% confidence intervals are computed using a linear probability model regressing the payment-scheme choice on dummy variables for the normalized row values, with standard errors clustered by participant. For each group size, there are 224 participants making ten decisions in each of five decision sets.

5.1. Payment-scheme choices

Before estimating the distribution of preferences for competition, we first analyze the participants' payment-scheme choices. Figure 1 depicts the fraction of competitive pay choices per row of the decision sets. Since decision sets varied between participants depending on their beliefs, we normalize rows such that row zero for participant i in decision set t, $r_{it} = 0$, corresponds to the highest row where the probability of earning the high amount in individual pay exceeds i's belief of being their group's winner. In other words, if $r_{it} < 0$, the expected value of competitive pay is higher than that of individual pay, and conversely, if $r_{it} \geq 0$, the expected value of competitive pay is lower than that of individual pay. If participants have no preferences for competition, they should choose competitive pay for all rows $r_{it} < 0$ and switch to individual pay for rows $r_{it} \geq 0$.

Figure 1 demonstrates that the expected payoff difference between competitive and individual pay strongly influences participants' payment-scheme choice. As individual pay becomes relatively more attractive, the fraction of participants choosing competitive pay gradually decreases. However, it is notable that a substantial number of participants choose competitive pay when the expected value of individual pay is higher. For instance, in row 1, 41.1% of participants

Table 2. Consistency of switching behavior with utility maximization within decision sets

Notes: Fraction of decision sets classified as consistent and inconsistent with expected utility maximization. Inconsistent decision sets contain *multiple switches* or a unique *non-monotonic switch* from individual to competitive pay. Consistent decision sets contain a *single switch* from competitive to individual pay or *no switch*. For each group size, there are 224 participants and a total of 1120 decision sets.

	Group size	
	Three	Six
Behavior inconsistent with utility maximization		
Multiple switches	3.1%	3.9%
Non-monotonic switch	1.0%	0.7%
Behavior consistent with utility maximization		
Single switch	77.0%	76.0%
No switch	18.9%	19.4%

in groups of three and 55.6% of participants in groups of six choose competitive pay, suggesting that they like to compete. Conversely, there are a substantial number of participants who choose individual pay even though the expected value of competitive pay is higher, as seen in row -1, where 47.6% of participants in groups of three and 37.1% of participants in groups of six choose individual pay, suggesting that they dislike competing. The figure also shows there are differences between group sizes, with more participants choosing competitive pay in groups of six than in groups of three.[14]

5.2. Switching behavior

Since we did not impose a single-switch restriction in the decision sets, we can evaluate whether behavior within sets is *consistent with expected utility maximization* in the absence of errors by looking at the number and direction in which participants switch within a decision set. Specifically, we classify decision sets as *inconsistent* if they contain multiple switches or a unique non-monotonic switch, meaning a switch from the payment scheme with the higher expected value (individual pay) to the one with the lower expected value (competitive pay). Similarly, we classify decision sets as *consistent* if they contain a single switch from competitive to individual pay or if there is no switch and all choices correspond to either competitive or individual pay.

Table 2 presents the fraction of consistent and inconsistent decision sets for the two group sizes. In groups of three, 95.9% of the decision sets display switching behavior consistent with

[14] A linear probability model regressing the payment-scheme choice on dummy variables for the normalized row values, interacted with a group size dummy variable, and clustering standard errors on participants, indicates that the difference between group sizes is statistically significant between rows -3 and 3 ($p < 0.020$).

Table 3. Consistency of switching behavior with utility maximization across decision sets

Notes: Fraction of participants according to the number of consistent and inconsistent decision sets. Inconsistent decision sets contain multiple switches or a unique switch from individual to competitive pay. Consistent decision sets contain a single switch from competitive to individual pay or no switch. For each group size, there are 224 participants.

	Group size	
	Three	Six
Behavior inconsistent with utility maximization		
At least one inconsistent decision set	8.9%	11.6%
Behavior consistent with utility maximization		
Five consistent decision sets	91.1%	88.4%
Five consistent decision sets and a switch in a majority of sets	75.9%	75.0%
Five consistent decision sets and a switch in all five sets	54.5%	45.5%

expected utility maximization, with 77.0% of sets featuring a single switch and 18.9% exhibiting no switches. A nearly identical pattern is observed in groups of six, where 95.4% of decision sets are classified as consistent: 76% with a single switch and 19.4% with no switches.[15]

Next, we examine participants' switching behavior across the five decision sets. The top section of Table 3 shows the percentage of participants with at least one inconsistent decision set. It reveals that only a small minority——8.9% when competing in groups of three and 11.6% when competing in groups of six—exhibit inconsistent decision sets. In other words, the vast majority of participants—91.1% in groups of three and 88.4% in groups of six—display switching behavior consistent with expected utility maximization in all five sets. The bottom section of Table 3 further shows that approximately three-quarters of participants in both group sizes make a single switch from competitive to individual pay in a majority of sets (three or more) and around half do so in all five sets.[16]

To sum up, participants' choices to compete are consistent with expected utility maximization. In nearly all decision sets, they either switch once from competitive to individual pay or do not switch at all. Remarkably, around 90% of the participants display this type of consistency with utility maximization in all five decision sets.

[15] The number of inconsistent decision sets per participant does not significantly differ across group sizes (Wilcoxon signed-rank test, $p = 0.255$).

[16] The fraction of participants with at least one inconsistent decision set does not vary significantly across group sizes (McNemar's χ^2 test, $p = 0.327$). The same is true for the fraction of participants who switch from competitive to individual pay in a majority of sets (McNemar's χ^2 test, $p = 0.894$). The fraction of participants who switch from competitive to individual pay in all sets is significantly higher in groups of three (McNemar's χ^2 test, $p = 0.033$).

5.3. Preferences for competition

In this subsection, we analyze when participants choose to switch from competitive to individual pay. First, we assess whether a participant likes or dislikes competition by looking at the median row at which they switch across the five decision sets, which we denote as $\bar{r}_i$. Second, we examine participants' preferences for competition using the median value of ω_{it} across the five decision sets, which we denote as $\bar{\omega}_i$ (see section 4).[17] To ensure an accurate representation of participants' preferences for competition, we focus only on those whose switching behavior is consistent with expected utility maximization in all five decision sets: 91.1% of participants in groups of three (204 out of 224) and 88.4% of participants in groups of six (198 out of 224). Nonetheless, in Appendix C, we show that our results are robust to these choices. Figure C1 illustrates the results using the mean row and value of ω_{it}s instead of the median. Tables C4 and C5 summarize the main results when we include all participants and when we exclude participants who do not switch in some decision sets.

Panel A of Figure 2 depicts the distribution of $\bar{r}_i$ for each group size, revealing that only a minority of participants have a median switching row of zero: 14.7% in groups of three and 13.1% in groups of six. In groups of three, 43.1% of participants have a positive $\bar{r}_i$, suggesting they like competition, while a similar 42.2% have a negative $\bar{r}_i$, suggesting they dislike competition. For groups of six, 59.1% have a positive $\bar{r}_i$ and 27.8% a negative one.

Panel B of Figure 2 displays the distributions of $\bar{\omega}_i$. These distributions confirm that a majority of participants are willing to sacrifice earnings to either engage in or avoid competition. For instance, in groups of three, 30.9% of participants have an $\bar{\omega}_i$ that indicates they are willing to forgo more than 5.0% of the utility gain of receiving the high amount to avoid competition. Conversely, 28.9% are willing to forgo at least 5.0% of this utility gain to ensure they compete. In groups of six, these percentages are 22.2% and 47.0%, respectively. To put these figures in perspective, the difference between getting the high and low amounts for the median participant amounts to €27 in groups of three and €42 in groups of six. Hence, a majority of participants are prepared to pay the utility equivalent of at least €1.35 in groups of three and €2.10 in groups of six to either avoid or ensure competition. These are non-negligible amounts considering that the median earnings in this task are €15.26 in groups of three and €14.00 in groups of six.

Figure 2 also demonstrates that increasing the number of competitors makes participants

[17] In sets where participants do not switch, we assign the highest value in the set for those who always choose competitive pay and the lowest value for those who always choose individual pay.

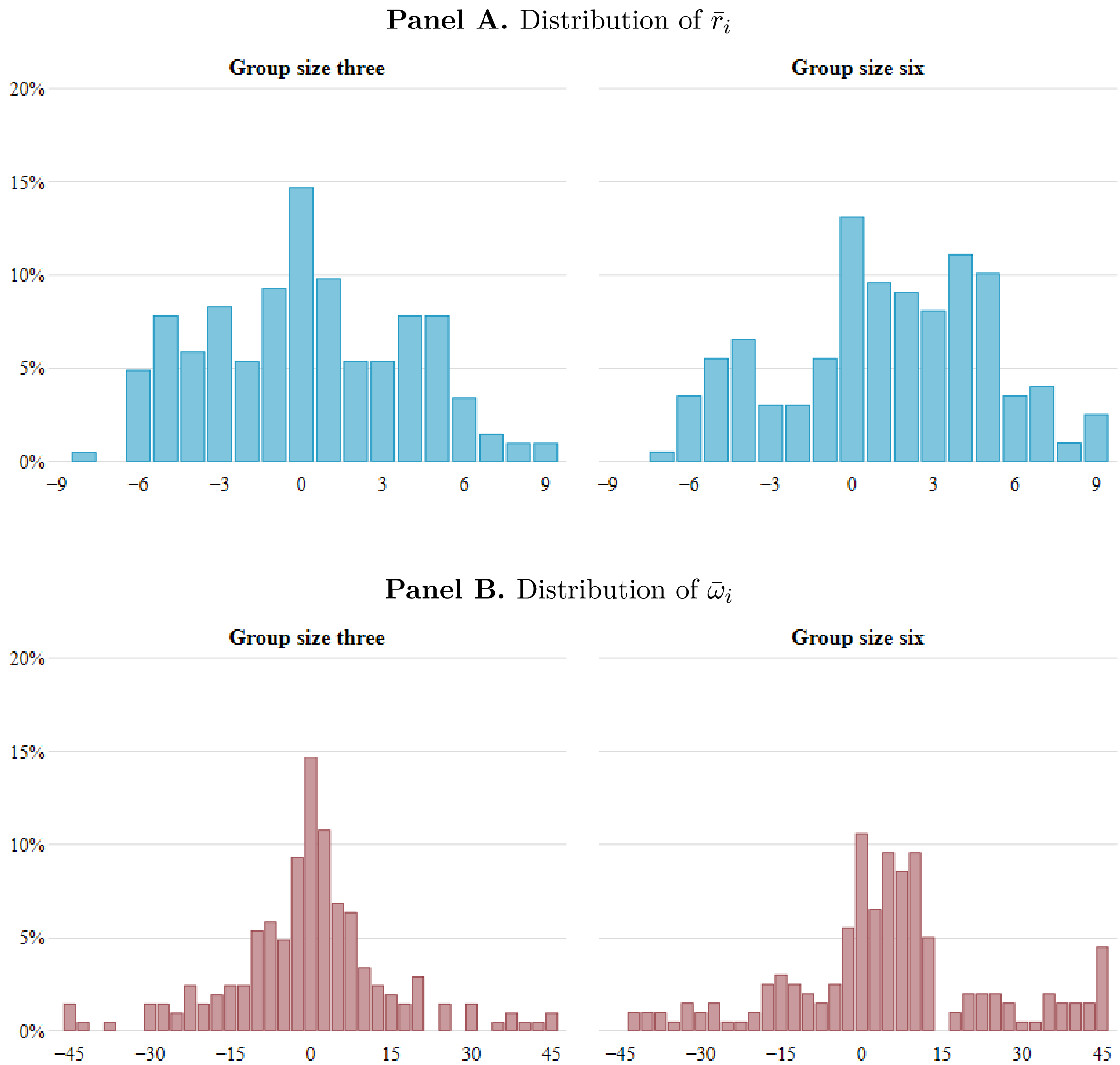


Figure 2. Distributions of measures of participants' preferences for competition

Notes: Panel A shows the distribution of $\bar{r}_i$, the median row at which participants switch from competitive to individual pay. Rows are normalized such that row zero corresponds to the highest row where the probability of earning the high amount in individual pay surpasses the participant's belief of being their group's winner. Panel B shows the distribution of $\bar{\omega}_i$, the median value of ω_{it} across decision sets, expressed as a percentage. ω_{it} represents the fraction of the utility gained from receiving the high instead of the low amount, $u_i(x_i \cdot \pi^H) - u_i(x_i \cdot \pi^L)$, participants are willing to forgo to either avoid or engage in competition. Data corresponds to participants whose switching behavior is consistent with expected utility maximization in all five decision sets. For visual purposes, the distribution in Panel B is censored at -45 and 45. For groups of three, 5 out of 204 participants fall outside this range, while for groups of six, it is 9 out of 198 participants.

more competitive. While, for groups of three, Wilcoxon signed-rank tests do not reject the null hypothesis that the distribution of $\bar{r}_i$ or $\bar{\omega}_i$ are centered around zero ($p = 0.727$ for $\bar{r}_i$ and $p = 0.387$ for $\bar{\omega}_i$), for groups of six this is clearly not the case ($p < 0.001$ for $\bar{r}_i$ and $\bar{\omega}_i$). Moreover, Wilcoxon signed-rank tests also reject the null hypothesis that the observed values of

$\bar{r}_i$ or $\bar{\omega}_i$ come from the same distribution in groups of three and six ($p < 0.001$ for $\bar{r}_i$ and $\bar{\omega}_i$).[18]

5.4. Persistence of competition-loving and competition-averse behavior

Next, we investigate whether our measures of participants' preferences for competition truly reflect their preferences or whether they are due to decision errors. As before, we conduct the analyses using the participants whose switching behavior is consistent with expected utility maximization in all five decision sets.

As we saw in Table 2, participants make very few errors in terms of switching behavior within decision sets. However, another way of thinking of errors is that participants switch from competitive to individual pay at the wrong row in the MPL. For example, if a participant truly has no preference for competition, they should switch to individual pay at $r_{it} = 0$, the row where the probability of receiving the high amount in individual pay surpasses their belief of being the group's winner. However, they could erroneously switch at rows above or below $r_{it} = 0$. This type of error could explain why the distribution of $\bar{r}_i$ is centered around zero, at least for groups of three (see Figure 2).

To assess whether switching errors are behind the observed distribution of $\bar{r}_i$, we analyze switching behavior across the five decision sets.[19] We start with the reasonable premise that switching errors are equally likely at rows above or below participants' intended switching row. It follows that if individual i makes a switching error in m sets, then the probability that i switches at a row *above* their intended row in x out of m sets is given by the probability mass function of the Binomial distribution $f(x, m, 0.5) = \binom{m}{x} 0.5^m$. Then, we consider the scenario where

[18] A potential explanation for the difference between group sizes is that participants exhibit a bias toward switching in the middle of MPLs, and the position of the competition-neutral row within MPLs differs by group size due to differences in beliefs. Given that the average position of the competition-neutral row is slightly higher in groups of three than in groups of six, we perform two checks to evaluate whether this difference influences the impact of group size on preferences for competition. First, we test whether ω_{it} differs by group size when restricting the sample to the 52.9% of decision sets where the competition-neutral row is located in the middle of the MPL (rows 5, 6, or 7). We find that the difference between group sizes narrows from 5.6 to 4.5 percentage points but remains statistically significant (Wilcoxon signed-rank test, $p = 0.002$). Second, we test once more whether ω_{it} differs by group size, but we restrict the sample to the 34.0% of decision sets where participants saw the competition-neutral row in exactly the same position in the MPL for both group sizes. Again, the difference between group sizes narrows (to 3.2 percentage points) but remains statistically significant (Wilcoxon signed-rank test, $p = 0.011$).

[19] Another simple way of using behavior across sets to assess whether it is driven by preferences is to examine variation in ω_{it}. If choices are mostly driven by (uncorrelated) decision errors, we should find relatively high within-participant variation. Conversely, if decision errors are minimal, we should find little variation within participants relative to between participants. In section C.4 of Appendix C, we show that variation in ω_{it} is much smaller within participants than between participants.

participants do not have preferences for competition. In this scenario, sets where participant i switches around $r_{it} = 0$ would represent correct switches, while sets where i switches away from $r_{it} = 0$ would be considered errors.

Table 4 displays the fraction of switches occurring above rows designated as *competition neutral* depending on the number of competition-neutral switches among the five decision sets.[20] For example, the first case corresponds to participants who did not switch at competition-neutral rows in any decision set, meaning they had five switches that could be either above or below the competition-neutral rows. Following the argument above, if participants are equally likely to switch above and below the competition-neutral rows by error, then we can calculate the probability of switching $x \in \{0, 1, 2, 3, 4, 5\}$ times above the competition-neutral rows and compare these probabilities to the observed switching behavior. In Panel A, we designate only $r_{it} = 0$ as the competition-neutral row. With this designation, participants switch at the competition-neutral row in 14.7% of all sets in groups of three and 13.1% in groups of six. In Panel B, we broaden the range of competition-neutral rows to include the adjacent rows $r_{it} = -1$ and $r_{it} = 1$. With the expanded designation, participants switch at the competition-neutral rows in 33.8% of all sets in groups of three and 28.3% in groups of six.

Table 4 demonstrates that the participants' switching behavior cannot be solely attributed to row-switching errors. Take Case I in Panel A, where participants do not switch at the competition-neutral row in any of the five sets. extreme outcomes, such as five switches or zero switches above the competition-neutral row, are predicted to be very rare, occurring only 3.1% of the time. However, these are the most commonly observed outcomes: in groups of three, 45.3% of participants always switch above the competition-neutral row, and 40.6% never do; in groups of six, 61.5% always switch above the competition-neutral row, and 26.2% always switch below. This pattern persists across other cases and when we broaden the set of competition-neutral rows in Panel B. For both Panel A and B, Chi-square goodness of fit tests reject the null hypothesis that the observed distribution of switches above the competition-neutral rows is drawn from the predicted distribution in all cases but two ($p < 0.011$ except for $p = 0.109$ in Case III of Panel A in groups of three and $p = 0.060$ in Case IV of Panel B in groups of three).

By and large, the evidence in Table 4 suggests that most participants reliably switch at rows either above or below the competition-neutral rows. We examine this pattern more closely by

[20]Sets in which participants do not switch at all are categorized as switches above the competition-neutral switch if participants always choose individual pay, and below if they always choose competitive pay.

Table 4. Fraction of switches above the competition-neutral rows depending on a participant's number of competition-neutral switches

Notes: Fraction of switches above rows considered as competition neutral depending on the number of competition-neutral switches a participant has in their five decision sets. In Panel A, only $r_{it} = 0$ is considered a competition-neutral row. In Panel B, $r_{it} = -1$ and $r_{it} = 1$ are also considered competition-neutral rows. Predictions are based on the assumption that switches that are not competition neutral are errors and are equally likely to occur above and below the competition-neutral rows. Data corresponds to participants whose switching behavior is consistent with expected utility maximization in all five decision sets.

Panel A. Competition-neutral rows defined as $r_{it} = 0$

	Number of switches at rows above $r_{it} = 0$					
	0	1	2	3	4	5
Case I: Participants with 0 switches at competition-neutral rows						
Prediction with 5 switching errors	3.1%	15.6%	31.3%	31.3%	15.6%	3.1%
Group size three ($n = 130$)	40.0%	2.3%	4.6%	2.3%	6.2%	44.6%
Group size six ($n = 123$)	26.0%	1.6%	0.8%	2.5%	8.1%	61.0%
Case II: Participants with 1 switch at competition-neutral rows						
Prediction with 4 switching errors	6.3%	25.0%	37.5%	25.0%	6.3%	
Group size three ($n = 41$)	34.2%	14.6%	12.2%	19.5%	19.5%	
Group size six ($n = 39$)	23.1%	15.4%	7.7%	7.7%	46.1%	
Case III: Participants with 2 switches at competition-neutral rows						
Prediction with 3 switching errors	12.5%	37.5%	37.5%	12.5%		
Group size three ($n = 17$)	29.4%	17.7%	35.3%	17.7%		
Group size six ($n = 22$)	22.7%	18.2%	22.7%	36.4%		
Case IV: Participants with 3 switches at competition-neutral rows						
Prediction with 2 switching errors	25.0%	50.0%	25.0%			
Group size three ($n = 11$)	63.6%	18.2%	18.2%			
Group size six ($n = 8$)	12.5%	12.5%	75.0%			

Panel B. Competition-neutral rows defined as $r_{it} = \{-1, 0, 1\}$

	Number of switches at rows above $r_{it} = 1$					
	0	1	2	3	4	5
Case I: Participants with 0 switches at competition-neutral rows						
Prediction with 5 switching errors	3.1%	15.6%	31.3%	31.3%	15.6%	3.1%
Group size three ($n = 85$)	50.6%	0.0%	0.0%	0.0%	0.0%	49.4%
Group size six ($n = 91$)	33.0%	0.0%	0.0%	0.0%	2.2%	64.8%
Case II: Participants with 1 switch at competition-neutral rows						
Prediction with 4 switching errors	6.3%	25.0%	37.5%	25.0%	6.3%	
Group size three ($n = 31$)	38.7%	9.7%	0.0%	19.4%	32.3%	
Group size six ($n = 32$)	18.8%	3.1%	3.1%	15.6%	59.4%	
Case III: Participants with 2 switches at competition-neutral rows						
Prediction with 3 switching errors	12.5%	37.5%	37.5%	12.5%		
Group size three ($n = 26$)	34.6%	11.5%	15.4%	38.5%		
Group size six ($n = 22$)	31.8%	4.6%	4.6%	59.1%		
Case IV: Participants with 3 switches at competition-neutral rows						
Prediction with 2 switching errors	25.0%	50.0%	25.0%			
Group size three ($n = 26$)	34.6%	26.9%	38.5%			
Group size six ($n = 28$)	28.6%	10.7%	60.7%			

Table 5. Fraction of participants who are persistently competition loving, persistently competition averse, and persistently competition neutral

Notes: Participants are classified as *persistently competition loving* if they switch above the competition-neutral rows in at least four sets, *persistently competition averse* if they switch below the competition-neutral rows in at least four sets, and *persistently competition neutral* if they switch at the competition-neutral rows in at least four sets. The remaining participants are classified as 'not defined.' In the first two columns, only $r_{it} = 0$ is considered a competition-neutral row. In the last two columns, $r_{it} = -1$ and $r_{it} = 1$ are also considered competition-neutral rows. Data corresponds to participants whose switching decisions are consistent with expected utility maximization in all five decision sets. There are 204 participants for groups of three and 198 for groups of six.

	Competition-neutral rows defined as:			
	$r_{it} = 0$		$r_{it} = \{-1, 0, 1\}$	
	Group size		*Group size*	
	Three	Six	Three	Six
Persistently competition loving	36.3%	52.0%	25.5%	40.4%
Persistently competition averse	33.8%	21.7%	27.0%	18.2%
Persistently competition neutral	2.5%	3.0%	17.6%	12.6%
Not defined	27.4%	23.2%	29.9%	28.8%

classifying participants based on their switching patterns across sets. We classify participants who switch above the competition-neutral rows in at least four out of five sets as *persistently competition loving*; those who switch below in at least four sets as *persistently competition averse*; and those who switch at the competition-neutral rows in at least four sets as *persistently competition neutral*. Participants who do not fit these criteria are categorized as 'not defined.'

Table 5 presents the percentage of participants classified as persistently competition loving, averse, and neutral, as well as those classified as 'not defined.' As before, we first classify participants considering $r_{it} = 0$ as the only competition-neutral row and then reclassify them, including rows $r_{it} = -1$ and $r_{it} = 1$ as competition-neutral rows.

Overall, a majority of participants are either persistently competition loving or persistently competition averse, which is not in line with there being no preferences for competition.[21] For both definitions of competition-neutral rows, Table 5 also reveals that the distribution of types is significantly different across group sizes (Stuart–Maxwell Marginal homogeneity tests, $p < 0.003$), with a higher fraction of participants classified are persistently competition loving

[21] Comparing our fractions of persistently competition loving and averse individuals to other studies is challenging. Most studies employ a binary measure of tournament entry to measure preferences for competition, and the few studies that use continuous measures do not have multiple measurements to assess persistence (see Table A1 in the Appendix). Buser et al. (2024), using a self-reported competitiveness scale ranging from 0 (not at all competitive) to 10 (extremely competitive), classify 18% of participants as competition averse, 35% as competition neutral, and 47% as competition loving. Hence, like us, they find that the modal individual is competition loving.

in groups of six compared to groups of three.

5.5. Discussion

So far, we have seen strong evidence that the behavior of a substantial number of participants is consistent with them having preferences for or against competition and is unlikely the result of errors in switching behavior within or across decision sets. Having said that, we acknowledge that our experimental design has some limitations.

It is clear that our measure of preferences for competition relies on an accurate measurement of participants' beliefs of being their group's winner. We use a well-accepted incentivized method to measure these beliefs. However, recent work on belief elicitation has shown that systematic errors can still occur (Danz et al., 2022). In section C.5 of Appendix C, we conduct two analyses to evaluate whether noise in belief measurement is driving our results. First, we use the two measures of beliefs per participant (one per group size) to check whether variance in beliefs correlates with variance in preferences for competition. We do not observe that participants with 'noisier' beliefs show more variation in their preferences for competition. Second, we apply the insight from Danz et al. (2022) that belief elicitation with the binarized scoring rule can lead to center bias in belief measurement. In our experiment, a center bias would imply that we overestimate ω_{it} for participants with beliefs above 50% and underestimate it for those with beliefs below 50%. However, we do not find that participants with beliefs above 50% have higher values of ω_{it} than those with beliefs below 50%. These analyses are reassuring, but one can still worry that the observed variation in ω_{it} could reflect noise in belief measurement rather than genuine differences in preferences for competition.

A second limitation is that, despite observing multiple decisions per participant, all decisions occur on the same day, which prevents us from assessing the stability of our measures over time. If we are capturing participants' preferences, we should expect substantial stability. Lastly, while we observe participants being persistently competition averse or competition loving within our setting, it remains an open question whether our measures of preferences for competition predict competitive behavior in other settings. We address these concerns with a follow-up experiment.

6. Follow-up experiment

We conducted a follow-up experiment to address some of the unresolved questions from our initial experiment. The follow-up experiment uses a similar design to the initial experiment but with some crucial differences.

6.1. Experimental design and procedures

A session of the follow-up experiment consists of three parts, with one part randomly selected to determine participants' session earnings.

The competition-entry choice

In part I, participants complete the well-known competition-entry choice introduced by Niederle and Vesterlund (2007) to study gender differences in competitiveness. Specifically, we first introduce participants to the addition task described in Section 3, giving them two minutes to familiarize themselves with it. Subsequently, we inform participants that they will perform the addition task under one of two payment schemes: individual pay, which gives them €1 per correct sum, or competitive pay, which gives them €3 per correct sum if they outperform two randomly selected participants and €0 otherwise.[22] After choosing their preferred payment scheme, participants perform the addition task. By giving the participants a competition-entry choice, we can compare our results to the literature and establish whether our measures of preferences for competition predict competitive behavior in this widely recognized benchmark. Additionally, the competition-entry choice informs us about the external validity of our measures as this choice is the one that has been shown to predict choices and outcomes outside the laboratory (see Lozano et al., 2023).

Payment-scheme choices and the probability of winning

In part II, participants are told they will perform the same addition task again, but first, they must choose between individual and competitive pay in five independent decision sets. As in the initial experiment, with competitive pay, participants earn π^H per correct sum if they outperform two other randomly selected participants and π^L otherwise. With individual pay, they earn π^H per correct sum with probability p and π^L with probability $1-p$. Each decision

[22]Due to budget considerations, we only use groups of three in the follow-up experiment.

set is an MPL with 10 choices. The values of π^H and π^L are constant within a decision set, while the probability of earning π^H in individual pay gradually increases from the first to the tenth choice. We use the same parameters to construct the five sets as in the initial experiment (for details, see Appendix B).

The key difference between the initial and follow-up experiments is that instead of eliciting participants' beliefs before their payment-scheme choices, we inform them of their probability of being the top performer in a randomly formed group of three. We explain to participants that these probabilities are calculated based on their performance in the addition task in Part I.[23] By directly informing participants of their probability of winning, we avoid any noise introduced by the belief elicitation task and mitigate the impact of belief-related biases such as overconfidence, difficulty calculating compound probabilities, or uncertainty about one's beliefs (i.e., ambiguity). To ensure an accurate measure of participants' preferences for competition, the probabilities used to construct the MPLs are centered around each participant's probability of winning. Once participants make their payment-scheme choices, a choice from one decision set is randomly selected and communicated to the participants, who perform the addition task once again.

Risk elicitation

In part III, participants engage in a risk-elicitation task consisting of four independent decision sets. Each decision set corresponds to an MPL of 10 choices between a certain amount, π^C, and a lottery that pays a high amount, π^H, with probability q and a low amount, π^L, with probability $1 - q$. Within a set, the certain amount, π^C, increases from the first to the tenth choice. The values of π^H, π^L, and q remain constant within a decision set but vary across sets. Appendix E contains a more detailed description of the procedure and the specific amounts used in each decision set. Once they have made their choices, one choice in a decision set is randomly selected to determine their earnings in this part. The risk elicitation task allows us to evaluate whether switching behavior and its consistency with utility maximization is similar in MPLs involving competition-entry choices and those involving risk.

[23]Participants' performance in Part I is an excellent predictor of their subsequent performance in Part II, as evidenced by a Pearson's correlation coefficient of $r = 0.785$ ($p < 0.001$).

Repetition

The last important difference between the initial and follow-up experiments is that in the follow-up, participants return a week later to repeat the same parts. Although earnings for each session are calculated independently, to ensure participants attend both sessions, they receive payment only at the end of the second session. Participants were informed at registration that they had to participate twice to receive payment. To make the sessions comparable, participants did not receive feedback about the outcomes in the first session. Hence, the only information revealed during the first session was their probability of winning. Repeating the session allows us to evaluate the stability of our measure of preferences for competition over time and compare it to the stability of the measure of risk preferences.

Experimental procedures

The follow-up experiment was conducted in December 2023 at the [blinded for review] University, adhering closely to the protocols of the initial experiment. A total of 118 participants (72 women and 46 men) attended two sessions held one week apart. Remarkably, there was no attrition between the first and second sessions. To maintain anonymity, each participant received a unique ID number during the first session, which they brought to the second session to receive payment. On average, participants earned €68.46 (including two show-up fees of €10 each). The instructions for the follow-up experiment can be found in Appendix G.

6.2. Replication of results from the initial experiment

Descriptive statistics of the participants' performance in the addition task and their choice to compete are presented in Table D1 in Appendix D. Below, we briefly show that the follow-up experiment replicates the main patterns observed in the initial experiment. The more detailed analysis is available in Appendix D.

- *Payment-scheme choices.* Figure D1 depicts the fraction of competitive pay choices per row in the decision set. As individual pay becomes relatively more attractive, the fraction of participants choosing competitive pay decreases. However, as in the initial experiment, many participants choose competitive pay when the expected value of individual pay is higher (e.g., in row 1, 42.8% choose competitive pay) and individual pay when the expected value of competitive pay is higher (e.g., in row -1, 34.9% choose individual pay).

- *Switching behavior.* Table D2 summarizes participants' behavior within decision sets, while Table D3 summarizes their switching behavior across decision sets. Similar to the initial experiment, we find that more than 95% of the decision sets exhibit behavior consistent with utility maximization. Additionally, around 90% of participants demonstrate this type of consistent behavior in all five sets. These fractions are not significantly different between the first and second sessions (Wilcoxon signed-rank test, $p > 0.424$).
- *Preferences for competition.* Panel A of Figure D2 depicts the distribution of $\bar{r}_i$, the median row at which participants switch across the five decision sets, while Panel B depicts the distribution of $\bar{\omega}_i$, the median value of ω_{it} across the five decision sets. As in the initial experiment, a majority of participants have non-zero values of $\bar{r}_i$ and $\bar{\omega}_i$, indicating they are willing to sacrifice money to either avoid competition or ensure they compete. The distributions across sessions look quite similar, although we see hints of a slight decrease in the values of $\bar{r}_i$ (from 1.72 to 0.85; Wilcoxon signed-rank test, $p = 0.077$) and $\bar{\omega}_i$ (from 2.86% to 1.15%; Wilcoxon signed-rank test, $p = 0.066$).
- *Persistence of competition-loving and competition-averse behavior.* Table D5 presents the percentage of participants classified as persistently competition loving, averse, and neutral, as well as those classified as 'not defined.' With the strict definition of competition-neutral behavior, around 60% of participants are classified as either persistently competition loving or persistently competition averse, with less than 6% classified as persistently competition neutral. With the broader definition of competition-neutral behavior, the fraction of participants classified as persistently competition neutral rises to around 25%, while the fraction classified as either persistently competition loving or persistently competition averse goes down to around 45%. These proportions do not vary significantly between the first and second sessions (Stuart–Maxwell Marginal homogeneity tests, $p > 0.458$).

The replication of the initial experiment's main findings increases our confidence that noise in belief elicitation is not driving the main results. The patterns of payment-scheme choices, the consistency of participants' switching behavior with utility maximization, and the large fraction of participants who are persistently competition loving or competition averse closely mirror those observed in the initial experiment. Remarkably, the average magnitudes of the measures of participants' preferences for competition are very similar despite participants seeing very different probabilities across experiments. Specifically, MPLs are centered on participants' beliefs in the initial experiment and on their probability of winning in the follow-up experiment, with the latter being about 20 percentage points lower due to overconfidence. That being said, one apparent difference between experiments is the larger variance in the distribution of preferences for competition in the initial experiment (cf. Figures 2 and D2). This higher variance is driven by the *Choice-first* treatment (see Table C4), suggesting that directly eliciting participants' beliefs is preferable to inferring them indirectly.

6.3. Stability

In this subsection, we evaluate the stability of our measures of preferences for competition over time. If preferences for competition are a stable trait and our experiment measures it accurately, we should see a significant correlation between the measures taken in the first and second sessions. To benchmark this correlation, we compare it to the correlation over time of measures of participants' risk preferences. In Appendix E, we present a detailed analysis of participants' risk preferences. Interestingly, we find that the consistency of switching behavior with utility maximization within and across sets is very similar for preferences for risk and competition.[24] Mirroring our analysis of preferences for competition, we use two measures of participants' risk preferences. The first measure corresponds to the median row at which participants switch across the four decision sets, $\bar{r}_i$. For the second measure, we calculate the CRRA coefficient that explains i's switching row in each set t and then use the median value across the four sets as the measure of participants' risk preferences, which we denote as $\bar{\alpha}_i$.

Figure 3 contains scatter plots illustrating the relationship between the first and second sessions of participants' preferences for competition and risk. Panels A and B show the median switching row of the competition and risk elicitation sets, respectively. Panel C shows the values of $\bar{\omega}_i$ and Panel D the values of $\bar{\alpha}_i$. To illustrate these correlations on a comparable scale, we standardized each variable to have a mean of zero and a standard deviation of one. Standardizing the variables also has the added advantage that the slope of the best-fit line corresponds to Pearson's correlation coefficient. The grids in the figure correspond to lengths of one standard deviation.

The figure illustrates a moderate correlation between participants' choices in the first and second sessions. The Pearson's correlation coefficient for the median switching row in the competition decision sets is 0.38 ($p < 0.001$), which is comparable to the correlation for the

[24] In both the initial and follow-up experiments, more than 95% of the sets used to elicit risk preferences display switching behavior consistent with expected utility maximization. Moreover, more than 90% of participants show this type of consistent behavior in all four decision sets. See Appendix E for details.

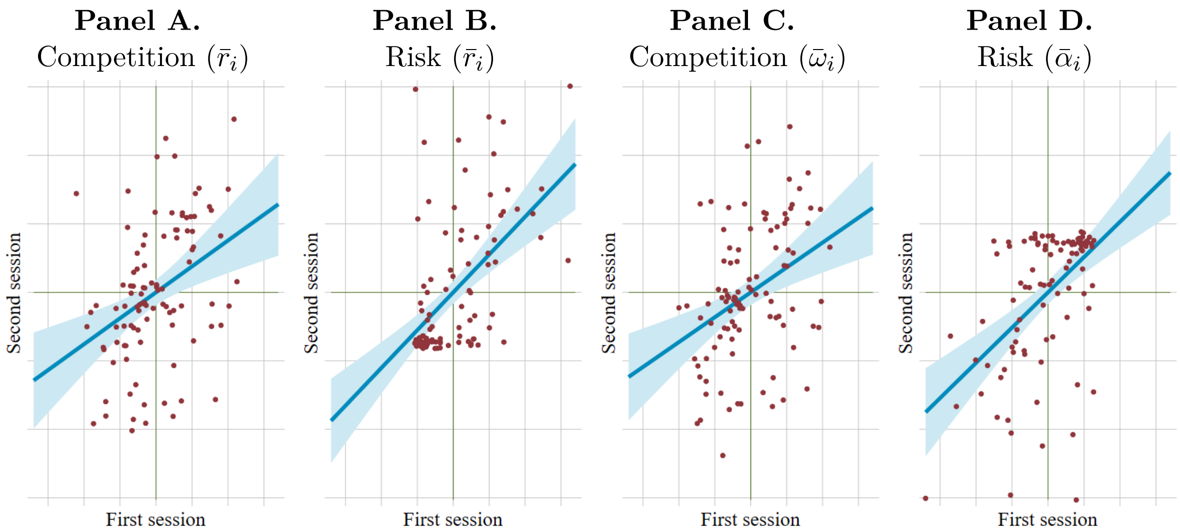


Figure 3. Correlation between the first and second sessions for measures of participants' preferences for competition and risk

Notes: Scatter plots of participants' preferences for competition and risk in the first and second sessions of the follow-up experiment. Panel A shows the median switching row of the competition elicitation sets. Panel B shows the median switching row of the risk elicitation sets. Panel C shows the values of $\bar{\omega}_i$. Panel D shows the values of $\bar{\alpha}_i$. All variables are standardized to have a mean of zero and a standard deviation of one. The best-fit lines correspond to Pearson's correlation coefficient. The grids have lengths of one standard deviation. Data corresponds to the 100 participants whose switching behavior is consistent with expected utility maximization in all decision sets of the first and second sessions of the follow-up experiment.

median switching row in the risk decision sets, 0.55 ($p < 0.001$). The difference between these two coefficients is not statistically significant (Wald test, $p = 0.245$). We obtain similar results for $\bar{\omega}_i$ and $\bar{\alpha}_i$, with correlation coefficients of 0.36 ($p < 0.001$) and 0.51 ($p < 0.001$), respectively, and no significant difference between them ($p = 0.277$).[25]

In sum, we find that the correlation coefficients for the measures of participants' preferences for competition are comparable to those for risk preferences and align with test-retest correlations of preferences reported in the literature. For example, Schildberg-Hörisch (2018) survey the literature on the stability of risk preferences and report correlation coefficients typically ranging between 0.27 and 0.57. Similarly, Fehr and Charness (2023) report that measures of social preferences exhibit test-retest correlation coefficients ranging from 0.42 to 0.56. Few studies examine preferences for competition over time, but Buser et al. (2024) report test-retest correlations of 0.58 and 0.40 for their unincentivized survey measures of competition and risk,

[25]The somewhat higher correlation of the risk measures is driven by decision sets without a switch. Table D6 in Appendix D shows that when the sample is restricted to participants who switch from competitive to individual pay in a majority of sets, the correlation for $\bar{\omega}_i$ increases to 0.48 while that of $\bar{\alpha}_i$ remains roughly the same at 0.53. If we further restrict the sample to participants who switch from competitive to individual pay in all decision sets, the correlations increase to 0.67 for $\bar{\omega}_i$ and 0.64 for $\bar{\alpha}_i$.

respectively.[26]

6.4. Competition-entry choice and preferences for competition

As a next step, we examine the relationship between our measures of preferences for competition and the traditional competition-entry choice used in the literature to study competitiveness (Niederle and Vesterlund, 2007). The competition-entry choice predicts important choices like high-school track choice (Buser et al., 2014, 2017b), participation in university entrance exams (Zhang, 2019), major choice (Reuben et al., 2017; Kamas and Preston, 2018; Buser et al., 2022), industry choice (Fallucchi et al., 2020; Reuben et al., 2024), and outcomes like educational attainment and income (Buser et al., 2024; Reuben et al., 2024). However, there is still an ongoing debate on the extent to which the competition-entry choice captures performance, beliefs, risk attitudes, or preferences for competition (Gillen et al., 2019; van Veldhuizen, 2022).

In the competition-entry choice, 44.9% of participants choose competitive pay in the first session, while 33.9% do so in the second session (McNemar's χ^2 test, $p = 0.041$). The decrease in the number of participants choosing to compete is likely due to receiving information about their probability of winning during the first session.

Table 6 explores the determinants of competition-entry through linear regressions of the competition-entry choice. Since we have up to two choices per participant, we cluster standard errors by participant and include a control for session timing in all regressions.[27] To have coefficients that are easy to interpret, we standardize the independent variables to have a mean of zero and a standard deviation of one, and we present coefficients in percent. As before, the analysis is limited to participants whose behavior is consistent with expected utility maximization in all five decision sets.

We start in column I with the participants' probability of being their group's winner as the independent variable. As expected, there is a strong positive relationship between the probability of winning and competition-entry. A one-standard-deviation increase in the probability of winning is associated with an 18.2 percentage point increase in competition entry. Next, we include our measures of participants' preferences for competition. In column II, we include participants' median switching row in the competition decision sets ($\bar{r}_i$), while in column III,

[26] Although group sizes change, we can also look at the correlations for these measures between groups of three and six in the initial experiment. We find a correlation coefficient of 0.34 for the median switching row and 0.38 for $\bar{\omega}_i$ ($p < 0.001$).

[27] Our results are unchanged if we use probit regressions or include participant random effects.

Table 6. Competition entry and preferences for competition

Notes: Linear regressions of the competition-entry choice (part I) in the follow-up experiment. The dependent variable equals one if participants choose competitive pay and zero otherwise. All regressions control for the session timing. Columns III and VI include sets of dummy variables for which we report the F-test for joint significance instead of individual coefficients. Standard errors clustered on participants are shown in parentheses. Data corresponds to participants whose behavior is consistent with expected utility maximization in all five decision sets. ** and * indicate statistical significance at 0.01 and 0.05.

	I	II	III	IV	V	VI	VII
Probability of winning	18.2**	23.0**	23.8**	21.5**	22.3**	$F_{9,113}$	$F_{9,113}$
	(3.2)	(3.1)	(3.0)	(3.2)	(3.2)	$= 5.7^{**}$	$= 6.9^{**}$
Competition $\bar{r}_i$		9.9**		8.7*		8.6*	
		(3.6)		(3.4)		(3.3)	
Competition $\bar{\omega}_i$			11.4**		10.0**		10.9**
			(3.7)		(3.5)		(3.5)
Risk $\bar{r}_i$				14.5**		$F_{5,113}$	
				(3.1)		$= 5.1^{**}$	
Risk $\bar{\alpha}_i$					-14.9^{**}		$F_{9,113}$
					(2.9)		$= 3.3^{**}$
Observations	214	214	214	214	214	214	214
Participants	144	144	144	144	144	144	144
R^2	0.15	0.18	0.19	0.27	0.28	0.31	0.32

we include their $\bar{\omega}_i$. Both measures show a strong positive association with competition-entry choices: a one-standard-deviation increase in preferences for competition increases the probability of choosing competitive pay by 9.9 percentage points in column II and 11.4 percentage points in column III. In columns IV and V, we further control for participants' risk preferences. In column IV, we include participants' median switching row in the risk decision sets, while in column V we include their median coefficient of CRRA, $\bar{\alpha}_i$. Higher levels of risk aversion are associated with a decreased probability of competition entry. However, controlling for risk preferences has little effect on the coefficients of the measures of preferences for competition. Finally, in columns VI and VII, we follow Gillen et al. (2019) and control for participants' probability of being their group winner and risk aversion in a more flexible way by dividing these variables into deciles and including them as dummy variables. Accordingly, for these variables, the regression table shows the F-test of joint significance instead of individual coefficients. The more flexible specifications slightly improve the regression fit but do not noticeably impact the significance or magnitude of the coefficients for preferences for competition.

Overall, we find that our measures of participants' preferences for competition are strong predictors of their competition-entry choices. We think this is reassuring as it validates our

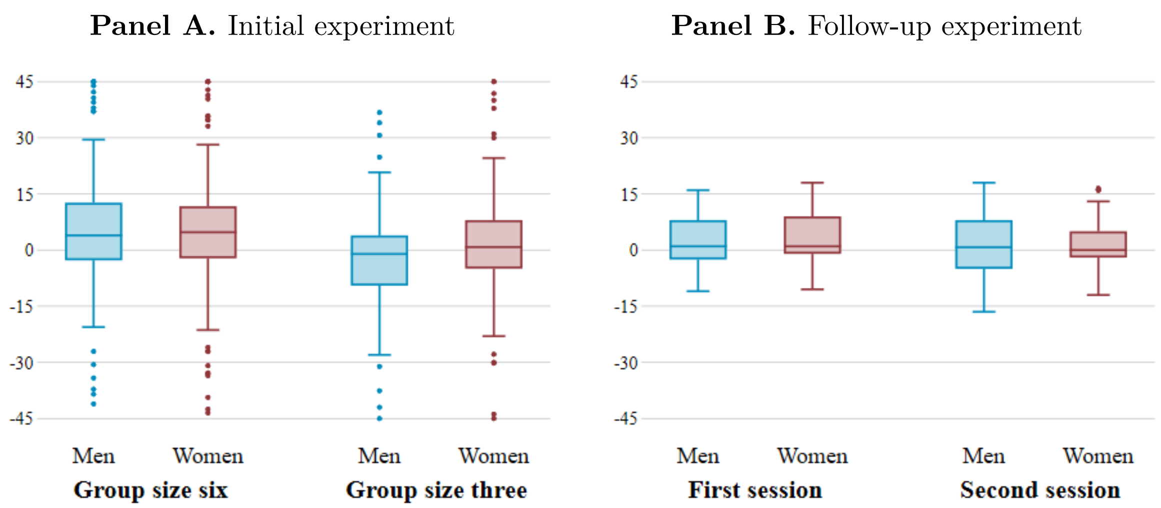


Figure 4. Preferences for competition by gender

Notes: Box plots of participants' $\bar{\omega}_i$ by gender for the initial and follow-up experiments. Each box plot shows the median, interquartile values, outliers as defined by Tukey (1977). Data corresponds to participants whose behavior is consistent with expected utility maximization in all five decision sets. In the initial experiment there are 82 men and 116 women in groups of six and 85 men and 119 women in groups of three. In the follow-up experiment there are 42 men and 63 women in the first session and 44 men and 65 women in the second session.

measures as capturing underlying preferences and demonstrates that these preferences influence competition-entry choices beyond performance, beliefs, and risk preferences.

7. Gender differences

While the main focus of this paper is on the measurement of preferences for competition, one of the reasons there has been considerable interest in this literature is the commonly-observed gender difference in competition entry (see Bertrand, 2011; Niederle, 2017). In this section, we report how our results are influenced by the participants' gender.

We do not find significant gender differences in the consistency of switching behavior with utility maximization either within or across decision sets in both the initial and follow-up experiments. Overall, around 98% of men's decision sets are classified as consistent, compared to 96% for women (see Tables C2 and D2; Mann-Whitney U tests, $p > 0.201$). Similarly, approximately 92% of men and 89% of women display behavior consistent with expected utility maximization in all five decision sets (see Tables C3 and D3; χ^2 tests, $p > 0.314$).[28]

Next, we turn to gender differences in preferences for competition. Figure 4 presents boxplots of participants' $\bar{\omega}_i$ by gender, visualizing the distributions' median, interquartile values, and

[28]There are also no gender differences in the consistency of switching behavior with expected utility maximization in the risk decision sets (see Appendix E.

outliers as defined by Tukey (1977). The figure shows no noticeable gender differences in either experiment. Hence, contrary to much of the literature, we do not find that women are more averse to competition than men. In fact, on average, women appear slightly more competition loving, although these differences are not statistically significant (see Tables C4 and D4).[29] Similarly, our analysis of persistence shows that men and women are equally likely to be classified as having a persistent preference for competition (either loving, averse, or neutral), with women being classified more often as competition loving than men, but again, these differences are not statistically significant (see TablesC5 and D5).[30] Finally, we do not find that the test-retest correlations of preferences for competition vary by gender (see Table D6; test of equality of correlation matrices tests $p > 0.764$).

Could the lack of gender differences be a consequence of this specific subject pool? We think this is unlikely since we replicate other gender differences observed in the literature. In particular, we replicate the gender gap in the competition-entry choice in the follow-up experiment: men choose to compete 53.3% of the time, while women do so 30.6% of the time (χ^2 test, $p = 0.003$). Given these results, we further examine the gender gap in competition entry and whether it can be explained by gender differences in other variables. Notably, in the same experiment, we find that women are significantly more risk-averse than men, with an average $\bar{\alpha}_i$ of 0.26 for women versus 0.17 for men (Mann-Whitney U test, $p < 0.001$). Women also tend to have a lower probability of being their group's winner, though this difference is not statistically significant (0.29 for women vs. 0.39 for men, Mann-Whitney U test, $p = 0.060$).[31]

Table 7 explores the determinants of the gender gap in competition entry. It presents the results of linear regressions of participants' competition-entry choices, using the same variables as in Table 6, plus gender.[32] Column I includes only the gender dummy as the independent

[29] Mann-Whitney U tests do not reject the null hypotheses of equal distributions for either $\bar{r}_i$ or $\bar{\omega}_i$ ($p > 0.339$ in the initial experiment and $p > 0.701$ in the follow-up experiment).

[30] χ^2 tests do not reject the null hypotheses of equal distributions of types in the follow-up experiment ($p > 0.842$ for the strict definition of competition-neutral behavior and $p > 0.427$ for the broader definition) and for groups of six in the initial experiment ($p = 0.901$ for the strict definition and $p = 0.136$ for the broader definition). For groups of three, the results are almost significant ($p = 0.069$ for the strict definition and $p = 0.056$ for the broader definition).

[31] We observe the same patterns in the initial experiment: women are significantly more risk-averse than men (see Table E4, Mann-Whitney U tests, $p = 0.049$) and have a lower probability of being their group's winner (see Table D1, Mann-Whitney U test, $p = 0.023$). Finally, while we cannot compare payment-scheme choices as they depend on participants' beliefs, which vary by gender, in the *Choice-first* treatment, participants complete one additional decision set where everyone faces the same ten choices (see section 3 and Appendix B). In this set, men choose competitive pay significantly more often than women (54.5% vs. 46.4%; Mann-Whitney U test, $p = 0.020$).

[32] As in Table 6, in all regressions, we cluster standard errors by participant, control for session timing, use

variable, indicating that women are 25.0% less likely to choose competitive pay. In column II, we add participants' probability of being their group's winner. Its inclusion reduces the gender coefficient by 5.2 percentage points to 19.8% (comparing the gender coefficients in columns I and II gives $p = 0.099$). Next, we include participants' preferences for competition, proxied by $\bar{r}_i$ in column III and $\bar{\omega}_i$ in column IV. Both measures are significantly associated with competition entry, but their inclusion has little effect on the gender coefficient, reducing it by 0.8 percentage points in column III ($p = 0.595$ when compared to column II) and and 0.6 in column IV ($p = 0.745$ when compared to column II). In columns V and VI, we control for participants' risk preferences instead of their preferences for competition. Using participants' median switching row in the risk decision sets (column V) significantly reduces the gender coefficient by 7.2 percentage points ($p = 0.016$ when compared to column II), while using participants' coefficient of CRRA (column VI) reduces it by 8.1 percentage points ($p = 0.012$ when compared to column II). Once we control for risk preferences, the gender coefficient is no longer statistically significant ($p = 0.098$ in column V and $p = 0.104$ in column VI). In columns VII and VIII, we include both participants' preferences for competition and risk, and in columns IX and X, we give more flexibility to these independent variables by dividing them into deciles and including them as dummy variables. In these specifications, the gender coefficient is further reduced, but only by 0.7 percentage points or less ($p > 0.752$ when compared to columns V and VI). By column X, the gender coefficient is less than half of that in column I and is not statistically significant ($p = 0.163$).

In sum, as in most of the literature on competition entry, we find that women choose to compete less frequently than men. Moreover, in line with Gillen et al. (2019) and van Veldhuizen (2022), we find that the gender difference in competition entry is no longer statistically significant once we control for risk preferences. Our experiment demonstrates that this is not due to a lack of preferences for competition but because men's preferences are not more competition loving than women's.[33] However, it is important to acknowledge that while the gender coefficient is not statistically significant once we control for risk preferences, its magnitude is

standardized independent variables, and conduct the analysis for participants whose behavior is consistent with expected utility maximization in all five decision sets.

[33]We find a similar result using data from the initial experiment. Specifically, we analyze the additional decision set in the *Choice-first* treatment where everyone faces the same ten choices. In this set, there is a significant gender gap of 8.0 percentage points, but it becomes statistically insignificant once we control for participants' beliefs. Preferences for competition predict choosing competitive pay but do not significantly reduce the gender coefficient. For details, see section C.6 in Appendix C.

Table 7. Competition entry and gender

Notes: Linear regressions of the competition-entry choice (part I) in the follow-up experiment. The dependent variable equals one if participants choose competitive pay and zero otherwise. All regressions control for the session timing. Columns IX and X include sets of dummy variables for which we report the F-test for joint significance instead of individual coefficients. Standard errors clustered on participants are shown in parentheses. Data corresponds to participants whose behavior is consistent with expected utility maximization in all five decision sets. ** and * indicate statistical significance at 0.01 and 0.05.

	I	II	III	IV	V	VI	VII	VIII	IX	X
Woman	-25.0^{**}	-19.8^{**}	-19.0^{*}	-19.2^{**}	-12.0	-11.7	-11.6	-11.7	-11.2	-11.0
	(7.9)	(7.5)	(7.3)	(7.2)	(7.2)	(7.1)	(7.2)	(7.0)	(7.0)	(7.8)
Probability of		16.7^{**}	21.3^{**}	22.2^{**}	16.5^{**}	16.6^{**}	20.7^{**}	21.5^{**}	$F_{9,113}$	$F_{9,113}$
winning		(3.2)	(3.1)	(3.1)	(3.2)	(3.2)	(3.3)	(3.3)	$= 4.7^{**}$	$= 5.6^{**}$
Competition $\bar{r}_i$			9.4^{**}				8.6^{**}		$F_{9,113}$	
			(3.3)				(3.2)		$= 2.5^{*}$	
Competition $\bar{\omega}_i$				11.1^{**}				10.0^{**}		$F_{9,113}$
				(3.4)				(3.4)		$= 2.2^{*}$
Risk $\bar{r}_i$					13.4^{**}		12.9^{**}		$F_{5,113}$	
					(3.1)		(3.1)		$= 3.9^{**}$	
Risk $\bar{\alpha}_i$						-14.0^{**}		-13.3^{**}		$F_{9,113}$
						(3.0)		(2.9)		$= 2.3^{*}$
Observations	214	214	214	214	214	214	214	214	214	144
Participants	144	144	144	144	144	144	144	144	144	144
R^2	0.08	0.19	0.22	0.23	0.26	0.27	0.28	0.30	0.35	0.35

sufficiently large that we cannot conclusively rule out the existence of gender differences due to other factors.

8. Conclusions

In this study, we advance our understanding of the factors that drive competition-entry decisions by measuring the extent to which individuals enjoy or dislike performing in competitive environments—in other words, their preferences for competition. Our findings provide compelling evidence that most participants are willing to forgo a significant portion of their expected earnings to either engage in or avoid competition. Additionally, our results demonstrate that competition-entry decisions are consistent with expected utility maximization and that most individuals behave as either persistently competition loving or persistently competition averse.

To measure preferences for competition, we adjust the experimental paradigm developed by Niederle and Vesterlund (2007) in three important ways. First, we introduce probabilistic incentives when participants do not compete to ensure that competition-entry choices are not influenced by risk preferences. Second, we record multiple competition-entry decisions from each participant using multiple price lists, which enables us to assess the consistency of participants' choices with expected utility maximization and whether they exhibit a persistent preference for or against competition. Third, we either measure participants' beliefs about their chances of winning using an incentive-compatible method (in the initial experiment) or inform them of their probability of winning based on their previous performance (in the follow-up experiment).

One intriguing finding of our study is that participants' preferences for competition vary with the intensity of competition, as determined by the number of competitors. On average, participants are more competition loving when competing in groups of six compared to groups of three. To our knowledge, few studies have explored the role of group size in competition-entry decisions. In the literature on contests, a few papers have looked at the impact of group size on effort provision (Dechenaux et al., 2015).[34] Although, both in theory and empirically, larger group sizes do not necessarily lead to higher effort levels (List et al., 2020), it is telling that the gap between observed and predicted effort increases with the number of competitors (e.g., see Sheremeta, 2011; Lim et al., 2014). Accounting for preferences for competition may explain this finding.

An important challenge in measuring preferences is the potential bias introduced by decision and measurement errors. While no experimental design can fully eliminate these biases, our analysis suggests that our findings are not driven by such errors or noisy measurements (Gillen et al., 2019). First, we observe a negligible fraction of violations of expected utility maximization in participants' switching behavior. Second, most participants exhibit behavior that is either persistently competition loving or persistently competition averse, a pattern unlikely to result from accidentally switching at the wrong point in the multiple price list. Third, we show that our results are robust to informing participants of their probability of winning, which should minimize the impact of errors in beliefs. Finally, we find that the test-retest correlations of our measures of preferences for competition are comparable to those observed for risk preferences, both for the same participants and in the broader literature on preference measurement.

In our follow-up experiment, we also examine the relationship between our measures of preferences for competition and the well-known competition-entry choice introduced by Niederle and Vesterlund (2007), which has been shown to predict many labor market and educational outcomes (Lozano et al., 2023). Reassuringly, we find that our direct measure of preferences for

[34] There are also studies investigating competition entry decisions where the number of competitors is unknown at the time of the decision (e.g., Morgan et al., 2012). It is unclear how our findings translate to settings where the number of competitors is endogenous.

competition is a strong predictor of choosing to compete, alongside individuals' performance and risk preferences. Having a direct measure of preferences for competition is valuable because it is not always possible to disentangle the impact of preferences for competition from other determinants of the competition-entry decision.

Although our experimental design offers several advantages, it does have some limitations. One limitation is that we ask participants to make a substantial number of choices based on relatively lengthy and detailed instructions. This makes our design well-suited for measuring preferences for competition in controlled environments but may not be ideal when participants are less attentive or time-constrained, as is often the case in many online samples and large surveys (for these samples see, Buser et al., 2024; Fallucchi et al., 2020).

A second limitation is that we did not elicit participants' beliefs about their absolute performance, which would be necessary to obtain an accurate monetary equivalent for our measures of preferences. We chose to omit this step to avoid burdening participants with an additional belief elicitation task. It is unclear whether participants systematically overestimate or underestimate their absolute performance. The literature on preferences for competition has primarily focused on beliefs about relative performance (see Table A1). However, when reported, participants tend to slightly overestimate the number of sums they will solve (Kamas and Preston, 2012; Wozniak et al., 2014; Banerjee et al., 2018; Saccardo et al., 2018). If this is true for our participants, we may be underestimating the monetary value of their preferences for competition.

A third limitation is that, while our design allows us to observe whether participants switch only once between competitive to individual pay or do not switch at all within decision sets, this is not a particularly strong test of consistency with expected utility maximization. A tougher test would involve evaluating whether participants' choices are consistent with the Generalized Axiom of Revealed Preference, as demonstrated by Choi et al. (2007) for risk preferences. This type of analysis would require more decision sets and greater within-participant variation of the trade-off between competitive and individual pay than our current design offers.

A fourth limitation is that our estimations do not consider probability weighting. While probability weighting does not affect the results concerning the consistency with utility maximization of participants' switching behavior, the analysis of their median switching row, or their classification as competition loving, adverse, or neutral, it can introduce a bias in the estimation of preferences for competition when measured as $\bar{\omega}_i$. In section E.5 of Appendix E, we estimate this potential bias using the probability weighting function proposed by Prelec (1998) and show that our main results remain unaffected. However, future research could benefit from examining how individually elicited probability weighting functions impact measures of preferences for competition.

Finally, it is important to note that our measures of preferences for competition are partial indicators of an individual's non-pecuniary utility of performing under competition. Participants' $\bar{\omega}_i$ provides a useful interpretation as the fraction of the utility gain from winning they are willing to forgo to perform under competitive pay. However, to fully understand how participants trade off their preferences for competition and monetary rewards, we would also need to precisely elicit individual utility functions, which is challenging without making strong assumptions about functional forms. Additionally, the interpretation of our measures of preferences for competition assumes that these preferences are additive in the utility function, which may not necessarily be the case.

In the last part of our analysis, we look at gender differences. We find that men and women are equally consistent with utility maximization in their switching behavior within and across sets, are willing to forgo similar fractions of the utility gain from winning to either engage in or avoid competition, and are similarly likely to be classified as having a persistent preference for competition (whether loving, averse, or neutral).

Additionally, while we do not find gender differences in preferences for competition, we do replicate the common finding that women choose to enter competitions less often than men. This suggests that the gender gap in competition entry is driven by other factors, such as differences in risk preferences (as seen in the follow-up experiment) or beliefs about winning the competition (as seen in the initial experiment). These findings align with Gillen et al. (2019) and van Veldhuizen (2022), who report that gender differences in competition entry disappear after accounting for risk preferences and beliefs. Importantly, unlike these studies, we show that the lack of a gender difference is not due to the absence of preferences for competition. In fact, our findings suggest that preferences for competition are widespread and distinct from risk preferences, as suggested by empirical work showing that competition and risk-taking behavior in the lab predicts different behaviors in the field (e.g., Buser et al., 2014; Reuben et al., 2017, 2024; Buser et al., 2024).

In sum, we find that preferences for competition are prevalent among our participants and play an important role in determining their competitive behavior. Given the ubiquity of competition in various aspects of our lives, understanding individuals' motivations to compete is essential. For instance, recognizing that non-pecuniary factors unrelated to performance can influence decisions to compete may affect the predicted effectiveness of competitions in identifying top performers and have implications for the welfare impact of tournaments. This paper advances this research by providing empirical evidence of a determinant of preferences for competition—the number of competitors—and an experimental design that directly measures them to facilitate further study.

References

Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., and Tungodden, B. (2016). Willingness to compete: Family matters. *Management Science*, 62(8):2149–2162.

Andersen, S., Ertac, S., Gneezy, U., List, J. A., and Maximiano, S. (2013). Gender, Competitiveness, and Socialization at a Young Age: Evidence from a Matrilineal and a Patriarchal Society. *Review of Economics and Statistics*, 95(4):1438–1443.

Apicella, C. L., Demiral, E. E., and Mollerstrom, J. (2017). No Gender Difference in Willingness to Compete When Competing against Self. *American Economic Review*, 107(5):136–140.

Apicella, C. L. and Dreber, A. (2015). Sex Differences in Competitiveness: Hunter-Gatherer Women and Girls Compete Less in Gender-Neutral and Male-Centric Tasks. *Adaptive Human Behavior and Physiology*, 1(3):247–269.

Balafoutas, L., Kerschbamer, R., and Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior & Organization*, 83(1):125–135.

Balafoutas, L. and Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science*, 335(6068):579–82.

Banerjee, R., Gupta, N. D., and Villeval, M. C. (2018). The spillover effects of affirmative action on competitiveness and unethical behavior. *European Economic Review*, 101:567–604.

Berge, L. I. O., Bjorvatn, K., Garcia Pires, A. J., and Tungodden, B. (2015). Competitive in the lab, successful in the field? *Journal of Economic Behavior & Organization*, 118:303–317.

Berlin, N. and Dargnies, M.-P. (2016). Gender differences in reactions to feedback and willingness to compete. *Journal of Economic Behavior & Organization*, 130:320–336.

Bertrand, M. (2011). New Perspectives on Gender. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier B.V.

Bönte, W., Lombardo, S., and Urbig, D. (2017). Economics meets psychology: Experimental and self-reported measures of individual competitiveness. *Personality and Individual Differences*, 116:179–185.

Brandts, J., Groenert, V., and Rott, C. (2015). The Impact of Advice on Women's and Men's Selection into Competition. *Management Science*, 61(5):1018–1035.

Buser, T., Cappelen, A., Gneezy, U., Hoffman, M., and Tungodden, B. (2021a). Competitiveness, gender and handedness. *Economics & Human Biology*, 43:101037.

Buser, T., Dreber, A., and Mollerstrom, J. (2017a). The impact of stress on tournament entry. *Experimental Economics*, 20(2):506–530.

Buser, T., Gerhards, L., and van der Weele, J. (2018). Responsiveness to feedback as a personal trait. *Journal of Risk and Uncertainty*, 56(2):165–192.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, Competitiveness, and Career Choices. *The Quarterly Journal of Economics*, 129(3):1409–1447.

Buser, T., Niederle, M., and Oosterbeek, H. (2024). Can competitiveness predict education and labor market outcomes? evidence from incentivized choice and survey measures. *Review of Economics and Statistics*, pages 1–45.

Buser, T., Peter, N., and Wolter, S. C. (2017b). Gender, Competitiveness, and Study Choices in High School: Evidence from Switzerland. *American Economic Review*, 107(5):125–130.

Buser, T., Peter, N., and Wolter, S. C. (2022). Willingness to compete, gender and career choices along the whole ability distribution. *Experimental Economics*, 25:1299–1326.

Buser, T., Ranehill, E., and van Veldhuizen, R. (2021b). Gender differences in willingness to compete: The role of public observability. *Journal of Economic Psychology*, 83:102366.

Cadsby, C. B., Servátka, M., and Song, F. (2013). How competitive are female professionals? A tale of identity conflict. *Journal of Economic Behavior & Organization*, 92:284–303.

Cárdenas, J.-C., Dreber, A., von Essen, E., and Ranehill, E. (2012). Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden. *Journal of Economic Behavior & Organization*, 83(1):11–23.

Cassar, A., Wordofa, F., and Zhang, Y. J. (2016). Competing for the benefit of offspring eliminates the gender gap in competitiveness. *Proceedings of the National Academy of Sciences*, 113(19):5201–5205.

Chapman, J., Snowberg, E., Wang, S., and Camerer, C. (2019). Loss Attitudes in the U.S. Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE).

Choi, S., Fisman, R., Gale, D. M., and Kariv, S. (2007). Consistency and Heterogeneity of Individual Behavior under Uncertainty. *American Economic Review*, 97(5):1921–1938.

Crosetto, P. and Filippin, A. (2013). The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, 47(1):31–65.

Danz, D., Vesterlund, L., and Wilson, A. J. (2022). Belief Elicitation and Behavioral Incentive Compatibility. *American Economic Review*, 112(9):2851–2883.

Dargnies, M.-P. (2012). Men Too Sometimes Shy Away from Competition: The Case of Team Competition. *Management Science*, 58(11):1982–2000.

Dariel, A., Kephart, C., Nikiforakis, N., and Zenker, C. (2017). Emirati women do not shy away from competition: Evidence from a patriarchal society in transition. *Journal of the Economic Science Association*, 3(2):121–136.

Dariel, A. and Nikiforakis, N. (2022). Is there a motherhood gap in the willingness to compete for pay? Working paper, NYU Abu DHabi.

Datta Gupta, N., Poulsen, A., and Villeval, M.-C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, 51(1):816–835.

Dechenaux, E., Kovenock, D., and Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4):609–669.

Dohmen, T. and Falk, A. (2011). Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender. *American Economic Review*, 101(2):556–590.

Dreber, A., von Essen, E., and Ranehill, E. (2014). Gender and competition in adolescence: Task matters. *Experimental Economics*, 17(1):154–172.

Eckel, C. C. and Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, 23:281–295.

Fallucchi, F., Nosenzo, D., and Reuben, E. (2020). Measuring preferences for competition with experimentally-validated survey questions. *Journal of Economic Behavior & Organization*, 178:402–423.

Fehr, E. and Charness, G. (2023). Social preferences: Fundamental characteristics and economic consequences. Working paper no. 10488, CESifo.

Fischbacher, U. (2007). Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Gächter, S. and Renner, E. (2010). The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, 13(3):364–377.

Geraldes, D. (2020). Women Dislike Competing Against Men. Working Paper 3741649, SSRN.

Gillen, B., Snowberg, E., and Yariv, L. (2019). Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study. *Journal of Political Economy*, 127(4):1826–1863.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society. *Econometrica*, 77(5):1637–1664.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Halko, M.-L. and Sääksvuori, L. (2017). Competitive behavior, stress, and gender. *Journal of Economic Behavior & Organization*, 141:96–109.

Harrison, G. W. and Phillips, R. D. (2014). Subjective Beliefs and Statistical Forecasts of Financial Risks: The Chief Risk Officer Project. In Andersen, T. J., editor, *Contemporary Challenges in Risk Management*, pages 163–202. Palgrave Macmillan UK, London.

Healy, A. and Pate, J. (2011). Can Teams Help to Close the Gender Competition Gap? *The Economic Journal*, 121(555):1192–1204.

Heckman, J., Jagelka, T., and Kautz, T. (2019). Some Contributions of Economics to the Study of Personality. Technical Report w26459, National Bureau of Economic Research, Cambridge, MA.

Holt, C. A. and Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, 92(5):1644–1655.

Hossain, T. and Okui, R. (2013). The Binarized Scoring Rule. *The Review of Economic Studies*, 80(3):984–1001.

Ifcher, J. and Zarghamee, H. (2016). Pricing competition: A new laboratory measure of gender differences in the willingness to compete. *Experimental Economics*, 19(3):642–662.

Jung, S. and Vranceanu, R. (2019). Willingness to compete: Between- and within-gender comparisons. *Managerial and Decision Economics*, page mde.3004.

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

Kamas, L. and Preston, A. (2018). Competing with confidence: The ticket to labor market success for college-educated women. *Journal of Economic Behavior & Organization*, 155:231–252.

Karni, E. (2009). A Mechanism for Eliciting Probabilities. *Econometrica*, 77(2):603–606.

Khachatryan, K., Dreber, A., von Essen, E., and Ranehill, E. (2015). Gender and preferences at a young age: Evidence from Armenia. *Journal of Economic Behavior & Organization*, 118:318–332.

Lee, S., Niederle, M., and Kang, N. (2014). Do single-sex schools make girls more competitive? *Economics Letters*, 124(3):474–477.

Lim, W., Matros, A., and Turocy, T. L. (2014). Bounded rationality and group size in Tullock contests: Experimental evidence. *Journal of Economic Behavior & Organization*, 99:155–167.

List, J. A., Van Soest, D., Stoop, J., and Zhou, H. (2020). On the role of group size in tournaments: Theory and evidence from laboratory and field experiments. *Management Science*, 66(10):4359–4377.

Lozano, L., Ranehill, E., and Reuben, E. (2023). Gender and Preferences in the Labor Market: Insights from Experiments. In Zimmermann, K. F., editor, *Handbook of Labor, Human Resources and Population Economics*. Springer.

Markowsky, E. and Beblo, M. (2022). When do we observe a gender gap in competition entry? a meta-analysis of the experimental literature. *Journal of Economic Behavior & Organization*, 198:139–163.

Mayr, U., Wozniak, D., Davidson, C., Kuhns, D., and Harbaugh, W. T. (2012). Competitiveness across the life span: The feisty fifties. *Psychology and Aging*, 27(2):278–285.

Molnar, A. and Paolacci, G. (2024). Competition increases the magnitude of dishonest reporting even when controlling for reward uncertainty. *Scientific Reports*, 14(1):31980.

Morgan, J., Orzen, H., and Sefton, M. (2012). Endogenous entry in contests. *Economic Theory*, 51:435–463.

Muller, J. and Schwieren, C. (2012). Can Personality Explain what is Underlying Women's Unwillingness to Compete? *Journal of Economic Psychology*, 33:448–460.

Niederle, M. (2017). A Gender Agenda: A Progress Report on Competitiveness. *American Economic Review*, 107(5):115–119.

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

Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Niederle, M. and Vesterlund, L. (2011). Gender and Competition. *Annual Review of Economics*, 3(1):601–630.

Petrie, R. and Segal, C. (2017). Gender differences in competitiveness: The role of prizes. Working paper, Texas A&M University.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3):497.

Price, C. R. (2012). Gender, Competition, and Managerial Decisions. *Management Science*, 58(1):114–122.

Reuben, E., Sapienza, P., and Zingales, L. (2024). Overconfidence and preferences for competition. *Journal of Finance*, 79(2):1087–1121.

Reuben, E., Wiswall, M., and Zafar, B. (2017). Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender. *The Economic Journal*, 127(604):2153–2186.

Saccardo, S., Pietrasz, A., and Gneezy, U. (2018). On the Size of the Gender Difference in Competitiveness. *Management Science*, 64(4):1541–1554.

Samak, A. C. (2013). Is there a gender gap in preschoolers' competitiveness? An experiment in the U.S. *Journal of Economic Behavior & Organization*, 92:22–31.

Schildberg-Hörisch, H. (2018). Are risk preferences stable? *Journal of Economic Perspectives*, 32(2):135–154.

Sheremeta, R. M. (2011). Contest design: An experimental investigation. *Economic Inquiry*, 49(2):573–590.

Shurchkov, O. (2012). Under Pressure: Gender Differences in Output Quality and Quantity Under Competition and Time Constraints. *Journal of the European Economic Association*, 10(5):1189–1213.

Sutter, M. and Glätzle-Rützler, D. (2015). Gender Differences in the Willingness to Compete Emerge Early in Life and Persist. *Management Science*, 61(10):2339–2354.

Sutter, M., Glätzle-Rützler, D., Balafoutas, L., and Czermak, S. (2016). Cancelling out early age gender differences in competition: An analysis of policy interventions. *Experimental Economics*, 19(2):412–432.

Toubia, O., Johnson, E., Evgeniou, T., and Delquié, P. (2013). Dynamic Experiments for Estimating Preferences: An Adaptive Method of Eliciting Time and Risk Parameters. *Management Science*, 59(3):613–640.

Trautmann, S. T. and van de Kuilen, G. (2015). Belief Elicitation: A Horse Race among Truth Serums. *The Economic Journal*, 125(589):2116–2135.

Tukey, J. W. (1977). *Exploratory data analysis*. Addison–Wesley, Reading, MA.

Tversky, A. and Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, 102(2):269–283.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

van de Kuilen, G. and Wakker, P. P. (2011). The Midweight Method to Measure Attitudes Toward Risk and Ambiguity. *Management Science*, 57(3):582–598.

van Veldhuizen, R. (2022). Gender Differences in Tournament Choices: Risk Preferences, Overconfidence or Competitiveness? *Journal of the European Economic Association*, page jvac031.

Wang, S. W. (2011). Incentive effects: The case of belief elicitation from individuals in groups. *Economics Letters*, 111(1):30–33.

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

Wu, G. and Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, 42(12):1676–1690.

Zhang, Y. J. (2019). Culture, Institutions and the Gender Gap in Competitive Inclination: Evidence from the Communist Experiment in China. *The Economic Journal*, 129(617):509–552.

Zhong, S., Shalev, I., Koh, D., Ebstein, R. P., and Chew, S. H. (2018). Competitiveness and stress. *International Economic Review*, 59(3):1263–1281.

Online appendices for: (Re)Measuring preferences for competition

Appendix A summarizes the methods used to elicit preferences for competition, beliefs, and risk preferences in various papers. Appendix B described in detail the procedure used to construct the multiple price lists used to measure preferences for competition. The auxiliary data analysis for the initial experiment is presented in Appendix C and for the follow-up experiment in Appendix D. Appendix E contains a detailed data analysis of risk preferences, while Appendix C.6 presents the auxiliary analysis of gender differences. Finally, a sample of the instructions from the initial experiment is available in Appendix F and from the follow-up experiment in Appendix G.

Appendix A. Papers measuring preferences for competition

Table A1 lists papers studying preferences for competition using slight variations of the Niederle and Vesterlund (2007) experimental design. This list is based on the papers reviewed by Dariel et al. (2017) plus a few more papers that share design choices with our experiment. The first two columns of the table indicate how the decision to compete is implemented: either as a single binary choice between individual and competitive pay or a more continuous measure of this choice using a multiple price list or a slider. The next six columns correspond to the belief elicitation task. Columns three through six describe the belief elicitation task. Most papers elicit the participants' expected rank within their group or the probability of being the group's winner. Other methods include eliciting the expected number of correct sums, self-reported assessments of their relative performance, and the expected number of sums performed by men and women. The seventh column indicates whether beliefs are elicited before or after the payment-scheme choice. The eighth column indicates whether the beliefs are incentivized, and the ninth column further specifies if the incentivization is done using the binarized scoring rule. Finally, the last three columns indicate whether risk preferences are measured with an unincentivized survey question (tenth column), incentivized lotteries in a multiple price list (eleventh column), or another incentivized method (twelfth column).

Table A1. List of studies using variations of the Niederle and Vesterlund (2007) experimental design

	Payment-scheme choice		Belief elicitation task						Risk elicitation task		
			Statistic			Timing	Incentivized				
Study	*Binary*	*MPL or slider*	*Rank*	*Prob. win*	*Other*	*Before / After*	*Yes*	*BSR*	*Survey*	*MPL*	*Other*
Niederle and Vesterlund (2007)	✓	×	✓	×	×	after	✓	×	×	×	✓
Gneezy et al. (2009)	✓	×	×	×	×	×	×	×	×	×	✓
Healy and Pate (2011)	✓	×	✓	×	×	after	✓	×	×	×	✓
Balafoutas and Sutter (2012)	✓	×	✓	×	×	after	✓	×	×	×	×
Balafoutas et al. (2012)	✓	×	✓	×	×	after	✓	×	×	✓	×
Cárdenas et al. (2012)	✓	×	×	×	✓	after	×	×	×	✓	×
Dargnies (2012)	✓	×	×	×	✓	after	×	×	×	×	✓
Kamas and Preston (2012)	✓	×	✓	×	✓	after	✓	×	✓	✓	×
Mayr et al. (2012)	✓	×	✓	×	×	before	✓	×	×	×	×
Muller and Schwieren (2012)	✓	×	✓	×	×	after	×	×	×	✓	×
Price (2012)	✓	×	✓	×	×	after	✓	×	×	×	×
Shurchkov (2012)	✓	×	✓	×	✓	after	×	×	×	×	×
Andersen et al. (2013)	✓	×	×	×	×	×	×	×	×	×	×
Cadsby et al. (2013)	✓	×	✓	×	×	after	✓	×	×	×	✓
Datta Gupta et al. (2013)	✓	×	×	×	✓	after	✓	×	✓	×	×
Niederle et al. (2013)	✓	×	✓	×	×	after	✓	×	×	×	✓
Samak (2013)	✓	×	×	×	✓	after	×	×	×	×	×
Buser et al. (2014)	✓	×	×	×	✓	after	×	×	✓	×	×
Dreber et al. (2014)	✓	×	✓	×	×	after	×	×	×	✓	×
Lee et al. (2014)	✓	×	×	×	✓	after	×	×	×	×	✓
Wozniak et al. (2014)	✓	×	×	×	✓	after	✓	×	×	✓	×
Apicella and Dreber (2015)	✓	×	×	×	×	×	×	×	×	×	×
Brandts et al. (2015)	✓	×	×	×	✓	after	✓	×	×	×	✓
Khachatryan et al. (2015)	✓	×	✓	×	×	after	×	×	×	✓	×
Sutter and Glätzle-Rützler (2015)	✓	×	✓	×	×	after	×	×	×	×	×
Almås et al. (2016)	✓	×	✓	×	×	before	×	×	✓	✓	×

Continuation Table A1. List of studies using variations of the Niederle and Vesterlund (2007) experimental design

	Payment-scheme choice		**Belief elicitation task**						**Risk elicitation task**		
			Statistic			**Timing**	**Incentivized**				
Study	*Binary*	*MPL or slider*	*Rank*	*Prob. win*	*Other*	*Before / After*	*Yes*	*BSR*	*Survey*	*MPL*	*Other*
Berlin and Dargnies (2016)	✓	×	×	✓	×	before	✓	✓	×	×	✓
Cassar et al. (2016)	✓	×	×	×	✓	after	×	×	×	✓	×
Ifcher and Zarghamee (2016)	×	✓	✓	×	×	after	✓	×	×	✓	✓
Sutter et al. (2016)	✓	×	✓	×	×	after	✓	×	×	×	×
Apicella et al. (2017)	✓	×	×	×	✓	after	✓	×	✓	×	×
Buser et al. (2017b)	✓	×	×	×	×	×	×	×	×	✓	×
Buser et al. (2017a)	✓	×	×	✓	×	after	×	×	✓	×	×
Bönte et al. (2017)	✓	×	✓	×	✓	after	×	×	✓	×	×
Dariel et al. (2017)	✓	×	×	×	✓	after	×	×	✓	×	×
Halko and Sääksvuori (2017)	✓	×	✓	×	×	after	✓	×	✓	✓	×
Petrie and Segal (2017)	×	✓	×	✓	×	after	✓	✓	×	×	✓
Reuben et al. (2017)	✓	×	✓	✓	×	after	✓	✓	×	✓	×
Banerjee et al. (2018)	✓	×	✓	✓	✓	after	✓	✓	×	✓	×
Buser et al. (2018)	✓	×	✓	×	×	after	✓	×	✓	×	×
Saccardo et al. (2018)	×	✓	×	✓	✓	after	×	×	✓	✓	×
Zhong et al. (2018)	✓	×	✓	×	×	after	✓	×	×	✓	×
Jung and Vranceanu (2019)	×	✓	✓	×	×	after	✓	×	✓	×	✓
Reuben et al. (2024)	✓	×	✓	×	×	after	✓	×	×	✓	×
Zhang (2019)	✓	×	✓	×	×	after	×	×	×	×	✓
Fallucchi et al. (2020)	×	✓	×	✓	×	after	✓	✓	✓	✓	×
Geraldes (2020)	✓	×	✓	×	×	after	✓	✓	✓	×	×
Buser et al. (2021b)	✓	×	✓	×	×	after	✓	×	×	×	✓
Buser et al. (2021a)	✓	×	×	×	✓	before	×	×	✓	×	×
van Veldhuizen (2022)	✓	×	×	✓	×	before	✓	✓	✓	✓	×
Danz et al. (2022)	✓	×	×	✓	×	after	✓	✓	×	×	✓
Molnar and Paolacci (2024)	✓	×	✓	✓	×	after	×	×	×	×	×

Appendix B. Details of the experimental design

B.1. High and low payments in the multiple price lists

In Table B1, we list the high π^H and low π^L amounts employed in the five decision sets used to measure the participants' preferences for competition. Each decision set is a multiple price list (MPL) with ten choices between individual and competitive pay. The order in which decision sets appear to participants is randomized.

Table B1. High and low values of the five decision sets (in euros)

	Decision set				
	1	2	3	4	5
Group size three					
π^H	4.00	6.00	1.50	4.00	2.00
π^L	0.00	0.00	0.00	1.00	0.50
Group size six					
π^H	6.00	9.00	3.00	4.00	3.50
π^L	0.00	0.00	0.00	1.00	0.50

B.2. Probabilities used in the multiple price lists

We use the procedure described below to have MPLs with a reasonable degree of accuracy (i.e., the steps between items are not too large) and ensure that the participants' belief of being their group's winner is contained within the MPL.

Each multiple price list contains ten rows $r \in [1, 10]$. Each row has a probability of winning in individual pay $p_r \in [0, 1]$. As one goes down the list, the probability of winning in individual pay increases by z percentage points (i.e., $p_{r+1} = p_r + z$). Hence, in a given MPL, the probabilities range from $p_1 = p_L$ to $p_{10} = p_L + 9z$. For a participant i, we use a reference probability $b_i \in [0, 1]$ to set the starting probability, p_L, for i's MPLs in the following way:

- If $b_i \leq 9z$ then $p_L = 0$.
- If $b_i \geq 1 - 9z$ then $p_L = 1 - 9z$.
- If $9z < b_i < 1 - 9z$ then $p_L = b_i - 5z + \epsilon$, where ϵ is a random number drawn from a uniform distribution with support $[-0.025, 0.025]$.

The values of z varied across the various decision sets as follows: $z = 0.02$ in decision set 1, $z = 0.01$ in decision set 2, $z = 0.03$ in decision set 3, $z = 0.02$ ($z = 0.03$) in decision set 4 for group sizes of three (six), and $z = 0.04$ in decision set 5. We varied the values of z and introduced the random component ϵ so that participants would not see the same probability range in every decision set.

In the initial experiment, the reference probability b_i depends on when beliefs are elicited. In the *Belief-first* treatment, b_i equals the participants' elicited belief of being their group's winner. In the *Choice-first* treatment, b_i is obtained with two extra MPLs designed to narrow the probability range in which participants switch from competitive to individual pay. In the first extra MPL, $p_L = 0.05$ and $z = 0.10$. The second extra MPL is based on the answers to the first extra MPL. Participants who switch from competitive pay to individual pay at $p_r \leq 0.35$ get $p_L = 0.05$, those who switch at $0.35 < p_r < 0.65$ get $p_L = 0.30$, and those who switch at $p_r \geq 0.65$ get $p_L = 0.50$. In all cases, $z = 0.05$. We set b_i as the probability at which the participant switches from competitive to individual pay in the second extra MPL.[35]

In the follow-up experiment, b_i equals the probability that a participant is the top performer in a randomly formed group within their session. These probabilities are calculated based on participants' previous performance in the addition task.

Appendix C. Auxiliary analysis of the initial experiment

C.1. Descriptive statistics of the addition task

For our initial experiment, Table C1 shows the means and standard deviations for the number of correct sums, the participants' belief of being their group's winner, and overconfidence calculated as the difference between their belief and the probability of being the winner of a randomly-formed group. Descriptive statistics are shown for the pooled data as well as depending on group size, the timing of belief elicitation, the order in which the tasks were played, and gender.

C.2. Switching behavior in various subsamples

In this section, we present the results for consistency of switching behavior with expected utility maximization displayed in Tables2 and 3 when we divide the sample into different subsamples. Specifically, we divide our sample depending on (i) the order in which participants played the experimental tasks, (ii) the timing of the belief elicitation task, and (iii) gender.

Table C2 shows whether behavior within sets is consistent with utility maximization in the

[35] For participants who switch multiple times or switch from individual to competitive pay, we took the number of competitive pay choices multiplied by z plus p_L as the switching probability.

Table C1. Descriptive statistics for performance and beliefs in the addition task

Notes: Mean and standard deviations in parenthesis for the number of correct sums in the addition task, participants' belief of being their group's winner (in percent), and overconfidence (participants' belief of winning minus their probability of winning, in percent). Statistics are presented by group size, the timing of the belief elicitation task, whether the task is played first or second, and gender.

		Correct sums		**Belief of winning**		**Over-confidence**	
Pooled data		11.4	(4.7)	44.9	(26.1)	19.9	(27.1)
By group size	Group size three	11.4	(4.7)	53.8	(22.3)	20.4	(25.7)
	Group size six	11.3	(4.7)	36.0	(26.7)	19.4	(28.4)
By timing of beliefs	Belief first	10.9	(4.6)	42.6	(26.2)	19.7	(27.3)
	Choice first	11.7	(4.7)	47.0	(26.0)	20.1	(26.9)
By order of play	Played first	10.8	(4.4)	42.0	(25.6)	20.2	(26.9)
	Played second	11.9	(4.9)	47.8	(26.4)	19.6	(27.3)
By gender	Men	12.2	(5.2)	48.3	(27.5)	17.7	(27.9)
	Women	10.8	(4.2)	42.6	(25.0)	21.4	(26.4)

absence of errors in the various subsamples. Decision sets are classified as *inconsistent* if they contain multiple switches or a unique non-monotonic switch and as *consistent* if they contain a single switch from competitive to individual pay or no switch. The table shows that in all subsamples at least 94% of all decision sets are classified as consistent with expected utility maximization. By and large, switching behavior across subsamples is similar, with one noticeable difference: participants in the *Belief-first* treatment are more likely to have sets with no switching and less likely to have sets with a single switch than those in *Choice-first*.

Table C3 shows participants' switching behavior across the five decision sets for the different subsamples. The table shows the fraction of participants with (I) at least one inconsistent decision set, (II) five consistent decision sets, (III) five consistent decision sets and a single switch in a majority of decision sets, and (IV) five consistent decision sets and a single switch in all five decision sets. Overall, over 86% of participants have five consistent decision sets in all subsamples. The only noticeable difference is that *Belief-first* shows a higher fraction of participants with five consistent decision sets and a switch in a majority of sets than *Choice-first*.

C.3. Preferences for competition in various subsamples

In this section, we present further analysis of our measures of participants' preferences for competition.

In Figure C1, we show the distribution of the participants' mean switching row across the

Table C2. Consistency of switching behavior with utility maximization within decision sets in various subsamples

Notes: Fraction of decision sets classified as consistent and inconsistent with expected utility maximization. Inconsistent decision sets contain *multiple switches* or a unique *non-monotonic switch* from individual to competitive pay. Consistent decision sets contain a *single switch* from competitive to individual pay or *no switch*.

	Inconsistent Behavior				**Consistent Behavior**			
	Multiple switches		*Non-monotonic switch*		*Single switch*		*No switch*	
Group size	Three	Six	Three	Six	Three	Six	Three	Six
Pooled data	3.1%	3.9%	1.0%	0.7%	77.0%	76.0%	18.9%	19.4%
By order of play								
Played first	4.7%	3.9%	1.6%	0.9%	76.9%	78.1%	16.7%	17.2%
Played second	1.6%	4.0%	0.4%	0.5%	77.0%	73.8%	21.1%	21.6%
By timing of beliefs								
Belief first	5.2%	5.0%	0.9%	0.6%	67.9%	68.2%	26.0%	26.2%
Choice first	1.2%	2.9%	1.0%	0.9%	85.3%	83.1%	12.5%	13.2%
By gender								
Men	2.4%	2.6%	0.2%	1.3%	74.7%	75.2%	22.6%	20.9%
Women	3.6%	4.8%	1.5%	0.3%	78.5%	76.5%	16.4%	18.3%

five decision sets (Panel A) and their preferences for competition when we measure them as the mean value of the five ω_{it}s instead of the median (Panel B). As with the median measures, we can reject the null hypothesis that the distribution is centered around zero for groups of six but not for groups of three (Wilcoxon signed-rank tests, $p < 0.001$ and $p > 0.227$, respectively). Moreover, we can reject the null hypothesis that the observed values in groups of three and six come from the same distribution (Wilcoxon signed-rank tests, $p < 0.001$).

Table C4 shows the mean and standard deviation of the participants' median switching row ($\bar{r}_i$) and preferences for competition ($\bar{\omega}_i$) for various subsamples. In addition, it also shows the results when we use other criteria to include participants in the analysis. First, we include all 224 participants. In sets with multiple switches, the value of the switching row and ω_{it} are calculated based on the first switch from competitive to individual pay. In sets with a single non-monotonic switch, we set the value of the switching row and ω_{it}s as the largest value in the set. Second, we include participants with five consistent decision sets and less than two sets without a switch, which guarantees that the median set is a set with a switch from competitive to individual pay (170 participants in groups of three and 168 in groups of six). Lastly, we include only participants with five consistent decision sets and no sets without a switch (122 participants in groups of three and 102 in groups of six).

Table C3. Consistency of switching behavior with utility maximization across decision sets in various subsamples

Notes: Fraction of participants according to the number of consistent and inconsistent decision sets. Inconsistent decision sets contain multiple switches or a unique switch from individual to competitive pay. Consistent decision sets contain a single switch from competitive to individual pay or no switch. Fraction of participants with at least one inconsistent decision set in column I, five consistent decision sets in column II, five consistent decision sets and a single switch from competitive to individual pay in a majority of decision sets in column III, and five consistent decision sets and a single switch from competitive to individual pay in all decision sets in column IV.

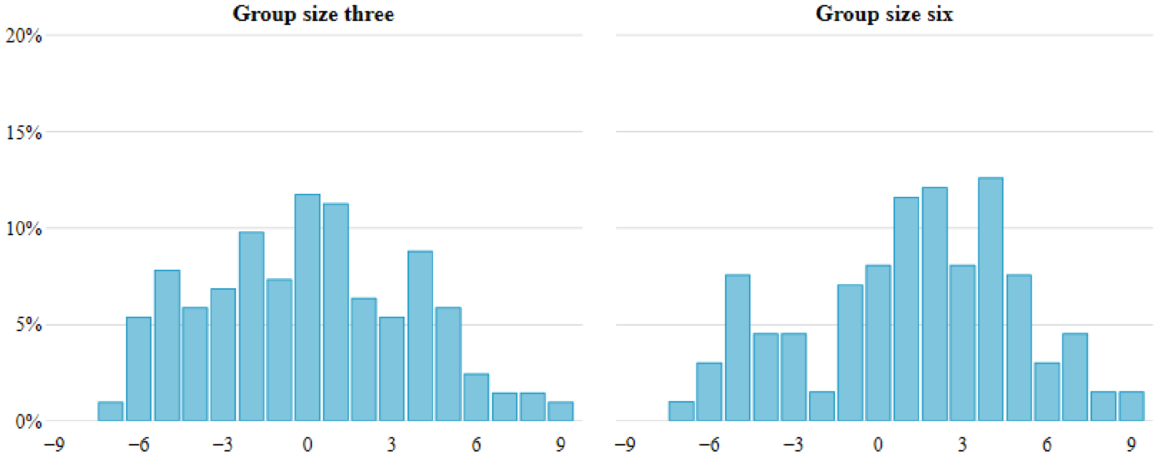
	Inconsistent Behavior		**Consistent Behavior**					
	I		II		III		IV	
Group size	Three	Six	Three	Six	Three	Six	Three	Six
Pooled data	8.9%	11.6%	91.1%	88.4%	75.9%	75.0%	54.5%	45.5%
By order of play								
Played first	11.8%	11.4%	88.2%	88.6%	79.1%	78.1%	52.7%	49.1%
Played second	6.1%	11.8%	93.9%	88.2%	72.8%	71.8%	56.1%	41.8%
By timing of beliefs								
Belief first	10.3%	13.1%	89.7%	86.9%	64.5%	66.4%	48.6%	41.1%
Choice first	7.7%	10.3%	92.3%	89.7%	86.3%	82.9%	59.8%	49.6%
By gender								
Men	6.6%	9.9%	93.4%	90.1%	73.6%	75.8%	49.5%	46.2%
Women	10.5%	12.8%	89.5%	87.2%	77.4%	74.4%	57.9%	45.1%

As in the pooled data, for groups of six, we reject the null hypothesis that the distributions of $\bar{r}_i$ and $\bar{\omega}_i$ are centered around zero in all subsamples (Wilcoxon signed-rank tests, $p < 0.027$). Moreover, we can reject the null hypothesis that the observed values in groups of three and six come from the same distribution in all subsamples except for the median switching row in the female subsample (Wilcoxon signed-rank tests, $p < 0.049$ and $p = 0.101$, respectively).

Table C5 displays the percentage of participants who are persistently competition loving, averse, and neutral for the various subsamples. Participants are classified as persistently competition loving if they switch above the competition-neutral rows in at least 4 out of 5 sets, persistently competition averse if they switch below the competition-neutral rows in at least 4 out of 5 sets, and persistently competition neutral if they switch at the competition-neutral rows in at least 4 out of 5 sets. The rest are classified as 'none.' In Panel A, participants are classified considering $r_{it} = 0$ as the only competition-neutral row. In Panel B, participants are classified considering $r_{it} = -1$, $r_{it} = 0$, and $r_{it} = 1$ as competition-neutral rows.

In all subsamples, the majority of participants are either persistently competition loving, persistently competition averse, or persistently competition neutral. With the more narrow defi-

Panel A. Distribution of mean switching rows



Panel B. Distribution of mean values of ω_{it}

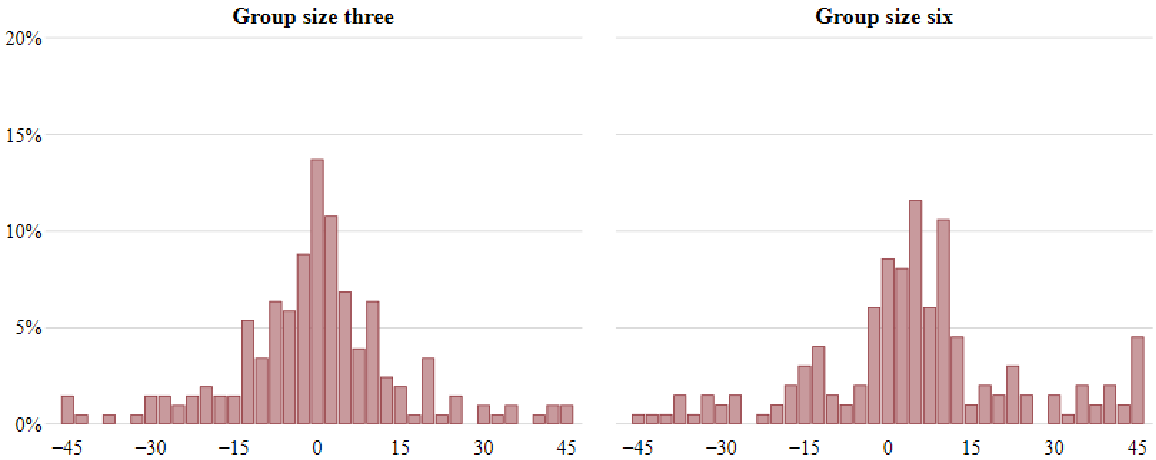


Figure C1. Distributions of additional measures of participants' preferences for competition

Notes: Panel A shows the distribution of the mean row at which participants switch from competitive to individual pay. Rows are normalized such that row zero corresponds to the highest row where the probability of earning the high amount in individual pay surpasses the participant's belief of being their group's winner. Panel B shows the distribution of the mean value of ω_{it} across decision sets, expressed as a percentage. ω_{it} represents the fraction of the utility gained from receiving the high instead of the low amount, $u_i(x_i \cdot \pi^H) - u_i(x_i \cdot \pi^L)$, participants are willing to forgo to either avoid or engage in competition. For visual ease, the distribution in Panel B is censored at −45 and 45. For groups of three, 5 out of 204 participants fall outside this range, while for groups of six, it is 10 out of 198 participants.

nition of competition-neutral rows (Panel A), in most subsamples, around 70% of participants are either persistently competition loving or persistently competition averse. With the larger definition of competition-neutral rows (Panel B), more participants are classified as persistently competition neutral, but the fraction of participants who are classified as 'none' remains roughly the same. The subsample that looks somewhat different is participants in the *Belief-first* treatment, where more than 40% of participants are classified as 'none.' However, note that, even

Table C4. Descriptive statistics of participants' median switching row ($\bar{r}_i$) and preferences for competition ($\bar{\omega}_i$) in various subsamples

Notes: Means and standard deviations in parentheses.

	Median switching row $\bar{r}_i$				**Median preference for competition $\bar{\omega}_i$**			
Group size	Three		Six		Three		Six	
Pooled data	0.14	(3.67)	1.28	(3.79)	−0.32	(15.80)	5.18	(20.03)
By order of play								
Played first	0.31	(3.88)	1.36	(3.97)	0.58	(14.91)	5.11	(21.35)
Played second	−0.02	(3.49)	1.21	(3.61)	−1.14	(16.58)	5.25	(18.68)
By timing of beliefs								
Belief first	0.33	(2.97)	1.66	(3.21)	0.53	(5.81)	3.66	(7.85)
Choice first	−0.04	(4.20)	0.95	(4.23)	−1.07	(21.03)	6.52	(26.49)
By gender								
Men	−0.32	(3.60)	1.37	(4.12)	−2.81	(14.84)	5.58	(20.55)
Women	0.46	(3.70)	1.22	(3.56)	1.46	(16.27)	4.90	(19.75)
Other inclusion criteria								
All participants	0.05	(3.71)	1.12	(3.84)	−0.66	(15.74)	4.88	(20.79)
Consistent with 3+ switches	0.16	(3.53)	1.17	(3.58)	−0.38	(16.16)	4.88	(20.76)
Consistent with 5 switches	0.00	(3.23)	0.91	(3.58)	−0.87	(16.49)	4.60	(20.35)

in this subsample, the distribution of participants classified as persistently competition loving or persistently competition averse is too high compared to a benchmark where participants do not have preferences for competition and make symmetric row-switching errors.

C.4. Variation between vs. within participants

In this section, we look at the extent to which variation in our measures of preferences for competition is explained by variation within participants versus variation between participants. The intuition is that if participants' choices are driven mostly by (uncorrelated) decision errors instead of differences in an underlying preference, then we should find relatively high variation in choices within participants compared to variation between participants. Conversely, if the impact of decision errors is minimal, then we should find that choices vary little within participants relative to the variation between participants. Since we have five sets per participant for a given group size, we can do a total variance decomposition analysis using the values of ω_{it}. Specifically, we regress the normalized rows at which participants switch or the values of ω_{it} on dummy variables indicating participant IDs and look at the resulting R^2, which corresponds to the fraction of the total variance accounted for by variance between participants. As

Table C5. Fraction of participants who are persistently competition loving, persistently competition averse, and persistently competition neutral in various subsamples

Notes: Participants are classified as persistently competition loving if they switch above the competition-neutral rows in at least 4 sets, persistently competition averse if they switch below the competition-neutral rows in at least 4 sets, and persistently competition neutral if they switch at the competition-neutral rows in at least 4 sets. The remaining participants are classified as 'None.' In Panel A, only $r_{it} = 0$ is considered as a competition-neutral row. In Panel B, $r_{it} = -1$ and $r_{it} = 1$ are also considered as competition-neutral rows.

Panel A. Competition-neutral rows defined as $r_{it} = 0$

	Group size three				**Group size six**			
	Loving	Averse	Neutral	None	Loving	Averse	Neutral	None
Pooled data	36.3%	33.8%	2.5%	27.5%	52.0%	21.7%	3.0%	23.2%
By order of play								
Played first	39.2%	34.0%	1.0%	25.8%	54.5%	22.8%	1.0%	21.8%
Played second	33.6%	33.6%	3.7%	29.0%	49.5%	20.6%	5.2%	24.7%
By timing of beliefs								
Belief first	34.4%	22.9%	3.1%	39.6%	51.6%	12.9%	4.3%	31.2%
Choice first	38.0%	43.5%	1.9%	16.7%	52.4%	29.5%	1.9%	16.2%
By gender								
Men	25.9%	38.8%	3.5%	31.8%	51.2%	20.7%	2.4%	25.6%
Women	43.7%	30.3%	1.7%	24.4%	52.6%	22.4%	3.4%	21.6%
Other inclusion criteria								
All participants	35.3%	34.8%	2.2%	27.7%	49.6%	23.2%	2.7%	24.6%
Consistent with 3+ switches	37.6%	33.5%	1.8%	27.1%	52.4%	22.0%	3.0%	22.6%
Consistent with 5 switches	36.9%	34.4%	2.5%	26.2%	49.0%	23.5%	3.9%	23.5%

Panel B. Competition-neutral rows defined as $r_{it} = \{-1, 0, 1\}$

	Group size three				**Group size six**			
	Loving	Averse	Neutral	None	Loving	Averse	Neutral	None
Pooled data	25.5%	27.0%	17.6%	29.9%	40.4%	18.2%	12.6%	28.8%
By order of play								
Played first	25.8%	24.7%	16.5%	33.0%	41.6%	18.8%	9.9%	29.7%
Played second	25.2%	29.0%	18.7%	27.1%	39.2%	17.5%	15.5%	27.8%
By timing of beliefs								
Belief first	17.7%	14.6%	26.0%	41.7%	36.6%	8.6%	18.3%	36.6%
Choice first	32.4%	38.0%	10.2%	19.4%	43.8%	26.7%	7.6%	21.9%
By gender								
Men	16.5%	32.9%	16.5%	34.1%	35.4%	15.9%	11.0%	37.8%
Women	31.9%	22.7%	18.5%	26.9%	44.0%	19.8%	13.8%	22.4%
Other inclusion criteria								
All participants	24.1%	27.2%	16.1%	32.6%	38.8%	19.2%	11.2%	30.8%
Consistent with 3+ switches	25.3%	27.1%	20.0%	27.6%	39.3%	17.9%	13.1%	29.8%
Consistent with 5 switches	22.1%	27.0%	25.4%	25.4%	36.3%	20.6%	18.6%	24.5%

Table C6. Total variance decomposition of preferences for competition

Notes: OLS regressions of participants' switching row and ω_{it} on dummy variables indicating participant IDs. Standard errors are shown in parentheses. ***, **, and * indicate statistical significance at 0.001, 0.01, and 0.05.

	Switching row		$\boldsymbol{\omega_{it}}$	
Group size	Three	Six	Three	Six
Constant	0.01	1.21***	−0.30*	5.28***
	(0.06)	(0.06)	(0.15)	(0.15)
Observations	1120	990	1120	990
Participants	204	198	204	198
R^2	0.81	0.84	0.94	0.95

before, we conduct the analyses using the participants whose switching behavior is consistent with expected utility maximization in all five decision sets. The results are presented in Table C6. We find that the within-participants component of the total variance is much smaller than the between-participants component for both the row at which participants switch, where more than 80% of the variance is due to between-participant variation, and the values of ω_{it}. where more than 93% of the variance is due to between-participant variation.

C.5. Noise in belief elicitation

As mentioned in the main body of the paper, our measure of preferences for competition relies on an accurate measurement of the participants' belief of being their group's winner. Hence, we conduct two analyses to evaluate whether noise in belief measurement is driving our main results.

For our first analysis, we check whether variance in beliefs correlates with variance in preferences for competition. For this analysis, we utilize the fact that we have two measures of beliefs per participant (one for each group size) and that we elicited beliefs both as the probability of being the group's winner and the expected percentile ranking in the performance distribution. Although the probability of being the group's winner mechanically decreases as one goes from groups of three to groups of six, the associated percentile ranking does not. Hence, an interpretation of observing a difference in a participant's expected percentile ranking is measurement error in beliefs. Using this interpretation, we classify participants as having noisy beliefs if the absolute difference between the two expected percentile rankings is larger than the median. If variation in ($\bar{r}_i$) and ($\bar{\omega}_i$) is driven by noise in belief measurement, we should observe higher variation in these measures among participants with noisy beliefs compared to the rest. Interestingly, we do not observe significantly more variation in $\bar{\omega}_i$ for participants with noisy beliefs. In groups of three, the standard deviation for participants with noisy beliefs equals 16.1%, while for participants without noisy beliefs, it is very similar at 15.1% (Levene's equality of variances test, $p = 0.896$). In groups of six, participants with noisy beliefs have a standard deviation of 21.7%, which is very similar to the standard deviation of those without noisy beliefs, 18.7% for $\bar{\omega}_i$ (Levene's equality of variances test, $p > 0.257$). Hence, we do not find evidence that variation in our measure of preferences for competition is driven noise in beliefs.

Our second analysis is based on the insights of Danz et al. (2022), who show that incentivized belief elicitation using the binarized scoring rule in some instances leads to a systematic center-bias in belief measurement. In our experiment, their findings imply that participants whose 'true' belief of being their group's winner is above 50% will tend to report a belief b_i' that is too low, $0.50 \leq b_i' \leq b_i$. Conversely, participants whose 'true' belief is below 50% will tend to report a belief that is too high, $b_i \leq b_i' \leq 0.50$. While it is not possible to test directly whether subjective beliefs are biased, a systematic center bias does have testable implications. Specifically, we would be overestimating ω_{it} for participants with beliefs above 50%, making them look more competition loving. Conversely, we would be underestimating ω_{it} for participants with beliefs below 50%, making them look more competition averse. We check for evidence of this effect by testing whether participants with beliefs above 50% have higher values of ω_{it} than participants with beliefs below 50%. We observe the opposite pattern. In both group sizes, participants with beliefs of being their group's winner above 50% have lower values of ω_{it} than those with beliefs below 50% (Mann-Whitney U tests, $p < 0.001$).[36] Hence, we do not find evidence that our measure of preferences for competition is actually capturing belief distortions due to incentivization.

C.6. Gender gap in competition entry in the initial experiment

In this section, we examine gender differences in competition entry in the initial experiment. As described in section 3, payment-scheme choices are generally not comparable between individuals since the probabilities in the decision sets depend on participants' beliefs. However, there is one exception in the *Choice-first* treatment, where participants complete an additional decision set

[36]We exclude from this analysis participants with beliefs equal to 50% (18 participants in groups of three and 3 in groups of six).

in which everyone faces the same choices. In this set, participants' make ten choices between individual and competitive pay with the probability of winning in individual pay ranging from 0.05 to 0.95. If they lose, individuals get €0, and if they win they get €3 per correct sum when competing in groups of three and €6 per correct sum when competing in groups of six. This decision set is used to narrow down the range of probabilities in subsequent sets, but we can also use them to compare the competition entry decisions of men and women.

Across the ten choices in this decision set, men choose competitive pay 54.5% of the time whereas women choose competitive pay 46.4% of the time (Mann-Whitney U test, $p = 0.020$). We explore the determinants of this gender gap in competition entry in Table C7. To do this, we use the fraction of competitive pay choices in each decision set as the dependent variable in linear regressions. Since each participant completes two of these decision sets (one per group size), We cluster standard errors on participants and control for group size and order in all regressions. We also standardize the continuous dependent variables to have a mean of zero and a standard deviation of one. Lastly, the analysis is restricted to decision sets where participants exhibited switching behavior consistent with expected utility maximization.

In column I, we include the gender dummy as an independent variable, confirming that women are significantly less likely to choose competitive pay than men ($p = 0.009$). In column II, we include as dependent variables the participants' belief and their actual probability of being the group's winner. Beliefs of being their group's winner have a strong significantly positive effect on the likelihood of choosing competitive pay. Interestingly, once we control for the participants' beliefs, the coefficient of the gender dummy is significantly smaller ($p = 0.021$) and no longer statistically significant ($p = 0.095$). In columns III and IV, we add the participants' preferences for competition. Reassuringly, both $\bar{r}_i$ (column III) and $\bar{\omega}_i$ (column IV) are strong significant predictors of their choice to compete in the additional decision set ($p < 0.001$ for both). Controlling for preferences for competition further reduces the coefficient of the gender dummy, but the change is not statistically significant ($p = 0.183$ when comparing the coefficients in columns II and III and $p = 0.139$ when comparing II and IV). In columns V and VI, we include participants' risk preferences as either their median switching row in the risk decision sets (column V) or their median coefficient of CRRA (column VI). Neither measure is significantly associated with competition entry ($p = 0.079$ in column V and $p = 0.185$ in column VI) and their inclusion has a negligible effect on the gender coefficient, reducing it by less than 0.1 percentage points ($p = 0.0.856$ when comparing columns III and V and $p = 0.670$ when comparing columns

Table C7. Choice of competitive pay in additional decision sets of the *Choice-first* treatment

Notes: Linear regressions of the fraction of competitive pay choices in the first additional decision set of the *Choice-first* treatment in the initial experiment. All regressions control for group size and order. Columns VII and VIII include sets of dummy variables for which we report the F-test for joint significance instead of individual coefficients. Standard errors clustered on participants are shown in parentheses. Data corresponds to participants whose switching behavior is consistent with expected utility maximization. ** and * indicate statistical significance at 0.01 and 0.05.

	I	II	III	IV	V	VI	VII	VIII
Woman	-8.0^{**}	-4.8	-2.2	-0.7	-2.2	-0.6	-0.7	-0.2
	(3.1)	(2.9)	(1.8)	(1.1)	(1.8)	(1.1)	(1.7)	(1.4)
Belief of being the		6.2^{**}	13.2^{**}	20.1^{**}	12.7^{**}	19.8^{**}	$F_{9,116}$	$F_{9,116}$
group's winner		(1.7)	(1.3)	(0.9)	(1.4)	(0.9)	$= 15.7^{**}$	$= 39.5^{**}$
Probability of being the		0.3	0.7	0.2	0.8	0.2	$F_{9,116}$	$F_{9,116}$
group's winner		(1.6)	(1.0)	(0.6)	(1.0)	(0.6)	$= 1.7$	$= 1.2$
Competition $\bar{r}_i$			13.2^{**}		12.8^{**}		$F_{9,116}$	
			(1.0)		(1.0)		$= 24.3^{**}$	
Competition $\bar{\omega}_i$				20.2^{**}		20.0^{**}		$F_{9,116}$
				(0.7)		(0.7)		$= 58.9^{**}$
Risk $\bar{r}_i$					1.8		$F_{5,116}$	
					(1.0)		$= 2.8^{*}$	
Risk $\bar{\alpha}_i$						-0.8		$F_{9,113}$
						(0.6)		$= 1.6$
Observations	228	228	228	228	228	228	228	228
Participants	117	117	117	117	117	117	117	117
R^2	0.09	0.20	0.55	0.83	0.56	0.83	0.63	0.79

IV and VI). The fact that risk aversion does not influence the payment-scheme choice is not surprising as the experiment is designed to remove the effect of risk preferences (see section 3). In columns VII and VIII, we give the independent variables more flexibility by dividing them into deciles and including them as dummy variables. For these variables, we report the F-test of joint significance instead of individual coefficients. In these columns, the gender coefficient is further reduced, but only by 1.5 percentage points in column VII ($p = 0.175$ when compared to column V) and 0.4 percentage points in column VIII ($p = 0.670$ when compared to column VI).

Overall, these findings align with those of the follow-up experiment. Our results indicate that women choose to compete less frequently than men, but this difference is better explained by gender differences in other factors, such as beliefs in the initial experiment and risk preferences in the follow-up experiment, rather than by differences in preferences for competition.

Table D1. Descriptive statistics for the addition task and competition-entry choice (follow-up experiment)

Notes: Mean and standard deviations in parenthesis for the number of correct sums in the addition task in parts I and II, the probability of winning in part I (in percent), and the fraction choosing competitive pay in part I (in percent). Statistics are presented by session timing and by gender.

	Correct sums part I		**Correct sums part II**		**Prob. winning part I**		**Compete part I**	
Pooled data	11.5	(4.6)	12.5	(4.7)	32.9	(31.1)	39.4	(49.0)
By session timing								
First session	9.8	(4.2)	11.5	(4.3)	32.8	(30.9)	44.9	(50.0)
Second session	13.2	(4.4)	13.5	(4.9)	33.0	(31.4)	33.9	(47.5)
By gender								
Men	12.3	(4.9)	13.4	(5.0)	38.0	(31.9)	53.3	(50.2)
Women	11.0	(4.3)	12.0	(4.5)	29.6	(30.3)	30.6	(46.2)

Appendix D. Auxiliary analysis of the follow-up experiment

D.1. Descriptive statistics of the addition task

For the follow-up experiment, Table D1 shows the means and standard deviations for the number of correct sums in parts I and II, the probability of winning in part I (i.e., the competition-entry choice), and the fraction choosing competitive pay in part I. Statistics are presented for the pooled data as well as depending on whether it was the first or second session and on gender. On average, participants answer 11.5 sums in part I and 12.5 in part II. Comparing the first and second sessions reveals that they improved their performance in the addition task in both parts (Wilcoxon signed-ranked tests, $p < 0.001$).

D.2. Payment-scheme choices

We begin with the analysis of participants' payment-scheme choices. Figure D1 depicts the fraction of competitive pay choices per row of the decision sets. Rows are normalized such that row zero, $r_{it} = 0$, corresponds to the highest row in set t where the probability of earning the high amount in individual pay exceeds i's probability of being their group's winner. If participants do not have preferences for competition, they should choose competitive pay for rows $r_{it} < 0$ and individual pay for rows $r_{it} \geq 0$.

Figure D1 shows that participants choose competitive pay less often as individual pay becomes relatively more attractive. However, as in the initial experiment, many participants choose

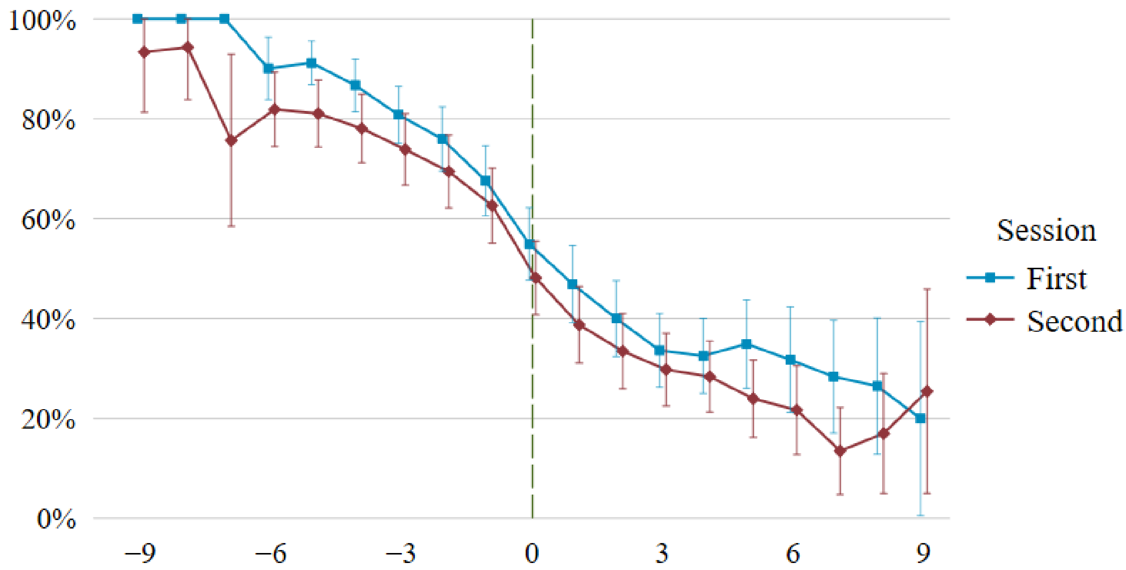


Figure D1. Fraction of competitive pay choices per normalized row in the decision sets (follow-up experiment)

Notes: Rows within decision sets are normalized such that row zero corresponds to the highest row where the probability of earning the high amount in individual pay exceeds the participant's probability of being their group's winner according to their performance in part I. 95% confidence intervals are computed using a linear probability model regressing the payment-scheme choice on dummy variables for the normalized row values, with standard errors clustered by participant.

competitive pay when the expected value of individual pay is higher (e.g., in row 1, 46.9% choose competitive pay in the first session and 38.8% in the second) and individual pay when the expected value of competitive pay is higher (e.g., in row −1, 32.4% choose individual pay in the first session and 37.4% in the second). The figure also shows no consistent differences between sessions. Regressing the payment-scheme choice on dummy variables of the normalized row values interacted with the session and clustering standard errors on participants finds statistically significant differences in rows −7, −5, −4, 5, and 7 ($p < 0.033$).

D.3. Switching behavior

Table D2 summarizes the number of times and the direction in which participants switch within a decision set. Specifically, it shows the fraction of decision sets in which there are multiple switches or a unique non-monotonic switch (a switch from individual to competitive pay), which we classify as *inconsistent*, and the fraction of decision sets with a single switch from competitive to individual pay or no switch at all, which we classify as *consistent*. Similar to the initial experiment, we observe that more than 95% of the decision sets are consistent, with a majority of sets having a single switch. The number of consistent decision sets per participant is

Table D2. Consistency of switching behavior with utility maximization within decision sets (follow-up experiment)

Notes: Fraction of decision sets classified as consistent and inconsistent with expected utility maximization. Inconsistent decision sets contain *multiple switches* or a unique *non-monotonic switch* from individual to competitive pay. Consistent decision sets contain a *single switch* from competitive to individual pay or *no switch*.

	Inconsistent Behavior				**Consistent Behavior**			
	Multiple switches		*Non-monotonic switch*		*Single switch*		*No switch*	
Session	First	Second	First	Second	First	Second	First	Second
Pooled data	2.5%	2.4%	0.7%	0.8%	66.9%	60.0%	29.8%	36.8%
By gender								
Men	1.3%	0.4%	0.9%	0.4%	65.7%	57.4%	32.2%	41.7%
Women	3.3%	3.6%	0.6%	1.1%	67.8%	61.7%	28.3%	33.6%

not significantly different across sessions (Wilcoxon signed-rank test, $p = 0.446$). Moreover, it is not significantly different from the number of consistent decision sets per participant in groups of three in the initial experiment (Mann-Whitney U tests, $p > 0.638$). Although, compared to the initial experiment, we find around 10% more decision sets with no switch.

Table D3 summarizes participants' switching behavior across the five decision sets. The table shows the percentage of participants with (I) at least one inconsistent decision set, (II) five consistent decision sets, (III) five consistent decision sets and a switch in a majority of sets (three or more), and (IV) five consistent decision sets and a switch in all five sets. In both sessions, around 90% of participants have five consistent decision sets, with a majority of participants switching once from competitive to individual pay in at least three sets. The fraction of participants in each category in Table D3 does not vary significantly across sessions (McNemar's χ^2 tests, $p > 0.417$). Compared to groups of three in the initial experiment, we find a similar fraction of participants with five consistent decision sets (Mann-Whitney U tests, $p > 0.658$).

D.4. Preferences for competition

Panel A of Figure D2 depicts the distribution of $\bar{r}_i$, the median row at which participants switch across the five decision sets. Panel B depicts the distribution of $\bar{\omega}_i$, the median value of ω_{it} across the five decision sets. To obtain an accurate representation of participants' preferences for competition, this analysis is based on participants whose decisions are consistent with expected utility maximization in all five decision sets (105 out of 118 participants in the first session and

Table D3. Consistency of switching behavior with utility maximization across decision sets (follow-up experiment)

Notes: Fraction of participants according to the number of consistent and inconsistent decision sets. Inconsistent decision sets contain multiple switches or a unique switch from individual to competitive pay. Consistent decision sets contain a single switch from competitive to individual pay or no switch. Fraction of participants with at least one inconsistent decision set in column I, five consistent decision sets in column II, five consistent decision sets and a single switch from competitive to individual pay in a majority of decision sets in column III, and five consistent decision sets and a single switch from competitive to individual pay in all decision sets in column IV.

	Inconsistent Behavior		**Consistent Behavior**					
	I		II		III		IV	
Session	First	Second	First	Second	First	Second	First	Second
Pooled data	11.0%	7.6%	89.0%	92.4%	61.0%	59.3%	40.7%	35.6%
By gender								
Men	8.7%	4.3%	91.3%	95.7%	58.7%	58.7%	39.1%	28.3%
Women	12.5%	9.7%	87.5%	90.3%	62.5%	59.7%	41.7%	40.3%

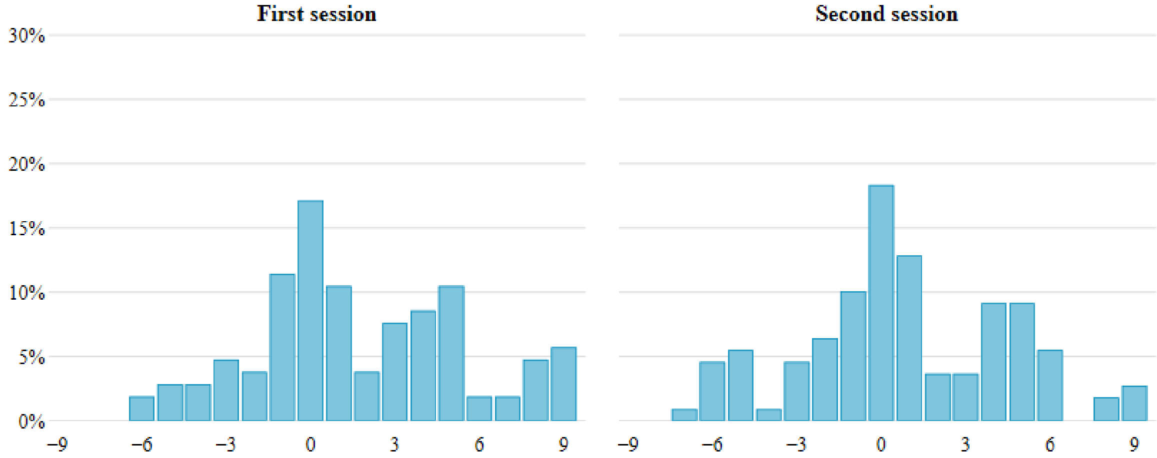
109 out of 118 in the second).

As in the initial experiment, Panel A shows that a majority of participants have a non-zero value of $\bar{r}_i$: 82.7% in the first session and 81.7% in the second. More specifically, 55.2% of participants in the first session and 48.6% in the second have a positive $\bar{r}_i$, suggesting they like competition, while 27.6% in the first session and 33.0% in the second have a negative $\bar{r}_i$, suggesting they dislike competition. Similarly, Panel B confirms that a majority of participants are willing to sacrifice money to either engage in or avoid competition. For instance, 38.1% of participants in the first session and 31.2% in the second have an $\bar{\omega}_i$ indicating they are prepared to forgo more than 5.0% of the utility gain of receiving the high amount to ensure they compete. Conversely, 14.3% in the first session and 19.3% in the second are prepared to forgo at least 5.0% of this utility gain to avoid competing. These are not small amounts considering that the median earnings in this task equal 45% of the additional return of receiving the high amount.

The distributions across sessions look similar. However, there is some evidence of a slight decrease in the values of $\bar{r}_i$ (from 1.72 to 0.85; Wilcoxon signed-rank test, $p = 0.077$) and $\bar{\omega}_i$ (from 2.86% to 1.15%; Wilcoxon signed-rank test, $p = 0.066$) from the first to the second session. Compared to the distributions of groups of three in the initial experiment, we find that the values of $\bar{r}_i$ and $\bar{\omega}_i$ are significantly higher in the first session but not in the second (Wilcoxon signed-rank tests, $p < 0.013$ and $p > 0.119$, respectively).

Table D4 shows the mean and standard deviation of the participants' median switching row

Panel A. Distribution of $\bar{r}_i$



Panel B. Distribution of $\bar{\omega}_i$

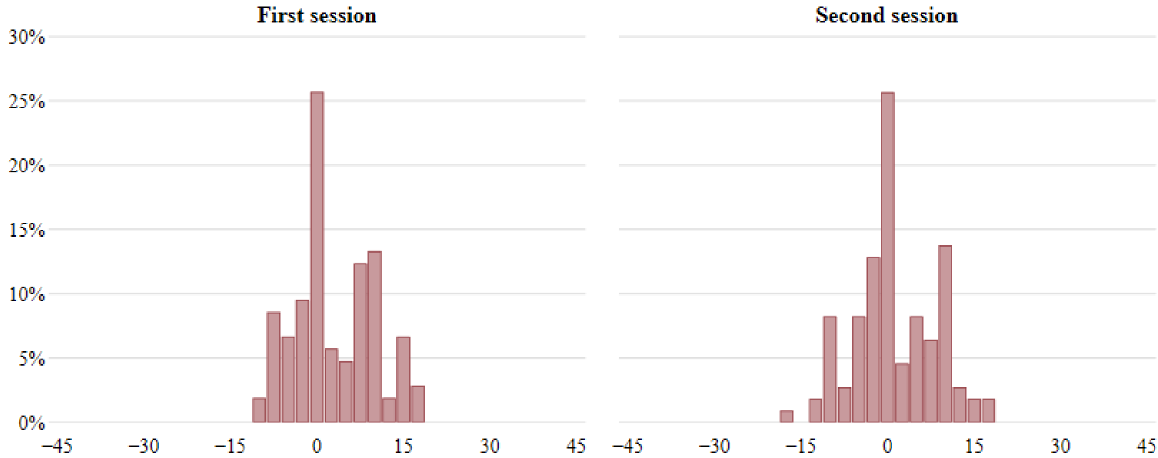


Figure D2. Distributions of measures of participants' preferences for competition (follow-up experiment

Notes: Panel A shows the distribution of $\bar{r}_i$, the median row at which participants switch from competitive to individual pay. Rows are normalized such that row zero corresponds to the highest row where the probability of earning the high amount in individual pay surpasses the participant's probability of being their group's winner. Panel B shows the distribution of $\bar{\omega}_i$, the median value of ω_{it} across decision sets, expressed as a percentage. ω_{it} represents the fraction of the utility gained from receiving the high instead of the low amount, $u_i(x_i \cdot \pi^H) - u_i(x_i \cdot \pi^L)$, participants are willing to forgo to either avoid or engage in competition.

($\bar{r}_i$) and preferences for competition ($\bar{\omega}_i$) depending on gender and when using other criteria to include participants in the analysis. These criteria include: all 118 participants, participants with five consistent decision sets and fewer than two sets without a switch (72 participants in the first session and 70 in the second), and participants with five consistent decision sets and no sets without a switch (48 participants in the first session and 42 in the second). When analyzing these subsamples, we observe that the significant differences in the values of $\bar{r}_i$ and $\bar{\omega}_i$ between

Table D4. Descriptive statistics of participants' median switching row ($\bar{r}_i$) and preferences for competition ($\bar{\omega}_i$) in various subsamples (follow-up experiment)

Notes: Means and standard deviations in parentheses.

	Median switching row ($\bar{r}_i$)				**Median preference for competition ($\bar{\omega}_i$)**			
Session	First		Second		First		Second	
Pooled data	1.72	(3.76)	0.85	(3.70)	2.86	(7.10)	1.15	(7.09)
By gender								
Men	1.57	(3.86)	0.68	(3.99)	2.29	(7.03)	0.81	(7.64)
Women	1.83	(3.72)	0.97	(3.52)	3.25	(7.19)	1.38	(6.74)
Other inclusion criteria								
All participants	1.58	(3.63)	0.88	(3.77)	2.58	(6.87)	1.14	(7.22)
Consistent with 3+ switches	0.79	(3.01)	0.56	(2.55)	1.06	(6.36)	0.64	(5.56)
Consistent with 5 switches	0.67	(2.56)	0.43	(1.85)	0.78	(5.37)	0.45	(4.37)

the first session and the initial experiment are present among men (Mann-Whitney U tests, $p < 0.017$ for both $\bar{r}_i$ and $\bar{\omega}_i$) but less so among women (Mann-Whitney U tests, $p = 0.049$ for $\bar{r}_i$ and $p = 0.186$ for $\bar{\omega}_i$). Moreover, these differences disappear when the sample is restricted to participants with five consistent decision sets with a switch in at least three sets (Mann-Whitney U tests, $p > 0.231$ for both $\bar{r}_i$ and $\bar{\omega}_i$).

Table D5 classifies participants based on their switching patterns across sets. Participants who switch above the competition-neutral rows in at least four sets are classified as persistently competition loving; those who switch below in at least four sets are persistently competition averse; and those who switch at the competition-neutral rows in at least four sets are persistently competition neutral. Participants who do not fit these criteria are categorized as 'none.' We use two definitions of competition-neutral rows. The first includes only $r_{it} = 0$ while the second uses a broader definition, including also rows $r_{it} = -1$ and $r_{it} = 1$. The table shows this classification for all participants whose switching behavior is consistent with expected utility maximization in all five sets, as well as by gender and using other inclusion criteria.

With both definitions of competition-neutral rows, a clear majority of participants are classified as either persistently competition loving, persistently competition averse, or persistently competition neutral. Naturally, fewer participants are classified as persistently competition neutral with the stricter definition of competition-neutral rows. The fraction of participants classified as one of the persistent types does not vary significantly from the first to the second session (McNemar's χ^2 tests, $p > 0.755$). The same holds when pooling the persistently competition

Table D5. Fraction of participants who are persistently competition loving, persistently competition averse, and persistently competition neutral in various subsamples (follow-up experiment)

Notes: Participants are classified as persistently competition loving if they switch above the competition-neutral rows in at least 4 sets, persistently competition averse if they switch below the competition-neutral rows in at least 4 sets, and persistently competition neutral if they switch at the competition-neutral rows in at least 4 sets. The remaining participants are classified as 'None.' In Panel A, only $r_{it} = 0$ is considered as a competition-neutral row. In Panel B, $r_{it} = -1$ and $r_{it} = 1$ are also considered as competition-neutral rows.

Panel A. Competition-neutral rows defined as $r_{it} = 0$

	First session				**Second session**			
	Loving	Averse	Neutral	None	Loving	Averse	Neutral	None
Pooled data	43.8%	19.0%	5.7%	31.4%	36.7%	25.7%	3.7%	33.9%
By gender								
Men	45.2%	16.7%	7.1%	31.0%	31.8%	27.3%	4.5%	36.4%
Women	42.9%	20.6%	4.8%	31.7%	40.0%	24.6%	3.1%	32.3%
Other inclusion criteria								
All participants	42.4%	20.3%	5.1%	32.2%	37.3%	25.4%	3.4%	33.9%
Consistent with 3+ switches	31.9%	22.2%	8.3%	37.5%	30.0%	21.4%	5.7%	42.9%
Consistent with 5 switches	29.2%	20.8%	12.5%	37.5%	23.8%	21.4%	9.5%	45.2%

Panel B. Competition-neutral rows defined as $r_{it} = \{-1, 0, 1\}$

	First session				**Second session**			
	Loving	Averse	Neutral	None	Loving	Averse	Neutral	None
Pooled data	39.0%	11.4%	21.9%	27.6%	27.5%	16.5%	26.6%	29.4%
By gender								
Men	38.1%	11.9%	21.4%	28.6%	27.3%	22.7%	20.5%	29.5%
Women	39.7%	11.1%	22.2%	27.0%	27.7%	12.3%	30.8%	29.2%
Other inclusion criteria								
All participants	36.4%	11.0%	19.5%	33.1%	27.1%	16.9%	24.6%	31.4%
Consistent with 3+ switches	26.4%	12.5%	29.2%	31.9%	17.1%	10.4%	35.7%	37.1%
Consistent with 5 switches	22.9%	10.4%	35.4%	31.2%	14.3%	4.8%	45.2%	35.7%

loving and averse categories (Stuart–Maxwell marginal homogeneity tests, $p > 0.458$). When considering all four types, there is weak evidence of a decrease in competition-loving behavior in the second session, although the difference is not statistically significant (Stuart–Maxwell marginal homogeneity tests, $p = 0.241$ for the strict definition and $p = 0.101$ for the broader definition of competition neutrality). Compared to groups of three in the initial experiment, there are no significant differences in the fraction classified as one of the persistent types (χ^2 tests, $p > 0.231$) or when pooling persistently competition loving and averse participants (χ^2 tests, $p > 0.159$). With the four types, we find some indication that, compared to the initial

Table D6. Test-retest correlation coefficients between the first and second sessions in various subsamples (follow-up experiment)

Notes: Pearson's correlation coefficients between the first and second sessions. We use a structural equation model with standardized variables to estimate the correlation matrix. The p-values in the table corresponds to Wald tests of the difference between the correlation coefficients of competition and risk.

	Median switching row $\bar{r}_i$			**Median preference** $\bar{\omega}_i$ **or** $\bar{\alpha}_i$		
Type of preference	Competition	Risk	Diff. p-value	Competition	Risk	Diff. p-value
Pooled data ($n = 100$)	0.38	0.55	0.245	0.36	0.51	0.277
By gender						
Men ($n = 41$)	0.41	0.44	0.891	0.38	0.43	0.852
Women ($n = 59$)	0.35	0.60	0.091	0.34	0.53	0.197
Other inclusion criteria						
All participants ($n = 118$)	0.41	0.55	0.322	0.41	0.51	0.447
Consistent with 3+ switches ($n = 52$)	0.47	0.60	0.356	0.48	0.53	0.742
Consistent with 5 switches ($n = 26$)	0.66	0.65	0.938	0.67	0.64	0.873

experiment, there is a higher fraction of participants classified as persistently competition loving in the first session of the follow-up experiment (χ^2 tests, $p = 0.241$) but not in the second (χ^2 tests, $p > 0.103$).

D.5. Stability of preferences for competition and risk

Table D6 shows the Pearson's correlation coefficients between the first and second sessions of the follow-up experiment, depending on gender and the criteria used to include participants in the analysis. In addition to the sample of participants with switching behavior consistent with expected utility maximization in all decision sets across both sessions (the pooled sample), these criteria include: all 118 participants, the 52 participants with five consistent decision sets and fewer than two sets without a switch in both sessions, and the 26 participants with five consistent decision sets and no sets without a switch in both sessions.

The first two columns present the correlation for the median switching row in the competition and risk decision sets. The third column reports the p-value for testing whether the difference between the coefficients for risk and competition is statistically significant. Specifically, we use a structural equation model with standardized variables to obtain the correlation matrix and then test the coefficients using Wald tests. The fourth and fifth columns correspond to the correlation for $\bar{\omega}_i$ and $\bar{\alpha}_i$, respectively, with the sixth column testing the difference between these two coefficients.

Appendix E. Auxiliary analysis of risk preferences

E.1. Details of the experimental design

Here, we describe the four decision sets used to measure participants' risk preferences in the initial and follow-up experiments. Each decision set is an MPL with ten rows $r \in [1, 10]$. Each row contains a choice between a certain amount, π_r^C, and a lottery that pays a high amount, π^H, with probability q and a low amount, π^L, with probability $1 - q$. As one goes down the MPL, the certain amount increases by z (i.e., $\pi_{r+1}^C = \pi_r^C + z$). Hence, in a given MPL, the certain amounts range from π_1^C to $\pi_1^C 0 = \pi_1^C + 9z$. The values of π^H, π^L, q, and z remain constant within a decision set but vary across sets. Table E1 contains the high π^H and low π^L amounts used in each decision set, along with the probability q of getting the high amount, the value of z, and the lowest certain amount π_1^C. As seen in the table, we also introduced small uniformly distributed random components to q and π_1^C. After participants make their choices, one choice in a decision set is randomly selected to determine their earnings from this task. The order in which decision sets appear to participants is randomized.

Table E1. Parameters of the four decision sets used to elicit risk preferences

	Decision set			
	1	2	3	4
π^H	36.00	72.00	72.00	108.00
π^L	0.00	0.00	0.00	0.00
q	0.50 ± 0.02	0.50 ± 0.02	0.30 ± 0.02	0.30 ± 0.02
π_1^C	14.40 ± 0.72	32.40 ± 1.44	14.40 ± 1.44	27.00 ± 2.16
z	0.72	0.72	1.44	1.08

E.2. Payment-scheme choices

We begin with the analysis of participants' choices. Figure D1 depicts the fraction of lottery choices per row of the decision sets. Rows are normalized such that row zero, $r_{it} = 0$, corresponds to the first row in set t where the certain amount exceeds the lottery's expected value. If participants are risk neutral, they should choose the lottery for rows $r_{it} < 0$ and the certain amount for rows $r_{it} \geq 0$.

As expected, participants choose the lottery less often as the certain amount increases. How-

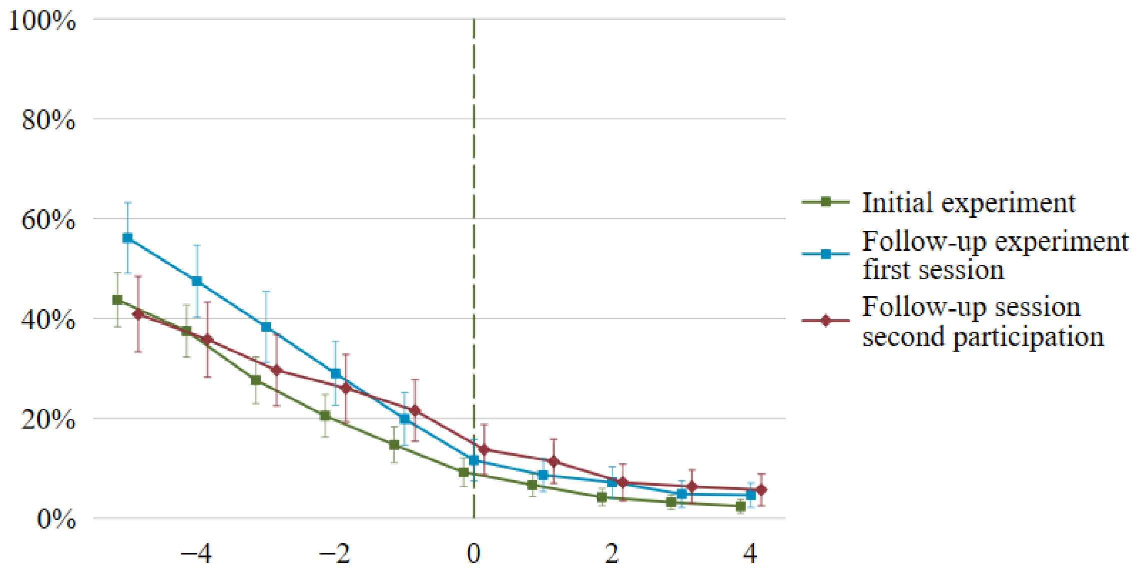


Figure E1. Fraction of lottery choices per normalized row in the decision sets (initial and follow-up experiments)

Notes: Rows within decision sets are normalized such that row zero corresponds to the first row where the certain amount equals or exceeds the expected value of the lottery. 95% confidence intervals are computed using a linear probability model regressing the lottery choice on dummy variables for the normalized row values, with standard errors clustered by participant.

ever, many participants choose the certain amount when the lottery pays more, suggesting they are risk averse. In fact, except for one case (row −5 in the first session), at least 50% of participants always choose the certain amount. By contrast, the fraction of participants who choose the lottery when the certain amount pays more never exceeds 11%. The figure also shows that there are no qualitative differences between experiments. There seems to be a shift toward fewer lottery choices from the first to the second session of the follow-up experiment.

E.3. Switching behavior

Table E2 describes participants' switching behavior within decision sets. It shows the fraction of decision sets *inconsistent* with expected utility maximization, either due to multiple switches or a single non-monotonic switch from the certain amount to the lottery, and the fraction *consistent* with expected utility maximization, with either a single switch from the lottery to the certain amount or no switch at all. We find that more than 95% of the decision sets of the risk MPLs are consistent. Comparing the number of consistent decision sets per participant across the risk and competition MPLs, we find they are very similar and statistically indistinguishable in both sessions of the follow-up experiment and in the initial experiment (Wilcoxon signed-rank tests, $p > 0.229$). However, within consistent decision sets, there are 14% to 36% more sets without

Table E2. Consistency of switching behavior with utility maximization within decision sets of lottery choices (initial and follow-up experiments)

Notes: Fraction of decision sets classified as consistent and inconsistent with expected utility maximization. Inconsistent decision sets contain *multiple switches* or a unique *non-monotonic switch* from the certain amount to the lottery. Consistent decision sets contain a *single switch* from the lottery to the certain amount or *no switch*. Data from the initial experiment and both sessions of the follow-up experiment.

			Initial	**Follow-up**	
				First	Second
Pooled data	Inconsistent decision sets	Multiple switches	2.5%	3.4%	3.4%
		Non-monotonic switch	0.9%	1.1%	0.2%
	Consistent decision sets	Single switch	41.2%	51.7%	34.1%
		No switch	55.5%	43.9%	62.3%
Men	Inconsistent decision sets	Multiple switches	2.5%	0.5%	1.1%
		Non-monotonic switch	0.3%	1.6%	0.0%
	Consistent decision sets	Single switch	42.6%	62.5%	47.8%
		No switch	54.7%	35.3%	51.1%
Women	Inconsistent decision sets	Multiple switches	2.4%	5.2%	4.9%
		Non-monotonic switch	1.3%	0.7%	0.3%
	Consistent decision sets	Single switch	40.2%	44.8%	25.3%
		No switch	56.0%	49.3%	69.4%

any switch in the risk MPLs compared to the competition MPLs.

Table E3 summarizes participants' switching behavior across the four decision sets of the risk MPLs. The table shows the percentage of participants with (I) at least one inconsistent decision set, (II) four consistent decision sets, (III) four consistent decision sets and a switch in a majority of sets (two or more), and (IV) four consistent decision sets and a switch in all four sets. As with the competition MPLs, we find that around 90% of the participants display behavior consistent with expected utility maximization in all sets of the risk MPLs. If we compare this fraction across the two types of MPLs, we do not find a statistically significant difference in either session of the follow-up experiment or in the initial experiment (Wilcoxon signed-rank tests, $p > 0.404$). However, in line with there being more sets without switching in the risk MPLs, we observe a smaller fraction of participants with either a majority or all sets involving a single switch.

E.4. Risk preferences

In this section, we present the analysis of participants' risk preferences. As in the previous analyses, we focus on participants whose switching behavior is consistent with expected utility

Table E3. Consistency of switching behavior with utility maximization across decision sets of lottery choices (initial and follow-up experiments)

Notes: Notes: Fraction of participants according to the number of consistent and inconsistent decision sets. Inconsistent decision sets contain multiple switches or a unique switch from the certain amount to the lottery. Consistent decision sets contain a single switch from the lottery to the certain amount or no switch. Data from the initial experiment and both sessions of the follow-up experiment.

		Initial	**Follow-up**	
		Group size three	First session	Second session
Pooled data	At least one inconsistent decision set	6.7%	9.3%	8.5%
	Four consistent decision sets	93.3%	90.7%	91.5%
	...and a switch in a majority of sets	42.0%	55.1%	39.0%
	...and a switch in all four sets	21.9%	28.8%	15.3%
Men	At least one inconsistent decision set	6.7%	9.3%	8.5%
	Four consistent decision sets	93.3%	90.7%	91.5%
	...and a switch in a majority of sets	42.0%	55.1%	39.0%
	...and a switch in all four sets	21.9%	28.8%	15.3%
Women	At least one inconsistent decision set	6.7%	9.3%	8.5%
	Four consistent decision sets	93.3%	90.7%	91.5%
	...and a switch in a majority of sets	42.0%	55.1%	39.0%
	...and a switch in all four sets	21.9%	28.8%	15.3%

maximization in all four decision sets: 209 out of 224 in the initial experiment, 107 out of 118 in the first session of the follow-up experiment, and 108 out of 118 in the second.

Mirroring the analysis of preferences for competition, we analyze various measures of participants' risk preferences. First, we use the row at which participants switch in a decision set. In sets where participants do not switch, we use the highest row if the lottery is always chosen and the lowest row if the certain amount is always chosen. Second, for each set t, we calculate participant i's normalized risk premium. Specifically, we first take the difference between the expected value of the lottery and the certain amount that makes i indifferent: $p_{it} \cdot \pi_{it}^{H} + (1 - p_{it}) \cdot \pi_{it}^{L} - \pi_{it}^{C*}$, where π_{it}^{C*} is the midpoint between the certain amount of the row at which i switched and the certain amount of the preceding row. Then, to make the risk premium comparable across lotteries, we normalize it with the difference between the lottery's high and low amounts, $\left(\pi_{it}^{H} - \pi_{it}^{L}\right)$. This gives us the risk premium $RP_{it} = p_{it} - \left(\pi_{it}^{C*} - \pi_{it}^{L}\right) / \left(\pi_{it}^{H} - \pi_{it}^{L}\right)$, which we can interpret as the fraction of the gain of winning the lottery that the participant is willing to give up to receive a certain amount. Positive values of RP_{it} imply the participant is risk averse, while negative values imply they are risk seeking. Finally, as a third measure, we follow the literature and calculate for each set the coefficient of constant relative risk aversion, α_{it}, that makes participants

Table E4. Descriptive statistics of participants' median switching row, risk premium, and CRRA coefficient (initial and follow-up experiments)

Notes: Means and standard deviations in parentheses. Data from participants whose switching behavior is consistent with expected utility maximization in all four decision sets in the initial experiment and both sessions of the follow-up experiment.

		Initial	**Follow-up**	
		Group size three	First session	Second session
Pooled data	Median switching row $\bar{r}_i$	−3.56 (2.05)	−2.98 (2.29)	−3.14 (2.69)
	Median risk premium $\bar{RP}_i$	5.21 (2.73)	4.62 (2.95)	4.94 (3.57)
	Median coefficient of CRRA $\bar{\alpha}_i$	0.24 (0.13)	0.22 (0.14)	0.23 (0.17)
Men	Median switching row $\bar{r}_i$	−3.36 (2.16)	−2.30 (2.54)	−2.23 (2.92)
	Median risk premium $\bar{RP}_i$	4.85 (2.83)	3.86 (3.22)	3.66 (3.85)
	Median coefficient of CRRA $\bar{\alpha}_i$	0.22 (0.13)	0.18 (0.15)	0.17 (0.18)
Women	Median switching row $\bar{r}_i$	−3.71 (1.97)	−3.46 (1.98)	−3.77 (2.35)
	Median risk premium $\bar{RP}_i$	5.45 (2.64)	5.15 (2.64)	5.83 (3.09)
	Median coefficient of CRRA $\bar{\alpha}_i$	0.25 (0.12)	0.24 (0.12)	0.28 (0.15)

indifferent at the switching point (Holt and Laury, 2002). That is, we solve numerically the expression:

$$\left(\left(\pi_{it}^{C*}\right)^{1-\alpha_{it}}-1\right)/\left(1-\alpha_{it}\right)=p_{it}\cdot\left(\left(\pi_{it}^{H}\right)^{1-\alpha_{it}}-1\right)/\left(1-\alpha_{it}\right)+\left(1-p_{it}\right)\cdot\left(\left(\pi_{it}^{L}\right)^{1-\alpha_{it}}-1\right)/\left(1-\alpha_{it}\right).$$

For each measure, we use the median value across the four sets as a participant's risk preferences.

Table E4 shows the mean and standard deviation of the participants' median switching row ($\bar{r}_i$), risk premium, ($\bar{RP}_i$), and CRRA coefficient ($\bar{\alpha}_i$) for participants whose switching behavior is consistent with expected utility maximization in all decision sets and depending on gender. As in most of the experimental literature, on average, participants are risk averse. For all three measures, we can reject the null hypothesis of risk neutrality (Wilcoxon signed-rank tests, $p < 0.001$). Also, consistent with most of the literature on gender differences, we see that women tend to be more risk averse than men. The gender difference is statistically significant for the

Table E5. Descriptive statistics of participants' preferences for competition adjusted for probability weighting $\bar{\omega}'_i$ in various subsamples (initial and follow-up experiments)

Notes: Means and standard deviations in parentheses.

	Initial experiment				**Follow-up experiment**			
Group size / Session	Three		Six		First		Second	
Pooled data	−0.30	(12.01)	4.14	(15.22)	2.48	(8.36)	0.90	(8.03)
By gender								
Men	−2.28	(12.39)	4.63	(15.90)	1.58	(7.83)	0.64	(8.34)
Women	1.11	(11.58)	3.80	(14.78)	3.08	(8.70)	1.07	(7.88)
Other inclusion criteria								
All participants	−0.55	(11.90)	3.86	(15.84)	2.22	(8.04)	0.98	(8.08)
Consistent with 3+ switches	−0.40	(12.21)	3.73	(15.40)	0.97	(8.28)	0.44	(7.36)
Consistent with 5 switches	−1.04	(12.19)	3.62	(15.17)	0.44	(7.27)	−0.02	(6.71)

three measures in the first and second sessions of the follow-up experiment (Mann-Whitney U tests, $p < 0.034$) and for the CRRA coefficient in the initial experiment (Mann-Whitney U tests, $p = 0.049$ for $\bar{\alpha}_i$, $p = 0.160$ for $\bar{RP}_i$, and $p = 0.174$ for $\bar{r}_i$).

E.5. Probability weighting

In this subsection, we analyze the potential impact of probability weighting on our main results. Probability weighting does not impact the analyses of switching behavior, the classification of participants as competition loving, adverse, or neutral, or of preferences for competition proxied by participants' median switching row. However, probability weighting can introduce bias in the estimation of preferences for competition when measured as $\bar{\omega}_i$.

Since we do not elicit participants' probability weighting function, we rely on functions reported in the literature. Specifically, we calculate the adjusted preference ω'_{it} and the potential bias β_{it} using $w(p) = e^{-(-\ln p)^{0.65}}$, the function proposed by Prelec (1998) as a good fit for the experiments reported in Tversky and Kahneman (1992), Tversky and Fox (1995), and Wu and Gonzalez (1996). Figure E2 depicts the distribution of $\bar{\omega}'_i$, the median value of ω'_{it} across the five decision sets for both the initial experiment in Panel A and the follow-up experiment in Panel B. As in our previous analysis, we use participants with five consistent decision sets. Table E5 shows the mean and standard deviation of $\bar{\omega}'_i$) depending on gender and when using other criteria to include participants in the analysis.

Recall that, in each decision set, we calculate $\omega_{it} = \hat{p}_{it} - b_i$ as i's preference for competition in set t. If i distorts probabilities according to the probability weighting function $w(.)$, then

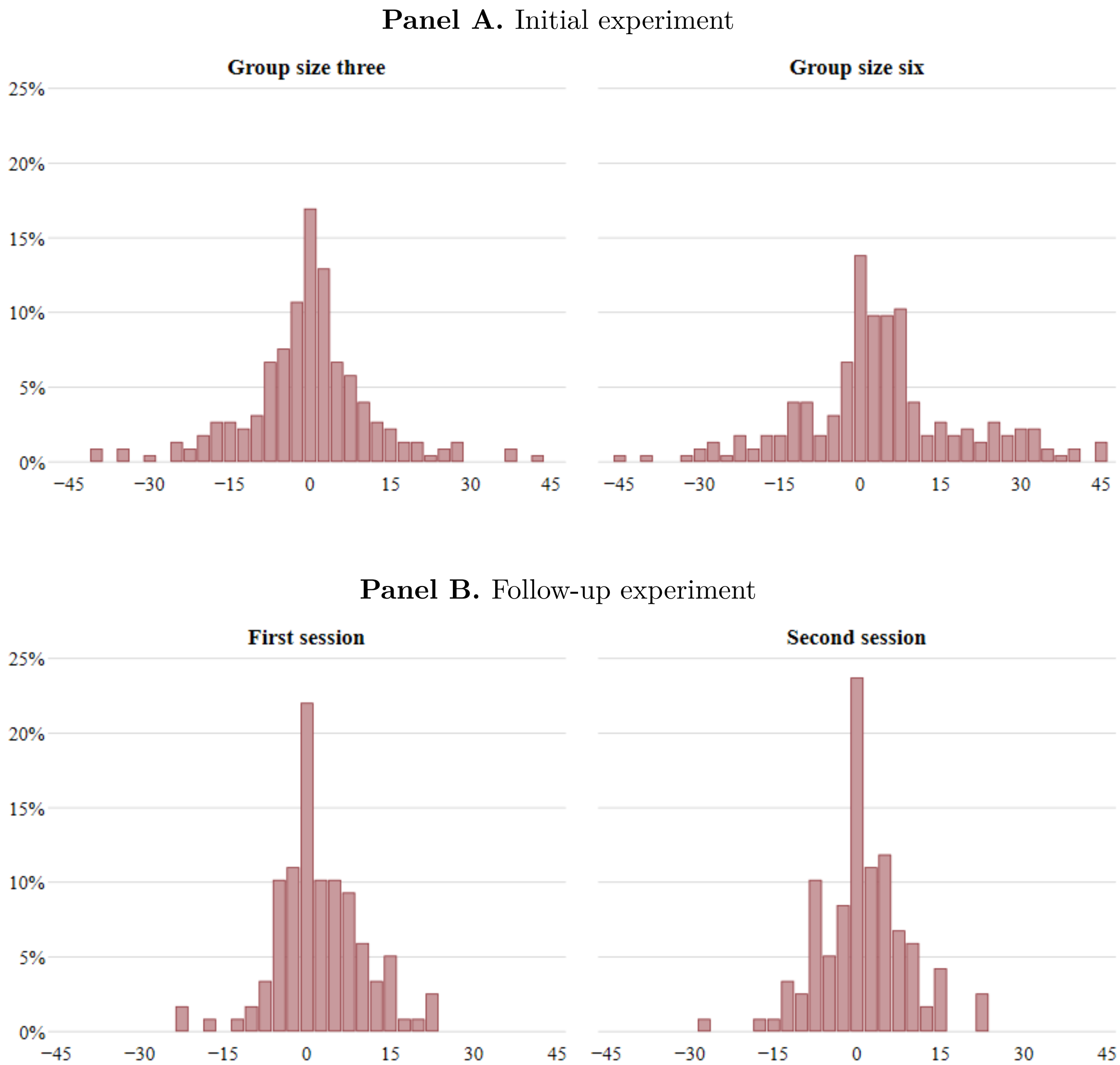


Figure E2. Distributions of participants' preferences for competition adjusted for probability weighting $\bar{\omega}'_i$ (initial and follow-up experiments)

Notes: The figure shows the distribution of $\bar{\omega}'_i$, the median value of ω'_{it} across decision sets, expressed as a percentage. $\omega'_{it} = w(\bar{p}_{it}) - w(b_i)$ represents the fraction of the utility gained from receiving the high instead of the low amount, $u_i(x_i \cdot \pi^H) - u_i(x_i \cdot \pi^L)$, participants are willing to forgo to either avoid or engage in competition adjusted with the probability weighting function $w(p) = \exp\left(-(-\ln p)^{0.65}\right)$. Panel A corresponds to the initial experiment and Panel B to the follow-up experiment.

the correct measure of i's preference would be ω'_{it} or $\omega'_{it} = w(\bar{p}_{it}) - w(b_i)$. Hence, probability weighting could introduce a bias in our measurements equal to $\beta_{it} = \omega'_{it} - \omega_{it}$. Note that the size and direction of this bias depends on how the probability weighting function distorts *differences* in probabilities. When the difference between the probability at the switching row and the participants' (expected) probability of being their groups' winner is small, we should not expect a large bias due to probability weighting. However, if this difference is large, the bias could be noticeable.

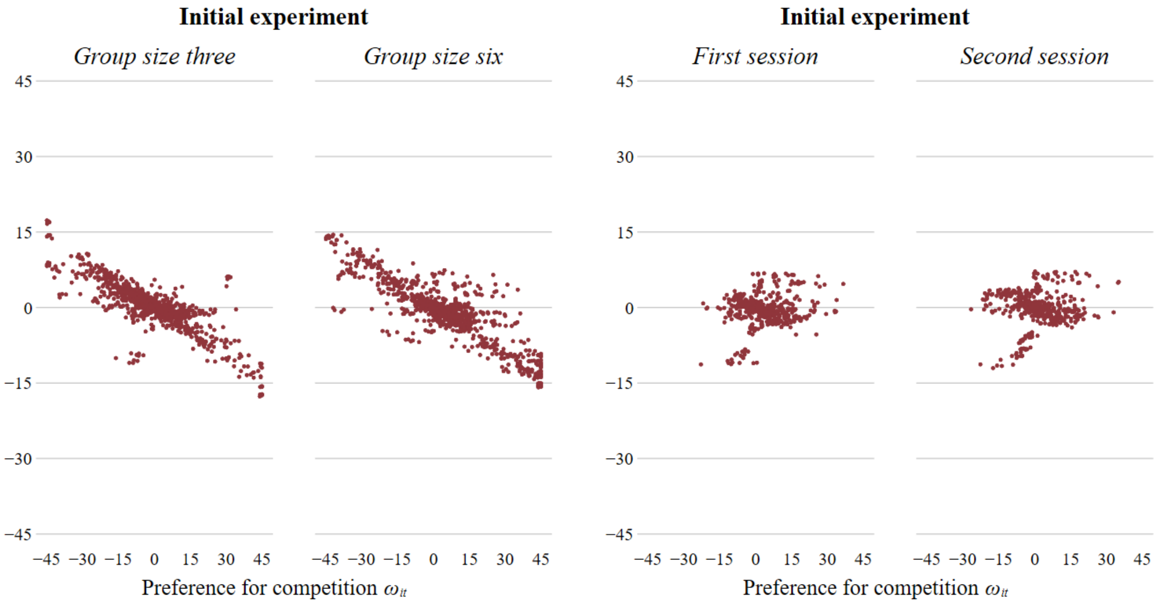


Figure E3. Potential bias of failing to account for probability weighting β_{it} depending on participants' observed preferences for competition ω_{it} (initial and follow-up experiments)

Notes: Scatter plots of the potential bias β_{it}—the difference between preferences for competition with and without probability weights $\omega'_{it} - \omega_{it}$—depending on participants' preferences for competition ω_{it}. Probabilities are adjusted with the probability weighting function $w(p) = \exp\big(-(-\ln p)^{0.65}\big)$.

By and large, the aggregate results using ω'_{it} look very similar to those using ω_{it}. As before, a majority of participants have non-zero values of $\bar{\omega}'_i$, indicating they forgo money to either engage in or avoid competition. For instance, the fraction of participants in the initial experiment with an $\bar{\omega}'_i$ greater than 5.0% in absolute value is 50.5% in groups of three and 63.1% in groups of six. In the follow-up experiment, these percentages are 53.3% in the first session and 51.4% in the second. Similarly, we reject the null hypothesis that the values of $\bar{\omega}'_i$ come from the same distribution in groups of three and six in the initial experiment (Wilcoxon signed-rank test, $p < 0.001$), but we do not reject this null hypothesis when comparing the values of $\bar{\omega}'_i$ across the first and second sessions of the follow-up experiment (Wilcoxon signed-rank test, $p = 0.164$). The test-retest correlation coefficient for $\bar{\omega}'_i$ equals 0.41, which is very similar to the 0.38 of $\bar{\omega}_i$. Lastly, we also find there are no gender differences in the values of $\bar{\omega}'_i$ (Mann-Whitney U tests, $p = 0.592$ in the initial experiment and $p = 0.926$ in the follow-up experiment).

The most noticeable difference between the distributions of $\bar{\omega}_i$ and $\bar{\omega}'_i$ is that the latter has fewer values at the tails of the distributions, particularly for the initial experiment (cf. Figures 2, D2, and E2). This is seen more clearly in Figure E3, which plots the bias β_{it} against the values of ω_{it}. As seen in the figure, for the initial experiment, the bias tends to be positive when $\omega_{it} < 0$, suggesting overestimation of i's aversion to competition, and negative when

$\omega_{it} > 0$, suggesting overestimation of i's enjoyment of competing. We do not see this pattern for the follow-up experiment. The most likely reason is the magnitude of the probabilities in the decision sets, which were considerably higher in the initial experiment due to participants' overconfidence.

Appendix F. Instructions for the initial experiment

Below are the instructions for the *before* treatment with first a group size of three followed by a group size of six. Instructions for the other treatments are very similar and available upon request.

General Instructions

Welcome to the experiment. In the experiment today, you will be asked to complete five tasks. Before each task, you will receive detailed instructions and description of how your earnings in that task are determined.

Your total earnings at the end of the experiment are the sum of the following two components:

1. A €10 show-up fee.
2. Your earnings in one of the five tasks. Specifically, at the end of the experiment, one of the five tasks you will complete during the experiment will be randomly chosen for payment purposes.

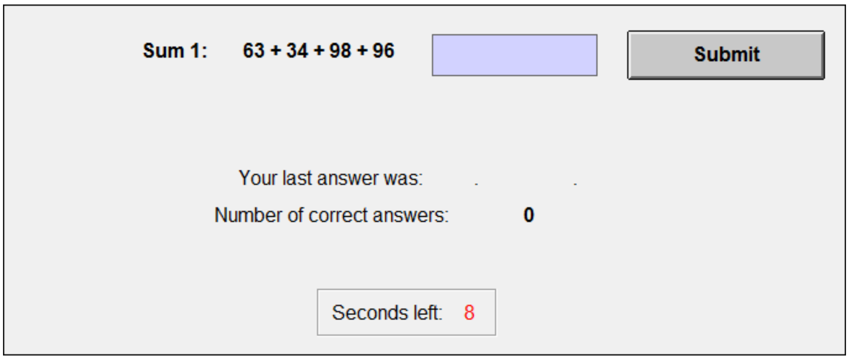
During the experiment, the use of cell phones is prohibited. All your information, decisions, and performance during this experiment are anonymous.

If you have a question, please raise your hand. An experimenter will come and answer your question in private. Now you will start Task 1, please read the instructions of Task 1 carefully.

Task 1

In Task 1, you will be randomly assigned to a group of three participants. In other words, you will be matched with two other participants in the room.

In Task 1 you will be given four minutes to calculate a series of sums of four two-digit numbers (see the screenshot below). You cannot use a calculator, but you are welcome to use the provided scratch paper. You submit an answer by clicking the button "Next". When you submit an answer, the computer will immediately tell you whether the answer is correct or incorrect and a new sum is generated.



Your earnings in Task 1 depend on your number of correct sums. Specifically, you can earn either a high amount or a low amount per correct sum. The high amount will vary between €1.5 and €6 per correct sum, and the low amount will vary between €0 and €1. You will be given the precise values before you perform the task. Whether you are paid a high amount or a low amount depends on your choices. Before you perform the task, you will choose between *individual pay* and *competitive pay*. The two payments schemes are as follows:

- *Individual pay:* if you choose individual pay, whether you receive a high or low amount per correct sum depends on chance. With individual pay your earnings do not depend on the performance of others in your group.
- *Competitive pay:* if you choose competitive pay, whether you receive a high or low amount per correct sum depends on your performance and the performance of the other two members of your group. Specifically, you will be your group's winner if you solve more sums in Task 1 than all others in your group in Task 1. If there are ties, the winner will be randomly determined among the tied group members. If you are your group's winner, you will receive the high amount per correct sum. If you are not your group's winner, then you are one of the two losers in the group. If you are one of the group's losers, you will receive the low amount per correct sum.

Practice round: Before Task 1 starts, you will have two minutes to get familiar with the screen and to practice the calculation of series of sums of four two-digit numbers. Please notice that your answers in this practice round will not be considered for your earnings in this experiment. Once you are done reading, click on the "Next" button on your screen.

Task 2

In this task, you can earn money by answering the following question: emph"How likely do you think it is that you are the winner of your group in Task 1?" Your answer can go from 0 (meaning you are completely certain that you are not the winner of your group) to 100 (meaning you are completely certain that you are the winner of your group).

Your earnings in Task 2 can be either *€0 or €20.* The probability of earning €20 depends on two things:

1. The actual outcome (whether you are the winner or a loser in your group)
2. The likelihood you selected as the answer to the question above.

The closer the likelihood you choose is to your actual outcome in Task 1, the higher the probability you have of earning €20. This probability is based on the formulas you see in footnote 1. [Footnote 1 text: Probability of earning €20 if you are the winner = $1 - (1 -$ Your selected likelihood/100$)^2$. Probability of earning €20 if you are one of the losers = $1 -$ (Your selected likelihood/100$)^2$.] It is not necessary for you to understand precisely the formulas, but it's important that you know that these formulas have been designed so that *your expected earnings are higher the closer your answer is to your actual likelihood of being your group's winner.*

To help you to think about your likelihood of being your group's winner, it is useful to think *how your performance in Task 1 ranks compared to the performance of all participants.* The table provided in the next page displays this information. In the table you can see for each possible rank (from being on the top 0% to being on the top 100%) the likelihood that someone with that rank is the winner of a group of three. The numbers on the table are calculated based on you being randomly assigned to groups of three. For example, imagine that your performance in Task 1 puts you in the Top 10%. This means that you performed better than 90% of all participants in the study and you performed worse than around 10% of all participants in the study. Then for you to be the winner it must be the case that *all two* of the other members of your group have a worse rank than you. In other words,

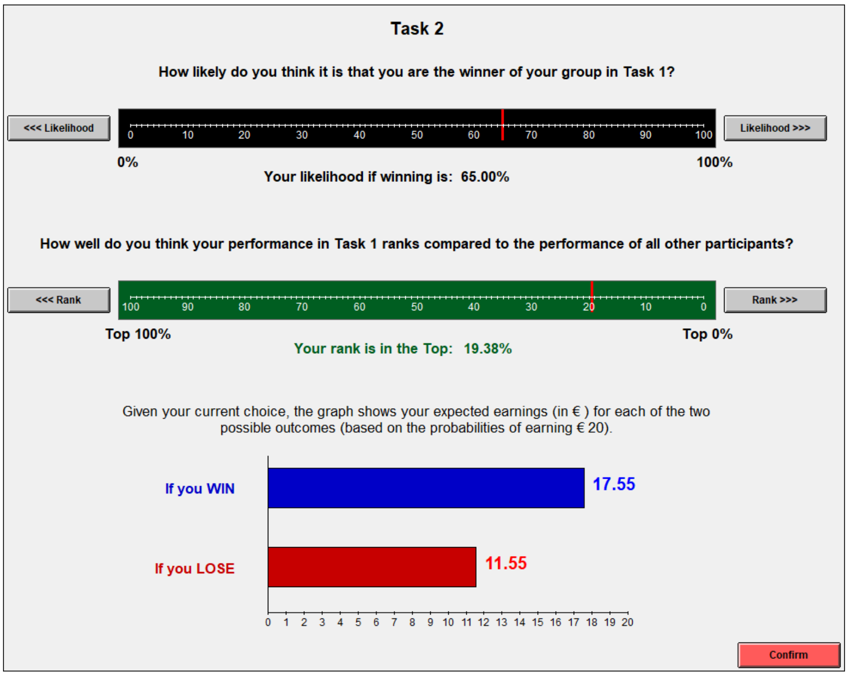
- You have been randomly matched *only* with participants who *all* come from the 90% of participants who performed worse than you, and
- You have *not* been randomly matched with *any* of the 10% of participants who performed better than you.

Group size 3

Your performance is in the Top ...	The likelihood that you are your group's winner is...	Your performance is in the Top ...	The likelihood that you are your group's winner is...
0%	100.00%	50%	25.00%
1%	98.01%	51%	24.01%
2%	96.04%	52%	23.04%
3%	94.09%	53%	22.09%
4%	92.16%	54%	21.16%
5%	90.25%	55%	20.25%
6%	88.36%	56%	19.36%
7%	86.49%	57%	18.49%
8%	84.64%	58%	17.64%
9%	82.81%	59%	16.81%
10%	81.00%	60%	16.00%
11%	79.21%	61%	15.21%
12%	77.44%	62%	14.44%
13%	75.69%	63%	13.69%
14%	73.96%	64%	12.96%
15%	72.25%	65%	12.25%
16%	70.56%	66%	11.56%
17%	68.89%	67%	10.89%
18%	67.24%	68%	10.24%
19%	65.61%	69%	9.61%
20%	64.00%	70%	9.00%
21%	62.41%	71%	8.41%
22%	60.84%	72%	7.84%
23%	59.29%	73%	7.29%
24%	57.76%	74%	6.76%
25%	56.25%	75%	6.25%
26%	54.76%	76%	5.76%
27%	53.29%	77%	5.29%
28%	51.84%	78%	4.84%
29%	50.41%	79%	4.41%
30%	49.00%	80%	4.00%
31%	47.61%	81%	3.61%
32%	46.24%	82%	3.24%
33%	44.89%	83%	2.89%
34%	43.56%	84%	2.56%
35%	42.25%	85%	2.25%
36%	40.96%	86%	1.96%
37%	39.69%	87%	1.69%
38%	38.44%	88%	1.44%
39%	37.21%	89%	1.21%
40%	36.00%	90%	1.00%
41%	34.81%	91%	0.81%
42%	33.64%	92%	0.64%
43%	32.49%	93%	0.49%
44%	31.36%	94%	0.36%
45%	30.25%	95%	0.25%
46%	29.16%	96%	0.16%
47%	28.09%	97%	0.09%
48%	27.04%	98%	0.04%
49%	26.01%	99%	0.01%
	continues →	100%	0.00%

The table shows that, for someone in the Top 10%, the likelihood that this happens is 81.00%.

You will indicate your likelihood of being your group's winner in a screen like the one below.



As you can see, there are two sliders in the top part of the screen. You can select your answer by moving the cursors in these two different sliders:

- *In the black slider*, you can select your likelihood of being the winner of your group. Your answer can go from 0% (meaning you are completely certain that you are not the winner of your group) to 100% (meaning you are completely certain that you are the winner of your group).
- *In the green slider*, you can select how your performance in task 1 ranks compared to the performance of all participants. Your answer can go from Top 100% (you performed worse than *all* other participants of the study) to Top 0% (you performed better than *all* other participants in the study).

Please notice that the information displayed in both sliders is always consistent with each other. In other words, when you select a likelihood on the black slider, the cursor on the green slider will automatically mark the rank associated with your selected likelihood. Similarly, when you select a rank on the green slider, the cursor on the black slider will automatically mark the likelihood associated with your selected rank. The values of the sliders are based on the numbers you can see in the table of the previous page. The cursors will appear on the sliders only after you have clicked on one of the sliders for the first time.

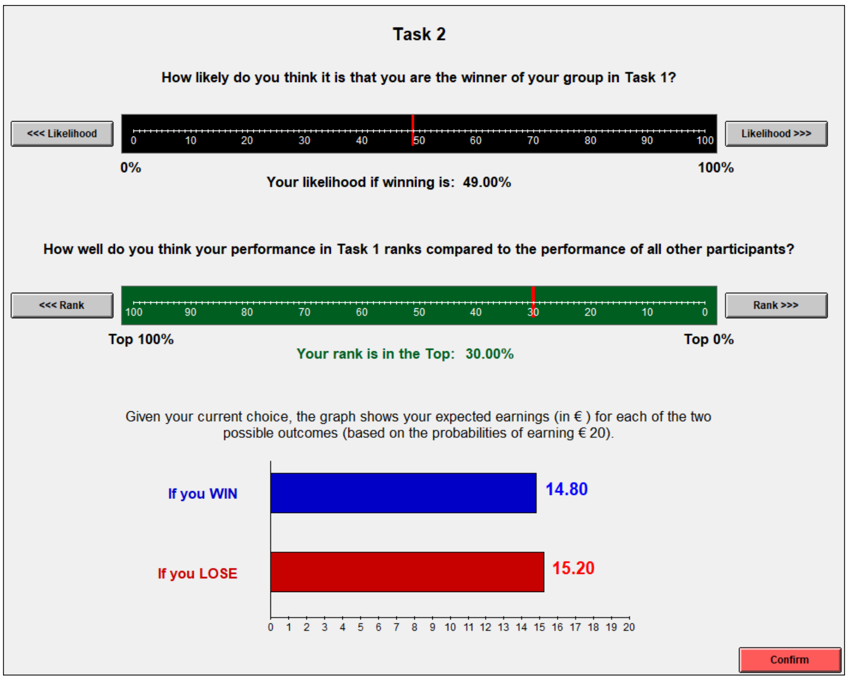
To help you to understand the consequences of your choice, below the sliders, you will also see the *expected earnings* associated to your choice in the two possible outcomes: in case *you are the winner*, and in case *you are one of the losers* of your group. You will obtain the highest expected earnings if *your answer equals the actual likelihood of you being the winner.*

Please remember that your earnings in Task 2 are either *€0 or €20*, therefore, your expected earnings are equal to €20 multiplied by the probability of earning the €20 (which is calculated with the formulas in footnote 1).

We provide an example below to illustrate how your earnings depend on your answers (note that the numbers used in this example are not indicative of what constitutes a good or bad answer in this task).

Example: Imagine that among the students taking part in this study, your performance in Task 1 puts you in the *Top 30%.* In other words, 70% of the study participants performed worse than you did and 30% performed better than you did. Recall that, for you to be the group's winner, it must be the case that *all two* of the other members of your group come from the 70% of participants who performed worse than you did. In this example, the probability that this occurs is *49.00%* (see the table).

Suppose that your answer is 49.00% in the black slider and Top 30% in the green slider, as shown in the screen below.

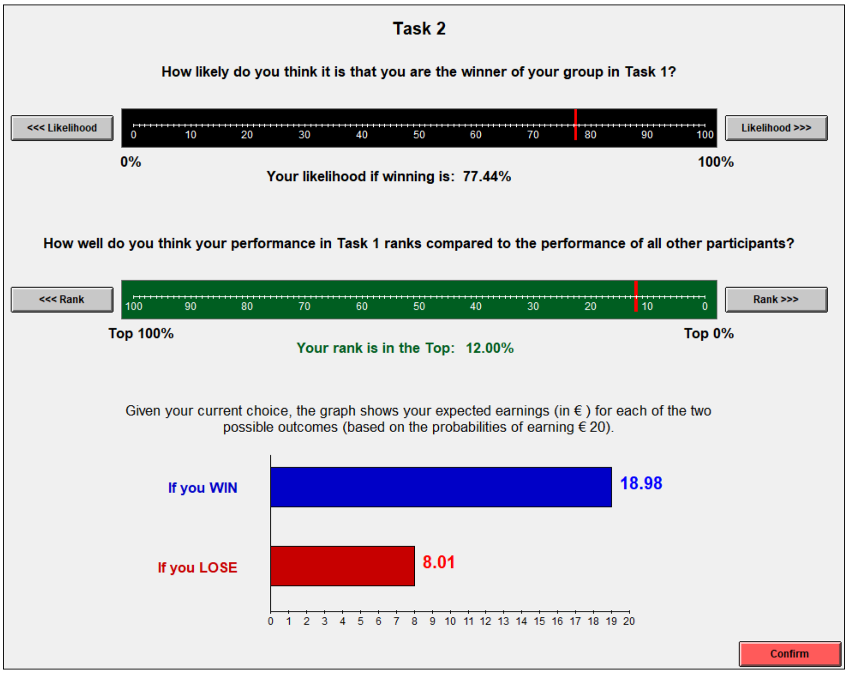


Then, as you can see with the bar graph in the screenshot:

- If you turn out to be the winner of your group, you can expect to earn in Task 2 *€14.80* on average (= €20 × probability of earning €20 if you are the winner).
- If you turn out to be one of the losers of your group, you can expect to earn in Task 2 *€15.20* on average (= €20 × probability of earning €20 if you are one of the losers).

Since the actual likelihood that you are the winner of your group is *49.00%*, this means that *49.00%* of the time you are the group's winner and *51.00%* of the time you are one of the losers. Overall, this means that you can expect to earn in Task 2 *€15.00 on average* (€15.00 = 0.49 × €14.80 + 0.51 × €15.20).

Now let's see what happens if you answer differently. Continue to suppose that your performance places you in the *Top 30%*. However, imagine that this time your answer is 77.44% in the black slider and Top 12% in the green slider, as shown in the screen below.



Then, as you can see with the bar graph in the screenshot:

- If you turn out to be the winner of your group, you can expect to earn in Task 2 *€18.98* on average (€20 × probability of earning €20 if you are the winner).
- If you turn out to be one of the losers of your group, you can expect to earn in Task 2 *€8.01* on average (€20 × probability of earning €20 if you are one of the losers).

Since the actual likelihood that you are the winner of your group is still *49.00%* (remember that you actually are in the Top 30%), this means you can expect to earn in Task 2 *€13.39 on average* (€13.39 = *0.49* × *€18.98* + *0.51* × *€8.01*). Note that *€13.39* is lower than *€15.00*, which are the expected earnings from reporting 49.00% in the black slider and Top 30% in the green slider.

In conclusion and to reiterate, you will obtain the highest expected earnings in Task 2 if *your answer equals your actual likelihood of being the group's winner in Task 1.* Once you are done reading, click on the "Next" button on your screen.

Your payment choice in task 1

Next you are going to perform Task 1, but before performing the task, you must choose how you want to be paid for each correct sum in Task 1. Recall that you can choose between *individual pay* and *competitive pay.*

You will be asked to make choices in *5 different decision sets.* All these decision sets are completely independent of each other. Each decision set consists of a table with a series of choices:

- The left choices correspond to *competitive pay.* Under competitive pay your earnings in Task 1 depend on your performance and the performance of others in your group. Specifically, if are the winner of your group then you earn the high amount per correct sum, otherwise you earn the low amount per correct sum.
- The right choices correspond to *individual pay.* Under individual pay your earnings in Task 1 depend on your performance and on chance. Specifically, you earn the high amount per correct sum with some probability X [a number between 1 and 100]. To determine your earnings, you will throw two ten-sided dice to randomly generate a number between 1 and 100. If the number you generate is lower than the probability X then you earn the high amount per correct sum, otherwise you earn the low amount per correct sum.

You must decide in every row whether you prefer *individual pay* or *competitive pay.* Note that in a decision set, the high and low amounts for competitive pay are the same in all rows. What varies from row to row is the probability of getting the high amount in individual pay. An example of one decision set is displayed below.

	Competitive Pay		Individual Pay
1.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **17%** chance and € 1.00 with **83%** chance
2.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **20%** chance and € 1.00 with **80%** chance
3.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **23%** chance and € 1.00 with **77%** chance
4.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **26%** chance and € 1.00 with **74%** chance
5.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **29%** chance and € 1.00 with **71%** chance
6.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **32%** chance and € 1.00 with **68%** chance
7.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **35%** chance and € 1.00 with **65%** chance
8.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **38%** chance and € 1.00 with **62%** chance
9.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **41%** chance and € 1.00 with **59%** chance
10.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **44%** chance and € 1.00 with **56%** chance

At the end of the experiment, one of the 5 decision sets will be randomly selected. Within the selected decision set, one of the 10 rows will be randomly selected. The type of payment you chose in the selected row will be used to determine how much you will receive per correct sum in Task 1.

Example: Take a look at the choices in the screenshot below.

	Competitive Pay		Individual Pay
1.	**€ 6.00 if you win and € 0.00 if you lose**	◉ ○	€ 6.00 with **21%** chance and € 0.00 with **79%** chance
2.	**€ 6.00 if you win and € 0.00 if you lose**	◉ ○	€ 6.00 with **23%** chance and € 0.00 with **77%** chance
3.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 25% chance and € 0.00 with 75% chance**
4.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 27% chance and € 0.00 with 73% chance**
5.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 29% chance and € 0.00 with 71% chance**
6.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 31% chance and € 0.00 with 69% chance**
7.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 33% chance and € 0.00 with 67% chance**
8.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 35% chance and € 0.00 with 65% chance**
9.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 37% chance and € 0.00 with 63% chance**
10.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 39% chance and € 0.00 with 61% chance**

Now, imagine that this decision set is randomly selected for payment and within this decision set, row number 4 is randomly selected. Given that *individual pay* was chosen instead of a *competitive pay* in this row, then:

- With 27% of chance, you will earn €6 per correct sum in Task 1 [the high amount].
- With 73% of chance, you will earn €0 per correct sum in Task 1 [the low amount].

Now, imagine that instead of row number 4, the row randomly selected for payment is row number 2. Given that *competitive pay* was chosen instead of *individual pay* in this row, then:

- If you are the group's winner in Task 1, you earn €6 per correct sum [the high amount].
- If you are one of the group's losers in Task 1, you earn €0 per correct sum [the low amount].

Task 3

In Task 3 you will perform again the same summation task you performed in Task 1. The main difference is that you will be randomly assigned to a *group of <u>six</u> participants* instead of three.

Task 4

Task 4 is like Task 2. In Task 4 you can earn money by answering the following question: *"How likely do you think it is that you are the winner of your group in Task 3?"* Again, your will be able to select your answer by moving the cursors in two different sliders:

- In the black slider, you can select your likelihood of being the winner of your group. Your answer can go from 0% (meaning you are completely certain that you are not the winner of your group) to 100% (meaning you are completely certain that you are the winner of your group).
- In the green slider, you can select how your performance in task 3 ranks compared to the performance of all participants. Your answer can go from Top 100% (you performed worse than *all* other participants of the study) to Top 0% (you performed better than *all* other participants in the study).

Your earnings in Task 4 will be calculated using the same formulas as in Task 2. Recall that *you will obtain the highest expected earnings if your answer equals the actual likelihood of you being the winner in Task 3.*

One important consideration for Task 4, is that to be the winner in Task 3, you need to be the best in a group of *six.* The table provided in the next page displays the likelihood of being your group's winner in Task 3 depending on each possible rank. Logically, it is harder to be the winner in a group of six than in a group of three. This is why the percentages in the table for Task 4 are lower than the percentages in the table for Task 2.

Your performance is in the Top ...	The likelihood that you are your group's winner is...
0%	100.00%
1%	95.10%
2%	90.39%
3%	85.87%
4%	81.54%
5%	77.38%
6%	73.39%
7%	69.57%
8%	65.91%
9%	62.40%
10%	59.05%
11%	55.84%
12%	52.77%
13%	49.84%
14%	47.04%
15%	44.37%
16%	41.82%
17%	39.39%
18%	37.07%
19%	34.87%
20%	32.77%
21%	30.77%
22%	28.87%
23%	27.07%
24%	25.36%
25%	23.73%
26%	22.19%
27%	20.73%
28%	19.35%
29%	18.04%
30%	16.81%
31%	15.64%
32%	14.54%
33%	13.50%
34%	12.52%
35%	11.60%
36%	10.74%
37%	9.92%
38%	9.16%
39%	8.45%
40%	7.78%
41%	7.15%
42%	6.56%
43%	6.02%
44%	5.51%
45%	5.03%
46%	4.59%
47%	4.18%
48%	3.80%
49%	3.45%
	continues →

Group of 6

Your performance is in the Top ...	The likelihood that you are your group's winner is...
50%	3.13%
51%	2.82%
52%	2.55%
53%	2.29%
54%	2.06%
55%	1.85%
56%	1.65%
57%	1.47%
58%	1.31%
59%	1.16%
60%	1.02%
61%	0.90%
62%	0.79%
63%	0.69%
64%	0.60%
65%	0.53%
66%	0.45%
67%	0.39%
68%	0.34%
69%	0.29%
70%	0.24%
71%	0.21%
72%	0.17%
73%	0.14%
74%	0.12%
75%	0.10%
76%	0.08%
77%	0.06%
78%	0.05%
79%	0.04%
80%	0.03%
81%	0.02%
82%	0.02%
83%	0.01%
84%	0.01%
85%	0.01%
86%	0.01%
87%	0.00%
88%	0.00%
89%	0.00%
90%	0.00%
91%	0.00%
92%	0.00%
93%	0.00%
94%	0.00%
95%	0.00%
96%	0.00%
97%	0.00%
98%	0.00%
99%	0.00%
100%	0.00%

Task 5

This is Task 5 of the experiment. The earnings from this part of the experiment are completely independent from the other tasks. The amount you earn depends *solely on your decisions and on chance*. Moreover, you will not perform further summation tasks.

You will be asked to make choices in *4 different decision tables*. All these decision tables are completely independent of each other. Each table has 10 different decisions, each in a different row. Each decision has two options:

- *Option A*, where you can earn a different certain amount in each of the 10 rows.
- *Option B*, where you can earn a high amount with some probability and a low amount with some other probability. Specifically, you earn the high amount with some probability X [a number between 1 and 100]. To determine your earnings, you will throw two ten-sided dice to randomly generate a number between 1 and 100. If the number you generate is lower than the probability X then you earn the high amount, otherwise you earn the low amount.

You can decide for every row whether you prefer *Option A* or *option B*. Option A is the same for every row, while option B takes 10 different amounts, one for each row. An example of one decision table is displayed below.

At the end of the experiment, one of the 4 decision tables will be randomly selected. Within the selected table, one of the 10 rows will be randomly selected. The choice you made in that row will determine your earnings of Task 5.

	Option A	Option B
1.	€ 28.80 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
2.	€ 30.24 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
3.	€ 31.68 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
4.	€ 33.12 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
5.	€ 34.56 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
6.	€ 36.00 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
7.	€ 37.44 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
8.	€ 38.88 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
9.	€ 40.32 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance
10.	€ 41.76 with certainty	€ 72.00 with 50% chance and € 0.00 with 50% chance

Appendix G. Instructions for the follow-up experiment

Below are the instructions for the first session of the follow-up experiment. Instructions for the second session are almost identical and available upon request.

General Instructions

Welcome to the first session of this experiment, which comprises *two separate sessions*. The second session is scheduled to take place in two/three work days from today. Your total earnings will be the sum of what you earn in each session, but the earnings from each session are entirely independent of one another.

Following today's session, you will receive an invitation to register for the second session. Note that attendance at both sessions is *mandatory to receive the total earnings*. Missing one session will result in a loss of your total earnings, and the show-up fee of €20 from both sessions.

To ensure the anonymity of your responses, you will be assigned an ID number at the beginning of today's session. *Please keep your ID number with you* as you will be asked to provide it both at the conclusion of today's session and during the second session. *Failure to do so will result in the loss of earnings for both experimental sessions.*

Today, you will receive instructions for the first experimental session, and you will receive instructions for the second session on that day.

Session #1 Instructions

In the experimental session of today, you will be asked to complete three tasks. Before each task, you will receive detailed instructions and a description of how your earnings in that task are determined.

Your total earnings of today's session are the sum of the following two components:

1. A 10 €show-up fee.
2. Your earnings in *one* of the three tasks. Specifically, at the end of the session, one of the three tasks will be randomly chosen for payment purposes.

During the experiment, the use of cell phones is prohibited. All your information, decisions, and performance during this experiment are anonymous.

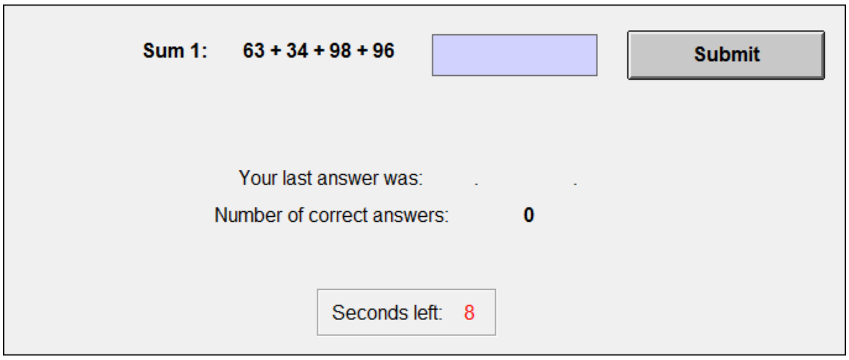
If you have a question, please raise your hand. An experimenter will come and answer your question in private.

Now you will start Task 1, please read the instructions of Task 1 carefully.

Task 1

In Task 1, you will be randomly assigned to a group of six participants. In other words, you will be matched with five other participants in the room.

In Task 1 you will be given four minutes to calculate a series of sums of four two-digit numbers (see the screenshot below). You cannot use a calculator, but you are welcome to use the provided scratch paper. You submit an answer by clicking the button "Next". When you submit an answer, the computer will immediately tell you whether the answer is correct or incorrect and a new sum is generated.



Your earnings in Task 1 depend on your number of correct sums and your choices. Before you perform the task, you will choose between *individual pay* and *competitive pay*. The two payment schemes are as follows:

- *Individual pay:* if you choose individual pay, you receive €1 per correct sum. With individual pay, your earnings do not depend on the performance of others in your group.
- *Competitive pay:* if you choose competitive pay, whether you receive a high or low amount per correct sum depends on your performance and the performance of the other two members of your group. Specifically, you will be your group's winner if you solve more sums in Task 1 than all others in your group in Task 1. If there are ties, the winner will be randomly determined among the tied group members. If you are your group's winner, you will receive the high amount of €3 per correct sum. If you are not your group's winner, then you are one of the two losers in the group. If you are one of the group's losers, you will receive the low amount of €0 per correct sum.

Practice round: Before Task 1 starts, you will have two minutes to get familiar with the screen and to practice the calculation of a series of sums of four two-digit numbers. Please notice that your answers in this practice round will not be considered for your earnings in this experiment. Once you are done reading, click on the "Next" button on your screen.

Task 2

In Task 2, you will be randomly assigned to a new group of three participants. In other words, you will be matched with two other participants in the room.

In Task 2, you are going to perform the same summation task you did in Task 1, but before performing the task, you choose how you want to be paid for each correct sum. You will choose again between *individual pay* and *competitive pay*. However, the way you select between *individual pay* and *competitive pay* differs from Task 1.

Your earnings in Task 2 depend on your number of correct sums. Specifically, you can earn either a high amount or a low amount per correct sum. The high amount will vary between €1.50 and €6 per correct sum, and the low amount will vary between €0.00 and €1.00. Whether you are paid a high amount or a low amount depends on your choices. You will be asked to make choices in *5 different decision sets*. All these decision sets are completely independent of each other. Each decision set consists of a table with a series of choices:

- The left choices correspond to *competitive pay*. Under competitive pay your earnings in Task 2 depend on your performance and the performance of others in your group. Specifically, if you are the winner of your group then you earn the high amount per correct sum, otherwise you earn the low amount per correct sum.
- The right choices correspond to *individual pay*. Under individual pay your earnings in Task 2 depend on your performance and on chance. Specifically, you earn the high amount per correct sum with some probability X [a number between 1 and 100]. To determine your earnings, you will throw two ten-sided dice to randomly generate a number between 1 and 100. If the number you generate is lower than the probability X then you earn the high amount per correct sum, otherwise you earn the low amount per correct sum.

You must decide in every row whether you prefer *individual pay* or *competitive pay*. Note that in a decision set, the high and low amounts for competitive pay are the same in all rows. What varies from row to row is the probability of getting the high amount in individual pay. An example of one decision set is displayed below.

	Competitive Pay		Individual Pay
1.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **17%** chance and € 1.00 with **83%** chance
2.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **20%** chance and € 1.00 with **80%** chance
3.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **23%** chance and € 1.00 with **77%** chance
4.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **26%** chance and € 1.00 with **74%** chance
5.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **29%** chance and € 1.00 with **71%** chance
6.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **32%** chance and € 1.00 with **68%** chance
7.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **35%** chance and € 1.00 with **65%** chance
8.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **38%** chance and € 1.00 with **62%** chance
9.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **41%** chance and € 1.00 with **59%** chance
10.	€ 4.00 if you win and € 1.00 if you lose	○ ○	€ 4.00 with **44%** chance and € 1.00 with **56%** chance

After making your choices, one of the 5 decision sets will be randomly selected. Within the selected decision set, one of the 10 rows will be randomly selected. The type of payment you chose in the selected row will be used to determine how much you will receive per correct sum in Task 2.

Importantly! Before you perform Task 2, you will be informed about the selected row, and the choice you made within that row. Hence, you will know whether your earnings are determined by *individual pay* or *competitive pay*, and the precise values of the high and low amounts.

Example: Take a look at the choices in the screenshot below.

	Competitive Pay		Individual Pay
1.	**€ 6.00 if you win and € 0.00 if you lose**	◉ ○	€ 6.00 with **21%** chance and € 0.00 with **79%** chance
2.	**€ 6.00 if you win and € 0.00 if you lose**	◉ ○	€ 6.00 with **23%** chance and € 0.00 with **77%** chance
3.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 25% chance and € 0.00 with 75% chance**
4.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 27% chance and € 0.00 with 73% chance**
5.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 29% chance and € 0.00 with 71% chance**
6.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 31% chance and € 0.00 with 69% chance**
7.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 33% chance and € 0.00 with 67% chance**
8.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 35% chance and € 0.00 with 65% chance**
9.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 37% chance and € 0.00 with 63% chance**
10.	€ 6.00 if you win and € 0.00 if you lose	○ ◉	**€ 6.00 with 39% chance and € 0.00 with 61% chance**

Now, imagine that this decision set is randomly selected for payment and within this decision set, row number 4 is randomly selected. Given that *individual pay* was chosen instead of a *competitive pay* in this row, then:

- With 27% of chance, you will earn €6 per correct sum in Task 2 [the high amount].
- With 73% of chance, you will earn €0 per correct sum in Task 2 [the low amount].

Now, imagine that instead of row number 4, the row randomly selected for payment is row number 2. Given that *competitive pay* was chosen instead of *individual pay* in this row, then:

- If you are the group's winner in Task 2, you earn €6 per correct sum [the high amount].
- If you are one of the group's losers in Task 2, you earn €0 per correct sum [the low amount].

What is your probability of winning the competition?

Before you choose between *individual pay* and *competitive pay* in Task 2, we will help you to think about *your chances of being your group's winner.* This assessment is based on your performance in Task 1. Specifically, the computer will use your and the other participants' performance in Task 1 and calculate the probability that you are the top performer in a new, randomly formed group of three participants.

Example: Imagine that you were the highest performer in your entire experimental session. This means that in any group of three, you will always be the top performer in the group. In this case, the computer will show your probability of winning as 100%.

Now, imagine that you were the second-highest performer in your experimental session. In this case, if the highest performer is in your new group, you don't win. If the highest performer is not in your group, then you are your group's top performer, and you win. Therefore, the probability of winning that the computer will show you is your likelihood of being randomly placed in a group without the highest performer when the groups are formed at random.

Task 3

This is Task 3 in today's experimental session. The earnings from this part are completely independent of the other tasks. The amount you earn depends *solely on your decisions and on chance.* Moreover, you will not perform further summation tasks.

You will be asked to make choices in *4 different decision tables.* All these decision tables are completely independent of each other. Each table has 10 different decisions, each in a different row. Each decision has two options:

- *Option A*, where you can earn a different certain amount in each of the 10 rows.
- *Option B*, where you can earn a high amount with some probability and a low amount with some other probability. Specifically, you earn the high amount with some probability X [a

number between 1 and 100]. To determine your earnings, you will throw two ten-sided dice to randomly generate a number between 1 and 100. If the number you generate is lower than the probability X then you earn the high amount, otherwise you earn the low amount.

You can decide for every row whether you prefer *Option A* or *option B*. Option A is the same for every row, while option B takes 10 different amounts, one for each row. An example of one decision table is displayed below.

After making your choices, one of the 4 decision tables will be randomly selected. Within the selected table, one of the 10 rows will be randomly selected. The choice you made in that row will determine your earnings for Task 3.

	Option A		Option B
1.	€ 28.80 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
2.	€ 30.24 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
3.	€ 31.68 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
4.	€ 33.12 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
5.	€ 34.56 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
6.	€ 36.00 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
7.	€ 37.44 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
8.	€ 38.88 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
9.	€ 40.32 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance
10.	€ 41.76 with certainty	○ ○	€ 72.00 with 50% chance and € 0.00 with 50% chance