1

Chapter 2: Effects of preparation on gender differences in willingness to compete

1.1 Introduction

1.1.1 Gender gaps persist despite women's success in education

Women have surpassed men in education outcomes, like college attendance and graduation rates [**Blau2017**, **Goldin2006**, **Stoet2015**], but continue to be underrepresented in top management positions in nearly all sectors [**Bertrand2001**]. And, a sizable gender wage gap still persists [**Blau2017**]. Traditional economic variables, such as household division of labor and discrimination, account for some, but not all, of these disparities [**Blau2017**]. As such, researchers have begun to consider psychological gender differences, including the predilection for competition, as means of understanding persistent gender gaps in labor market outcomes [**Niederle2011**]. Specifically, they expected that gender differences in competitiveness may explain why women may not compete for promotions and are underrepresented in top-level or highly competitive positions (e.g., CEO) [**Niederle2007**].

1.1.2 Gender differences in competitiveness

Research suggests women are, on average, less competitive than men [**Niederle2011**, **Croson2009**, **Niederle2017b**, **Shurchkov2018**], conditional on ability and other factors relevant to competitiveness. Foundational work on gender differences in competitiveness operationalized competitiveness as the choice of a tournament payment scheme, that reaps potentially higher earnings but requires outperforming an opponent, over a piece-rate scheme, where participants are paid per unit of work they produce [**Niederle2007**, **Saccardo2018**]. In this paradigm, women are less likely to enter tournaments while completing mathematical problems, even when they would have earned more by competing [**Niederle2007**]. Numerous conceptual replications over the past 15 years suggest that the gender difference in willingness to compete is robust [**Niederle2011**, **Niederle2017a**, **Niederle2017b**]. Notably, this effect has been replicated in diverse populations (e.g., across age groups and cultures) [**Apicella2015**, **Buser2014**, **Sutter2016**, **Andersen2013**, **Buser2017b**, **Sutter2010**, **Dreber2014**, **Mayr2012**] and with a diverse set of tasks [**Apicella2015**, **Saccardo2018**, **Bjorvatn2016**, **Sutter2015**, **Frick2011**, **Samek2019**]. However, there is evidence that the task used during competition affects the size of the gender gap. For instance, some research suggests that when the task is female-typed or gender-neutral, the gender gap in willingness to compete may be reduced or eliminated [**Iriberri2017**, **Boschini2014**, **Boschini2019**, **Apicella2015**, **Grosse2010**, **Gunther2010**, **Dreber2014**, **Dreber2011**, **Shurchkov2012**]. Drawing from the psychology literature on stereotype threat [**Steele1997**, **Spencer1999**, **Spencer2016**], negative stereotypes about women's ability to perform male-typed tasks (e.g., math, mental rotation) may produce anxiety and undermine performance. As a result, women may decide not to engage in a competition because they either believe the stereotype or because the stereotype provokes enough anxiety to reduce performance [**Gunther2010**, **Grosse2010**, **Iriberri2017**, **Shurchkov2012**].

Importantly, this laboratory measure of competitiveness based on the preference between a tournament and piece-rate payment scheme predicts labor market outcomes, including education choices [**Buser2014**, **Zhang2012**, **Buser2017c**, **Buser2017b**], entrepreneurial decisions (e.g., investment, employment) [**Berge2015**], and both expected and actual earnings [**Reuben2015**, **Reuben2017**]. Thus, understanding why men and women differ in levels of competitiveness and whether interventions exist that can reduce or eliminate the difference may be key for solving the pernicious gender gaps in the labor market. In fact, many of the most crucial time points during one's career that determine later success (e.g., job interviews, university and post-baccalaureate entrance exams, promotion) require competing against others for a limited number of slots and substantially higher earnings.

1.1.3 Risk attitude and confidence as proposed mechanisms for gender difference in competitiveness

Both confidence and risk attitudes have been implicated in driving gender differences in willingness to compete [**Niederle2011**, **Veldhuizen2017**]. However, the extent to which confidence and risk attitudes account for the gender difference in willingness to compete is debated. The foundational research in this literature suggests that confidence and risk attitudes do not completely explain gender differences in competitiveness, since a residual gap remained after controlling for these factors [**Niederle2007**]. As a result, the unexplained component of the original gender effect was taken as evidence of a distinct "competitiveness" trait, separate from risk attitudes and confidence [**Niederle2007**, **Niederle2011**]. Conversely, recent work correcting for measurement error [**Gillen2019**] and using experimental techniques to isolate the effects of the competitiveness trait [**Veldhuizen2017**] suggests that risk attitudes and confidence can fully explain the gender gap in the choice to compete.

Regardless of whether competitiveness is a "stand-alone" trait, it is clear that both confidence and risk attitudes influence how men and women react to competitions.

For instance, even in the original study by **Niederle2007**, 27% of the gender gap in tournament entry was explained by men being more overconfident than women about their relative performance on the task. As such, interventions designed to correct men's overconfidence, increase women's confidence, or decrease women's perceptions of risk and uncertainty in competitive contexts may help reduce the gender gap in competitiveness.

Gender differences in confidence

Within the literature on the gender gap in competitiveness, confidence is operationalized as the belief about one's relative performance during a competition, where individuals who have inaccurately high (low) ratings of their performance are deemed overconfident (underconfident) [**Niederle2011**, **Moore2008**, **Moore2017**]. If an individual does not believe their performance is higher than the individuals they are competing against, they are unlikely to make the decision to compete for fear of losing.

While most individuals are overconfident [**Alicke2013**, **Dunning2004b**], there is ample research to suggest that women are less (over)confident on average than men across a number of domains [**Mobius2011**, **Niederle2011**, **Croson2009**, **Lundeberg1994**, **Niederle2007**, **Bertrand2010a**, **Beyer1990**, **Beyer1997**, **Jakobsson2013**, **Bandiera2022**]. Because women are less overconfident, they compete less often than they should, given their actual ability [**Niederle2007**]. Confidence too may help explain why, in some situations, the gender gap in competitiveness may be reduced or eliminated. For instance, women tend to compete more when tasks are female-typed or gender-neutral [**Iriberri2017**, **Boschini2014**, **Boschini2019**, **Apicella2015**, **Grosse2010**, **Gunther2010**, **Dreber2014**, **Dreber2011**, **Shurchkov2012**], when they are facing other women [**DattaGupta2013**, **Booth2012**], or when competing against themselves [**Apicella2017a**, **Bonte2018**, **Carpenter2018**, **Apicella2020**]. For example, **Apicella2017a** document a gender difference in confidence when women and men are competing against other individuals, but not when they are competing against themselves (i.e., their own past performance). There are several non-mutually exclusive and potentially interacting explanations that could account for women's relatively lower (over)confidence, including differences in performance or ability, experience, innate psychological differences, and stereotype threat [**Steele1997**, **Spencer1999**, **Spencer2016**]. In the latter case, for instance, women may decide to forgo competitions because they either believe negative stereotypes about their ability to perform certain tasks, or because stereotypes provoke enough anxiety to reduce confidence [**Gunther2010**, **Grosse2010**, **Iriberri2017**, **Shurchkov2012**, **Burow2017**]. Taken together, this body of research suggests that interventions designed to increase confidence in women may embolden them to compete more.

Gender differences in risk attitude

A second variable that has been identified as a possible explanation for gender differences in competitiveness is risk attitudes, typically construed as the preference for a certain gain over a gamble, even if the gamble has an equal or greater monetary expectation [**Kahneman1982**]. Researchers investigating gender differences in risk attitudes find that men are typically more risk-seeking than women [**Eckel2008**, **Charness2012**, **Croson2009**, **Bertrand2010a**], including in hunter-gatherers [**Apicella2017**], but see [**Harrison2007**] for an exception. While most studies report a gender difference, the difference appears to be small to medium [**Filippin2016**] and culturally-dependent [**Gneezy2009**].

Competitive payment schemes are inherently riskier than piece-rate payment schemes because the variance in returns is greater. With piece-rate payment schemes, individuals are guaranteed a certain amount for every unit they produce. Moreover, there typically exists uncertainty in competitions since one's relative performance is unknown [**Niederle2011**]. Indeed, some of the gender gap in competitiveness is explained by men and women's differing risk attitudes [**Niederle2011**]. In fact, some recent work suggests that nearly 30% of the gender gap in competitive choices can be explained by risk attitudes [**Gillen2019**, **Veldhuizen2017**].

1.1.4 Why preparation may eliminate gender gap in competitiveness by changing confidence and/or risk attitude

Preparation or training on a task may increase one's confidence [**Gist1992**, **Schunk1981**, **Schunk1982**, **Usher2008**], since people usually are able to observe an improvement in their performance over time. In fact, there is evidence that preparation (as measured through studying) does causally affect academic performance among university students [**Stinebrickner2008**]. In line with increased confidence after preparation or training, **Lent1996** found that college students listed past accomplishments as the most influential factor in determining their confidence. Moreover, research directly comparing the effects of mastery experiences (via preparation), vicarious experiences (e.g., watching others perform a task), and a control treatment without any intervention on confidence found that mastery increased confidence significantly more than vicarious experiences and the control treatment [**Bandura1977a**]. Other research suggests that women are less likely to be confident in their answers when asked about topics outside of their field of expertise, leading to a confidence gap [**Sarsons2021**]. This suggests that gaining expertise, perhaps through practicing, may selectively boost women's confidence. **Roll2011** found that practicing mathematics problems, using an intelligent tutoring system, significantly decreases underconfidence but has no effect on rates of overconfidence. This too suggests that practicing may preferentially benefit women who are more likely than men to be underconfident. That said, if practicing only helps with underconfidence, when most people, including women, are overconfident, then its application may be limited.

Preparation, and the feelings of preparedness or self-efficacy that follow, may also decrease the perceived riskiness of competitions. With increased self-efficacy, individuals may believe their probability of winning in a competition is higher than those with lower self-efficacy. Surprisingly little work has explored how preparation impacts men's and women's risk attitudes. However, some experimental work suggests that manipulating perceived competence on a task by giving participants positive feedback about their performance on a task can lead to significantly more risk-taking behavior [**Krueger1994**]. The researchers were able to rule out the role of mood in driving the results by giving some participants positive feedback on one task and negative feedback on another. For these participants, risk-taking increased in the positive feedback condition and decreased in the negative feedback condition. Interestingly, **Gysler2002** find that knowledge – in this case, understanding of financial markets – and confidence in that knowledge, negatively correlate with women's risk aversion, but positively correlate with men's risk aversion. This suggests that preparation may disproportionately increase risk-taking in women. Finally, there is evidence that risk attitudes play a greater role in predicting decisions to compete when individuals are competing against other individuals, rather than themselves (i.e., their own past performance), possibly because there is more uncertainty in estimating an opponent's ability versus one's own ability [**Apicella2017a**].

1.1.5 Previous literature looking at preparation on competition entry

Surprisingly, little work has explored how preparation impacts men and women's confidence, risk attitudes, or their willingness to compete. We know of only one study that has explored how delaying a competition, and in some cases, offering an opportunity to study, affects decisions to compete. In a working paper, **Charness2021** examine whether gender gap entry rates change when a future opportunity to study for the task is made available. The authors hypothesized that women would be more likely to compete when there is an opportunity to study for the task. Contrary to their prediction, the authors found that providing an option to study leads men to be more likely to enter into future planned tournaments and leads women to be less likely to enter into these tournaments, resulting in a significant gender gap. However, this gap was only present during the initial, provisional sign-up period. When the actual choice was made later – sometime between one and five days – the gender difference disappeared. Of those men who returned to complete the study, some switched into the non-competitive payment scheme. The authors suggest that the results may be explained by men being overly confident in their future selves' resolve to study.

1.1.6 The current experiments

We examine the role of preparation on the gender differences in willingness to compete through three experiments. However, unlike **Charness2021**, we do not introduce a significant time delay in our studies. That is, experiments took place in a single session, thus minimizing any potential for gender differences in beliefs about future behavior to affect our results.

In the first experiment, we test whether simply knowing that there will be an opportunity to prepare before performing a task affects the gender gap in willingness to compete. That is, we manipulate participants' knowledge of whether they will have time to prepare before they make their decision to compete. We anticipated that participants with this information would be more inclined to compete compared to participants without this information and that this effect would be stronger for women. Thus, we expected a main effect of condition and an interaction between gender and condition on the choice to compete. In the second experiment, we examined how actual preparation influences the decision to compete. That is, we manipulated whether participants were required to prepare before making the decision to compete. Again, we expected that women in the preparation condition would be more inclined to compete than women in the no preparation condition, while we did not expect men's competitiveness to be significantly affected by the preparation manipulation.

Finally, in experiment 3, we examine how an unlimited amount of preparation affects gender differences in the willingness to compete. Across all experiments, we measured gender differences in actual preparation after administering the treatment and eliciting preferences to compete. Finally, we monetarily incentivized participants in both studies to correctly predict whether men or women would prepare and compete more. The research design, pre-registered hypotheses, measures and analyses are available on OSF and all analyses were conducted in R statistical software (version 4.0.4). All participants across the three experiments were recruited through Amazon Mechanical Turk (or MTurk), with a guaranteed payment and the opportunity to earn bonuses depending on their performance and the performance of others. Recruiting participants on this platform allowed for efficient data collection while meeting acceptable psychometric standards, such as high test-retest and alpha reliability [**Rand2012**, **Buhrmester2011**, **Paolacci2014**, **Chandler2016**].

1.2 Study 1

1.2.1 Methods

Participants

All study measures described below are publicly available on OSF both as a .pdf and .qsf. Participants on Amazon Mechanical Turk who opted into the study had to pass several screening questions. Specifically, participants included in the paid portion of the study had to (i) identify their nationality as American and live in the United States, (ii) identify as a man or a woman, and (iii) be using a computer (rather than a phone or tablet). If they did not meet these criteria, they did not proceed to the paid portion of the study. Additionally, upon reviewing the data, we had reason to suspect that some participants completed the study more than once. Specifically, some participants had the same IP address, MTurk ID, and were of the same gender. When entries matched on all three identifiers, we included only the first entry and excluded all subsequent entries. The final sample consisted of 1056 participants (53.6% women), with an average age of 37.74 ($SD = 13.19$) years.[1] 54 participants (53.7% women) dropped out of the study before finishing and we use their data when available. See Table 1.1 for all demographic information.

[1]Note: Per our screening criteria for participants, the number of participants who do not identify as men or women is 0 across all studies. Therefore, the proportion of men in the sample is represented by whatever proportion of the full sample does not identify as women

Demographics Across Conditions

Characteristic	Control condition, N = 522	Preparation condition, N = 522	**p-value**
Age	38 (13)	38 (14)	>0.9
Unknown	24	17	
Gender			>0.9
Man	243 (47%)	244 (47%)	
Woman	279 (53%)	278 (53%)	
Race/ethnicity			0.3
Black/African-American	56 (11%)	49 (9.7%)	
East Asian	41 (8.2%)	34 (6.7%)	
Hispanic/Latino	33 (6.6%)	23 (4.6%)	
Middle Eastern	0 (0%)	1 (0.2%)	
Native American	1 (0.2%)	4 (0.8%)	
Pacific Islander	0 (0%)	2 (0.4%)	
Selected more than one option	14 (2.8%)	21 (4.2%)	
Selected other	5 (1.0%)	8 (1.6%)	
White/Caucasian	348 (70%)	363 (72%)	
Unknown	24	17	
Household Income			0.4
Less than $10,000	73 (15%)	69 (14%)	
$10,000 to $20,000	75 (15%)	55 (11%)	
$20,000 to $30,000	71 (14%)	69 (14%)	
$30,000 to $40,000	66 (13%)	71 (14%)	
$40,000 to $50,000	51 (10%)	60 (12%)	
$50,000 to $60,000	48 (9.6%)	51 (10%)	
$60,000 to $70,000	28 (5.6%)	34 (6.7%)	
$70,000 to $80,000	21 (4.2%)	29 (5.7%)	
$80,000 to $90,000	11 (2.2%)	15 (3.0%)	
$90,000 to $100,000	15 (3.0%)	12 (2.4%)	
$100,000 to $200,00	39 (7.8%)	35 (6.9%)	
Over $200,000	0 (0%)	5 (1.0%)	
Unknown	24	17	
Education			0.2
Less than a high school degree	3 (1.0%)	3 (1.0%)	
High School Diploma	65 (21%)	55 (18%)	
Vocational Training	9 (3.0%)	17 (5.5%)	
Some College	167 (55%)	155 (50%)	
Bachelor's degree	0 (0%)	0 (0%)	
Graduate Degree	61 (20%)	78 (25%)	
Unknown	217	214	

Table 1.1: Size of sample in Study 1 with corresponding percentage listed for gender, race, education, and household income, with p-values derived from Fisher's exact test. Mean with corresponding standard deviation listed for age, with p-values derived from Kruskal-Wallis test. If a participant did not respond to a given question, we list their response as 'Unknown'.

Demographics Based on Participant Gender

Characteristic	**Man**, N = 490	**Woman**, N = 566	**p-value**
Age	36 (13)	39 (13)	<0.001
Unknown	25	28	
Race/ethnicity			0.092
Black/African-American	43 (9.2%)	62 (12%)	
East Asian	42 (9.0%)	33 (6.1%)	
Hispanic/Latino	34 (7.3%)	22 (4.1%)	
Middle Eastern	1 (0.2%)	0 (0%)	
Native American	2 (0.4%)	3 (0.6%)	
Pacific Islander	1 (0.2%)	1 (0.2%)	
Selected more than one option	12 (2.6%)	23 (4.3%)	
Selected other	5 (1.1%)	8 (1.5%)	
White/Caucasian	325 (70%)	386 (72%)	
Unknown	25	28	
Household Income			0.8
Less than $10,000	63 (14%)	79 (15%)	
$10,000 to $20,000	60 (13%)	70 (13%)	
$20,000 to $30,000	60 (13%)	80 (15%)	
$30,000 to $40,000	62 (13%)	75 (14%)	
$40,000 to $50,000	52 (11%)	59 (11%)	
$50,000 to $60,000	49 (11%)	50 (9.3%)	
$60,000 to $70,000	30 (6.5%)	32 (5.9%)	
$70,000 to $80,000	30 (6.5%)	20 (3.7%)	
$80,000 to $90,000	10 (2.2%)	16 (3.0%)	
$90,000 to $100,000	12 (2.6%)	15 (2.8%)	
$100,000 to $200,00	34 (7.3%)	40 (7.4%)	
Over $200,000	3 (0.6%)	2 (0.4%)	
Unknown	25	28	
Education			0.9
Less than a high school degree	2 (0.7%)	4 (1.2%)	
High School Diploma	61 (21%)	59 (18%)	
Vocational Training	13 (4.5%)	13 (4.0%)	
Some College	153 (52%)	169 (53%)	
Bachelor's degree	0 (0%)	0 (0%)	
Graduate Degree	63 (22%)	76 (24%)	
Unknown	198	245	

Table 1.2: Size of sample in Study 1 with corresponding percentage listed for race, education, and household income, with p-values derived from Fisher's exact test. Mean with corresponding standard deviation listed for age, with p-values derived from Kruskal-Wallis test. If a participant did not respond to a given question, we list their response as 'Unknown'.

Procedures

Participants were told they would be completing a timed multiplication task where they could choose how they would be paid for their performance. We chose a multiplication task because we expected participants' performance to improve with practice. Indeed, research suggests that rehearsing and recalling associative memories can speed up retrieval of those memories [**Rundus1971**]. The task involved solving problems from multiplication tables 1-12 as quickly as possible within a two-minute period. They were provided an example of a question with the correct response and had to answer three practice problems correctly to proceed, as a test of their comprehension. After completing the comprehension questions, participants were randomly assigned to either a "knowledge of preparation" condition or a control condition. Participants in the "knowledge of preparation" condition were presented the following text:

"There is an option to practice/study before completing the multiplication task that is available to all participants. If you take this opportunity to practice/study, we will provide you with materials that may help boost your performance in the multiplication task. You will have unlimited time to practice/study before completing the task. You can stop practicing/studying at any point."

Participants assigned to the control condition simply proceeded without seeing this text. An equal number of participants were assigned to both conditions (control= 50%). Of the men who completed the study, 50.1% were randomly assigned to the control condition. Of the women who completed the study, 49.91% were randomly assigned to the control condition, $\chi^2(1, n = 1056) = 0.00$, $p > .999$.

Then, all participants were asked to choose how they wanted to be paid. They were given two options, either a piece-rate payment scheme or a tournament payment scheme. They first read a description of each payment scheme, and had to correctly answer three comprehension questions before making their selection.

Under the piece-rate scheme participants were told that they would be paid \$.10 for every problem answered correctly. Under the tournament scheme, participants were told that they would be paid \$.20 for every problem they answered correctly, but only if they answered more questions correctly than a randomly assigned competitor. The order of presentation of the tournament and piece-rate payment options was randomized for participants. Participants in the experimental condition were reminded that they had the option to prepare before completing the task. After choosing a payment scheme, participants in both conditions were given an opportunity to prepare before the multiplication task. If they chose not to prepare, they proceeded to the timed multiplication task. If they chose to prepare, participants were presented with each multiplication table, 1 through 12, in sequential order. Each multiplication table provided products of numbers up to 12. Thus, participants could use the tables to study. Additionally, participants were asked if they wanted to complete practice problems for each times table. If they said yes, participants were asked to solve all multiples in that table and could only proceed to the next table if they answered all the questions correctly. Once they completed all practice questions for a given times table, they were shown the multiplication table again and were asked if they would like to continue solving problems from that table or move onto the next multiplication table. This process was repeated for each multiplication table. We originally pre-registered that we would use time spent preparing as a secondary dependent variable of interest outside of the binary choice to prepare. However, upon reflection, we decided to use a different measure, "number of practice rounds completed" (M = 1.11, SD = 2.22), which was collected in the Qualtrics survey instead of the time preparing variable. This decision was made prior to conducting any analyses and was incorporated into the preregistration plans for Studies 2 and 3. We decided against using time spent preparing as a secondary dependent variable because of concerns that we were not able to monitor whether participants were actually practicing during that time or were doing other activities, making interpretation of any significant gender differences in the results difficult. For instance, gender differences may arise if women (or men) on MTurk are more likely to be interrupted by other family members while they are working on the computer in general, rather than because of gender differences in desire to prepare. With the number of practice rounds completed variable, participants had to make the conscious decision to continue preparing by clicking a "Yes" button every time they wanted to complete a new round of preparation. This variable was encoded as follows: if participants chose not to practice at all, they were assigned a "0" whereas participants who said they wanted to practice were assigned a value of at least "1", and this number increased incrementally by one for each additional round of practice completed.

Overall, we had two measures of preparation behavior: 1) binary choice to practice, and 2) number of practice rounds completed. The decision to practice measure conceptually captures a participants' baseline willingness to prepare, before they know what the preparation will involve. The number of practice rounds serves as a way to quantify the number of times participants continue to practice after having made the initial decision to prepare, and having seen what the practice rounds look like, which we imagine may reflect different underlying decision processes.

Following the preparation portion of the study, participants moved on to the paid portion of the study. They were required to solve as many problems as possible in two minutes. Participants' scores on the task were quantified as the number of questions correct within the two-minute time frame allotted, without any penalties for incorrect responses. After completion, participants were told how many problems they answered correctly. We do not include any information about their relative performance since we ask them to guess their relative performance in the confidence measure. Thus, participants following the tournament payment scheme were not told whether they won, since this serves as an indicator of relative performance. We employ this design across all studies in this dissertation.

Then, they completed a series of incentivized follow-up questions, including measures of confidence and perceptions of gender differences. For these measures, participants were told one of these questions would be selected for a possible bonus payment, and if they answered the selected question correctly, they would earn a bonus of $.10.

For the measure of confidence, participants were asked to correctly predict their relative performance compared to all other participants completing the task by indicating the decile of their score. We used a measure of relative performance, rather than a measure of absolute performance (e.g., asking participants to guess their score on the task) because perceptions of relative performance will likely be predictive of the choice to compete given competition inherently requires a comparison of one's performance to the performance of one's competitors. The confidence measure draws from previous research [**Niederle2007**], but instead of asking participants to indicate whether they won against a randomly selected opponent, we asked them to guess their relative decile to provide us with more information about their relative confidence. Given the difficulty of guessing one's exact percentile without any information about other participants, deciles are used rather than percentiles to make earning the bonus seem more achievable. Also, the item was phrased so participants did not need to understand the word "decile," but were asked instead: "If my performance is compared to that of all participants that completed the task, I think my score was. . . " with the options for responses ranging from "Better than all other participants" to "Better than none of the other participants" with 10% increments in between (e.g., "Better than 50% of participants"). Since task-specific confidence measures tend to be better predictors of behavior than general measures of confidence [**Oney2015**], the confidence measure assesses participants' beliefs within the context of the multiplication task used.

Participants were also asked to correctly predict whether men or women 1) correctly solved more problems 2) spent more time practicing before completing the multiplication task, and 3) chose the tournament payment option more. An additional question about perceptions of general gender differences in willingness to prepare that was not incentivized was included after participants responded to the incentivized questions: "For most tasks, do you think men or women generally prepare (i.e., practice and/or study) more?"

Finally, participants completed a measure of risk attitudes, where they answered if they generally are willing to take risks or try to avoid taking risks [**Dohmen2011b**] on a 10-point scale with 0 meaning participants are "Not at all willing to take risks" and 10 indicating participants are "Very willing to take risks." There is evidence that risky behavior (i.e., lottery choices) is strongly associated with the risk measure used in this study [**Dohmen2011b**]. Additionally, risk attitude tends to be explained by one underlying trait, with a relatively smaller amount of variation in risk attitude explained by context (e.g., risk attitude during career, health, or financial decisions). Thus, across contexts, risk attitude is likely to be stable and predictive of behavior [**Dohmen2011b**]. To determine whether participants used calculators to improve their performance on the task and whether there were gender differences in the use of calculators, we also asked participants about their use of calculators and perceptions of calculator use on the multiplication task. Neither of these measures was incentivized.

1.2.2 Results

Describing main variables of interest

Contrary to previous data in this literature [**Niederle2007**], a minority of participants (15.41%) chose to compete. Despite the small proportion of participants who chose to compete, we still replicate the gender gap in the choice to compete when gender is included as the only predictor in the model, where a greater share of men (20.25%) compared to women (11.19%) chose to compete. A logistic regression revealed that this gender difference in the choice to compete is significant, $b = -0.70$, 95% CI $[-1.05, -0.36]$, $z = -3.95$, $p < .001$. However, when including control variables, such as risk attitudes, confidence, task scores, and the hypothesized interaction between gender and competition choice, we find that the effect of gender is no longer significant, $b = -0.42$, 95% CI $[-0.95, 0.10]$, $z = -1.56$, $p = .118$, while risk attitudes, $b = 0.31$, 95% CI $[0.23, 0.39]$, $z = 7.60$, $p < .001$ and task scores, $b = 0.02$, 95% CI $[0.01, 0.02]$, $z = 3.34$, $p = .001$, are significant, suggesting those variables may fully explain the observed gender difference in willingness to compete (see Table 1.3).

	(1)			(2)			(3)		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.25	0.20 – 0.32	**<0.001**	0.26	0.18 – 0.35	**<0.001**	0.01	0.01 – 0.03	**<0.001**
gender [Woman]	0.50	0.35 – 0.70	**<0.001**	0.48	0.29 – 0.78	**0.004**	0.66	0.38 – 1.11	0.118
condition [pract]				0.98	0.62 – 1.53	0.926	0.88	0.54 – 1.41	0.590
condition [pract] * gender [Woman]				1.06	0.53 – 2.14	0.861			
task score							1.02	1.01 – 1.03	**0.001**
risk							1.36	1.26 – 1.47	**<0.001**
conf rank							1.01	1.00 – 1.02	0.147
gender [Woman] * condition [pract]							1.11	0.53 – 2.32	0.780
Observations	1019			1019			1004		
R^2 Tjur	0.016			0.016			0.108		

Table 1.3: All models are logistic regressions with choice to compete as the dependent variable, where man and control are the reference categories for participant gender and preparation condition, respectively. The gender difference in the choice to compete is not reduced by preparation condition, but is explained by risk attitudes and task scores. p < .05 is considered significant and bolded.

In separate linear regressions with gender as the only predictor, we also observed gender differences in both risk attitudes, $b = -0.85$, 95% CI $[-1.16, -0.54]$, $t(1002) = -5.36$, $p < .001$, and confidence, $b = -8.25$, 95% CI $[-10.97, -5.54]$, $t(1002) = -5.97$, $p < .001$. When included as the sole predictor in a linear regression, we find that gender significantly predicts task scores on the paid multiplication task, $b = -7.31$, 95% CI $[-9.81, -4.81]$, $t(1005) = -5.73$, $p < .001$, such that women have lower scores on average, $M_{\text{women}} = 49.14$, $SD_{\text{women}} = 18.83$; $M_{\text{men}} = 56.44$, $SD_{\text{men}} = 21.63$. The effect of gender on task scores holds in a separate linear regression with other variables included as predictors in the model, $b = -6.41$, 95% CI $[-8.95, -3.87]$, $t(998) = -4.96$, $p < .001$ (see Table 1.5). See Table 1.4 for a summary of gender differences in the main variables of interest.

Variable	**Man**, N = 490[1]	**Woman**, N = 566[1]
task_score	53 (41, 71)	48 (35, 63)
comp_choice		
piecerate	378 (80%)	484 (89%)
tournament	96 (20%)	61 (11%)
pract_choice	205 (43%)	304 (56%)
risk	5.00 (3.00, 7.00)	4.00 (2.00, 6.00)
conf_rank	60 (50, 80)	50 (40, 70)

[1] Median (IQR); n (%)

Table 1.4: Gender differences in the main variables of interest, including: task scores, choice to compete, choice to practice, confidence, and risk attitudes. Medians are reported for task score, risk attitudes, and confidence, with IQRs in parentheses. For choice to practice and choice to compete, we report the number and percentage of participants that fall into each category for each respective gender.

	(1)			(2)			(3)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	56.44	54.61 – 58.28	**<0.001**	56.21	54.17 – 58.25	**<0.001**	43.06	39.00 – 47.12	**<0.001**
gender [Woman]	-7.31	-9.81 – -4.81	**<0.001**	-8.25	-10.97 – -5.53	**<0.001**	-6.41	-8.95 – -3.87	**<0.001**
comp choice [tournament]				1.16	-3.36 – 5.68	0.615	2.14	-2.07 – 6.35	0.319
gender [Woman] * comp choice [tournament]				9.46	2.43 – 16.50	**0.008**	8.05	1.62 – 14.48	**0.014**
risk							-1.51	-1.98 – -1.04	**<0.001**
conf rank							0.35	0.30 – 0.41	**<0.001**
Observations	1007			1007			1004		
R^2 / R^2 adjusted	0.032 / 0.031			0.046 / 0.043			0.203 / 0.199		

Table 1.5: All models are linear regressions with task score as the dependent variable, where man and piece-rate payment scheme are the reference categories for participant gender and competition choice, respectively. After controlling for risk attitudes, confidence, and competition choice, women still have lower scores on the multiplication task than men, p < .05 is considered significant and bolded.

Effects of knowledge of preparation condition on gender differences in choice to compete

Contrary to our predictions, we do not find evidence of a significant interaction between gender and condition on the decision to compete, $b = 0.06$, 95% CI [-0.63, 0.76], $z = 0.18$, $p = .861$ (see Figure 1.1). When included as a sole predictor of the choice to compete in a logistic regression, we did not find evidence that assignment to the knowledge of preparation affected the choice to compete, $b = 0.01$, 95% CI [-0.33, 0.35], $z = 0.05$, $p = .963$.

Gender differences in preparation

As hypothesized, a logistic regression with gender predicting the choice to practice shows that a greater proportion of women (55.88%) took advantage of the opportunity to practice relative to men (43.25%), $b = 0.51$, 95% CI [0.26, 0.76], $z = 4.01$, $p < .001$ (see right panel of Figure 1.2). Gender remains a significant predictor of the binary choice to prepare after adding participants' choice to compete and the interaction

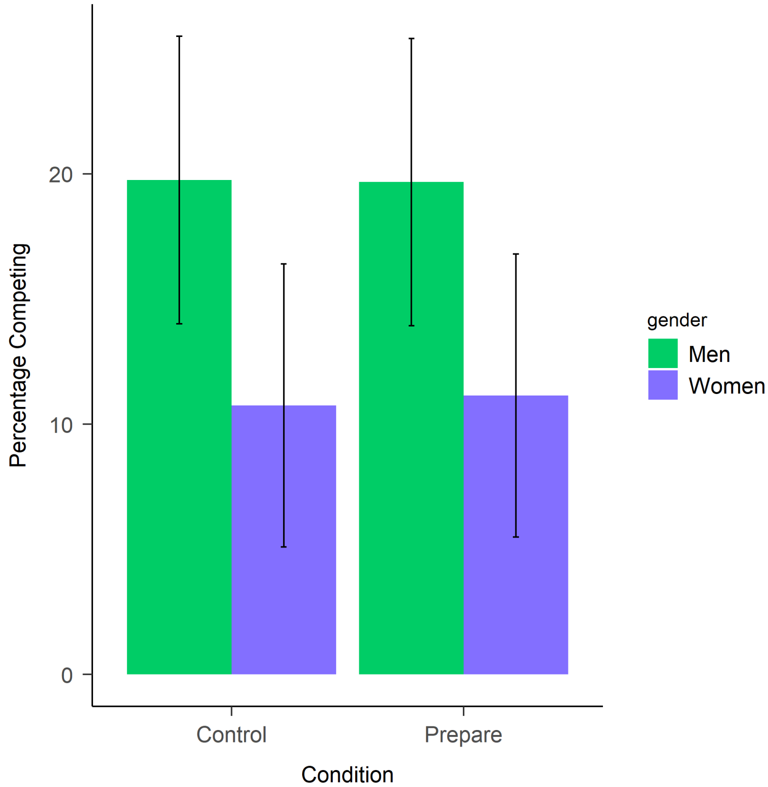

Figure 1.1: Proportion of men and women in Study 1 who chose to compete based on preparation condition. Knowledge of preparation did not reduce the gender difference in competitiveness. Error bars represent standard errors.

between gender and the choice to compete in the model, $b = 0.54$, 95% CI [0.27, 0.82], $z = 3.92$, $p < .001$, but we do not find an interaction between gender and the choice to compete, (see left panel of Figure 1.2). We also find that the choice to compete positively predicts a participants' likelihood of choosing to practice, $b = 0.50$, 95% CI [0.05, 0.95], $z = 2.18$, $p = .030$. In a subsequent logistic regression with additional possible predictors of the decision to practice, we find that the gender effect holds, $b = 0.55$, 95% CI [0.27, 0.84], $z = 3.79$, $p < .001$, suggesting that it is not explained by the observed gender differences in risk attitudes, confidence, nor task scores (see Table 1.6). We also ran a two-part hurdle model with gender, competition choice, and the interaction between those variables predicting the number of practice rounds variable with the assumption that there may be different decision-making processes

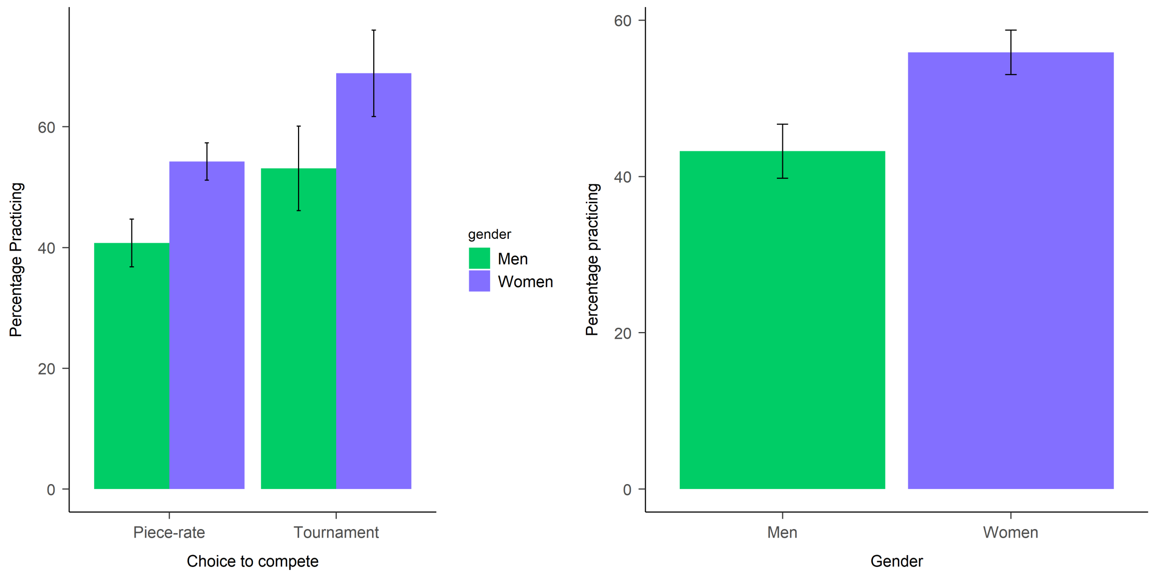

Figure 1.2: Right panel shows the proportion of men and women in Study 1 who chose to prepare. Left panel shows the proportion of men and women in Study 1 who chose to prepare based on choice to compete. Women choose to prepare more than men, regardless of their decision to compete. Error bars represent standard errors.

underlying the choice to prepare when first offered the opportunity versus the choice to continue preparing thereafter. However, we do not find evidence of gender differences in the choice to continue preparing after the initial decision to prepare - that is, the gender predictor in the count part of the model did not show a significant effect on the dependent variable, $b = 0.11$, 95% CI [-0.12, 0.34], $z = 0.96$, $p = 0.34$.

	(1)			(2)			(3)		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.76	0.63 – 0.91	**0.003**	0.69	0.56 – 0.84	**<0.001**	0.81	0.47 – 1.39	0.448
gender [Woman]	1.66	1.30 – 2.13	**<0.001**	1.72	1.31 – 2.27	**<0.001**	1.74	1.31 – 2.31	**<0.001**
comp choice [tournament]				1.65	1.05 – 2.59	**0.030**	1.71	1.07 – 2.73	**0.024**
gender [Woman] * comp choice [tournament]				1.13	0.55 – 2.37	0.740	1.06	0.51 – 2.25	0.872
task score							1.00	1.00 – 1.01	0.209
risk							1.03	0.97 – 1.08	0.322
conf rank							0.99	0.98 – 1.00	**0.002**
Observations	1018			1018			1004		
R^2 Tjur	0.016			0.025			0.035		

Table 1.6: All models are logistic regressions with choice to prepare as the dependent variable, where man and piece-rate payment scheme are the reference categories for participant gender and competition choice, respectively. Women prepare more than men regardless of competition choice, task score, risk attitudes, or confidence. $p < .05$ is considered significant and bolded.

Perceptions of gender differences in preparation, performance, and competitiveness

This gender difference in preparation aligned with participants' incentivized predictions about gender differences in preparation, where a greater proportion of participants (83.37%) expected women to spend more time preparing for the multiplication task, $\chi^2(1, n = 1056) = 447.11$, $p < .001$ (see Figure 1.3). They also expected women to prepare more in general, $\chi^2(1, n = 1056) = 625.06$, $p < .001$, with 89.51% indicating women prepare more in general versus 10.49% indicating that men prepare more in general (see Figure 1.6). However, participants did not expect any gender differences in performance on the task, $\chi^2(1, n = 1056) = 1.02$, $p = .313$ (see Figure 1.4). Additionally, participants accurately predicted that women were less likely to choose to compete, $\chi^2(1, n = 1056) = 716.24$, $p < .001$ (see Figure 1.5). See Table 1.7) for a summary of participants' responses to the questions about gender differences in preparation, performance, and competitiveness.

Characteristic	N = 1,056[1]
better_gender_guess	
Men	518 (52%)
Women	486 (48%)
perc_task_gender_pract	
Men	167 (17%)
Women	837 (83%)
perc_gender_comp	
Men	926 (92%)
Women	78 (7.8%)
perc_gen_gender_pract	
Men	105 (10%)
Women	896 (90%)

[1] n (%)

Table 1.7: Number and percentage of participants that selected each respective option when asked which gender would correctly solve more problems on the multiplication task, spend more time preparing for the multiplication task, choose the tournament payment scheme more often, and spend more time preparing on most tasks.

Effects of gender and perceptions on practicing

Given our evidence that women choose to prepare more and that participants believe women choose to prepare more, we explored whether women who believed other women prepare more were especially likely to prepare. To that end, we ran a logistic regression with the choice to practice as the dependent variable and gender, beliefs about gender differences in preparation on most tasks, and the interaction between those two variables as the predictors. We find that women who said women generally prepare more on most tasks were especially likely to prepare, $b = 1.00$,

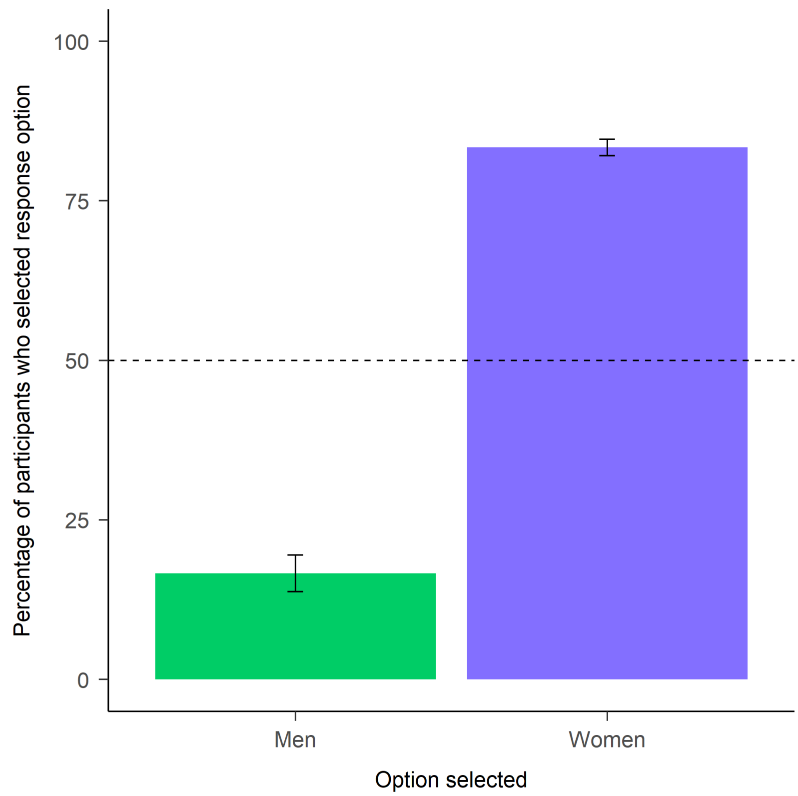

Figure 1.3: Proportion of participants that predicted women or men would spend more time preparing for the multiplication task. A significantly larger proportion of participants expected women to spend more time preparing for the multiplication task. Error bars represent standard errors.

95% CI [0.16, 1.85], $z = 2.33$, $p = .020$ (see Figure 1.7). We ran a nearly identical analysis with participants' beliefs about gender differences in preparation for the multiplication task, gender, and the interaction between the two as predictors instead, and replicate the interaction effect, $b = 0.77$, 95% CI [0.08, 1.46], $z = 2.19$, $p = .029$, such that women who said women spent more time preparing for the multiplication task were especially likely to prepare (see Figure 1.8).

Responses to calculator use questions

Finally, we analyzed participants' responses to the questions about their calculator use and thoughts on using a calculator for the multiplication task. 86% of participants indicated that they thought using a calculator to answer the multiplication questions

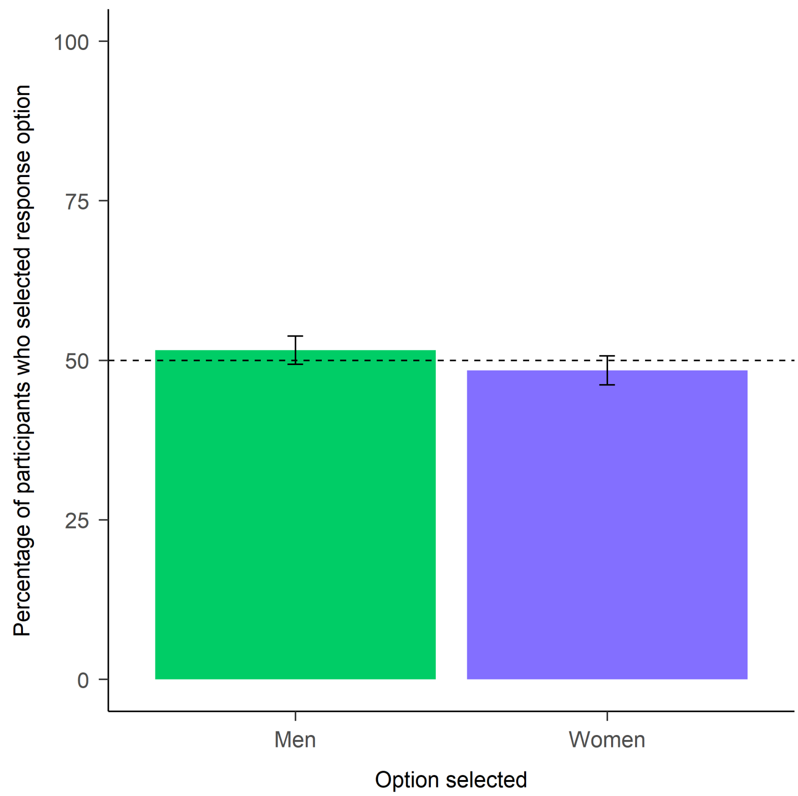

Figure 1.4: Proportion of participants that predicted women or men would correctly solve more problems on the multiplication task. There was no significant difference in the proportion of participants that expected women or men to perform better on the multiplication task. Error bars represent standard errors.

would slow them down and 93% of participants said they did not use a calculator. Importantly, there were no gender differences in perceptions of how calculators would affect performance, $\chi^2(1, n = 1056) = 0.42$, $p = .519$. Additionally, we did not find evidence of gender differences in actual calculator use, $\chi^2(1, n = 1056) = 1.70$, $p = .193$.

1.2.3 Discussion

Summary of main experimental results

Here we examine how knowledge of an opportunity to prepare influences gender differences in competitiveness. Our results suggest that knowledge of an opportunity to prepare did not change women and men's behavior to significantly different degrees. We also do not find evidence of any effect of condition on the choice to compete

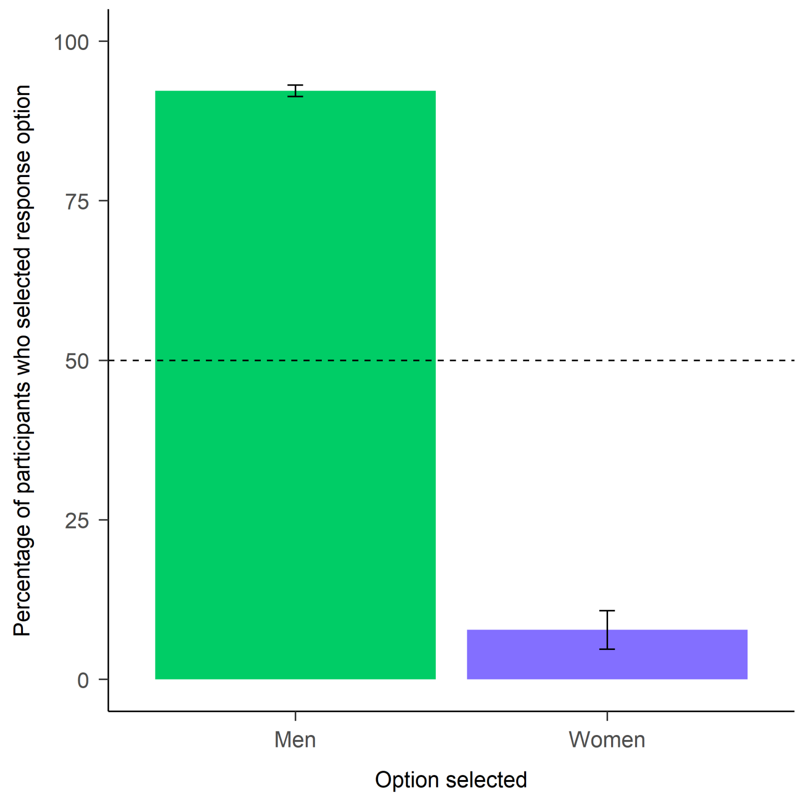

Figure 1.5: Proportion of participants that predicted women or men would choose the tournament payment scheme more often. A significantly larger proportion of participants expected men to be more likely to choose to compete. Error bars represent standard errors.

in general, such that participants' who knew they could prepare before choosing to compete were not significantly more likely to choose to compete than participants who were not provided this information before deciding whether to compete.

Despite no evidence of our anticipated effects of condition and the interaction between gender and condition on the choice to compete, we still find that men choose to compete significantly more than women, an effect that is explained by differences in risk attitudes and task scores.

Relative to other studies, only a small proportion of men (20.25%) and women (11.19%) in the current study chose to compete. For instance, **Niederle2007** find that 35% of women and 73% of men chose to compete. It is possible that the online nature of the task may have contributed to this result, which we explore further in

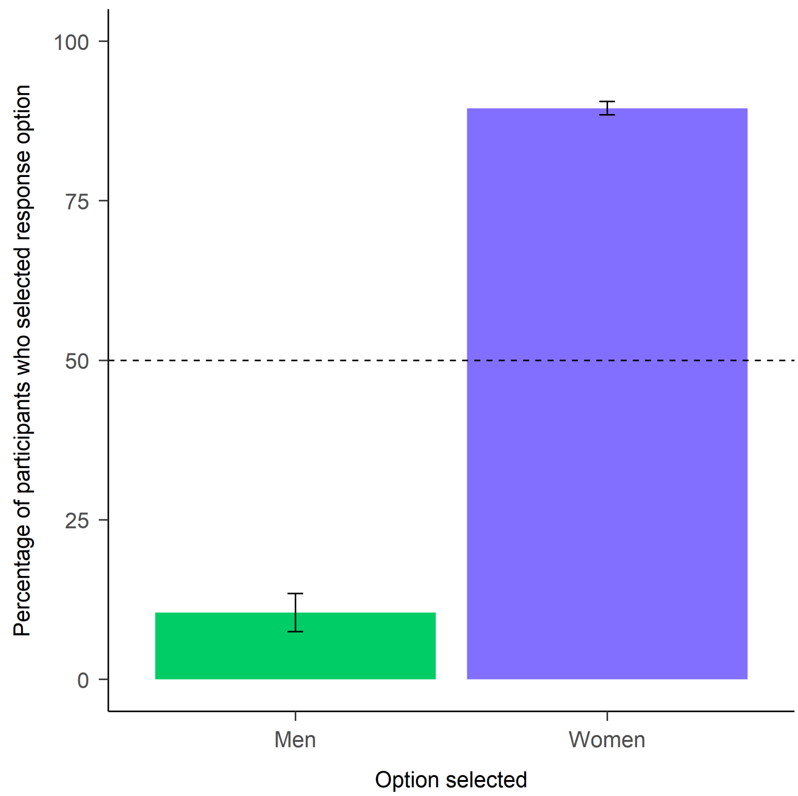

Figure 1.6: Proportion of participants that predicted women or men would spend more time preparing on most tasks. A significantly larger proportion of participants expected women to spend more time preparing on most tasks. Error bars represent standard errors.

the broader discussion section summarizing results across studies.

Contrary to previous research using similar arithmetic tasks [**Niederle2011**], we find that gender predicts task score, over and above other possible variables that likely drive performance (e.g., risk attitudes, choice in a payment scheme, confidence). We discuss implications of this finding further in the context of the rest of the results across studies in the overall discussion section.

Gender differences in preparation

Based on the binary measure of preparation, we find that women choose to prepare more than men, independent of whether they had chosen the competitive tournament or non-competitive piece-rate payment scheme. However, we do not find evidence

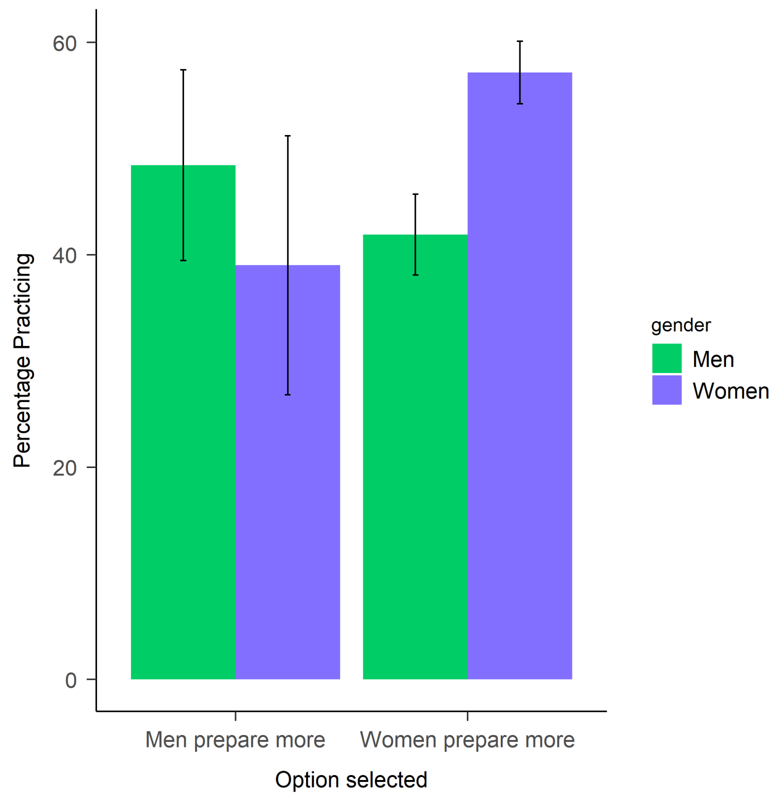

Figure 1.7: Proportion of men and women in Study 1 who chose to practice based on whether they thought men or women spend more time preparing on most tasks. Women who thought women generally prepare more were especially likely to choose to practice. Error bars represent standard errors.

that gender predicts the number of practice rounds completed.

We also included several incentivized questions eliciting participants' beliefs about gender differences across many of our behavioral variables of interest. Both men and women were significantly more likely than chance to say that women prepare more than men both before completing the multiplication task and in general[2]. Though we cannot directly attest to the general accuracy of these gender differences in preparation, since there is little research that has directly tested these gender differences, our study shows that women do indeed choose to prepare more than men, regardless of the payment scheme that they chose for their performance. Thus, we have our first piece of evidence from this study that gender differences in the choice to prepare and gender stereotypes

[2]Note that the question about general differences in practicing was unincentivized

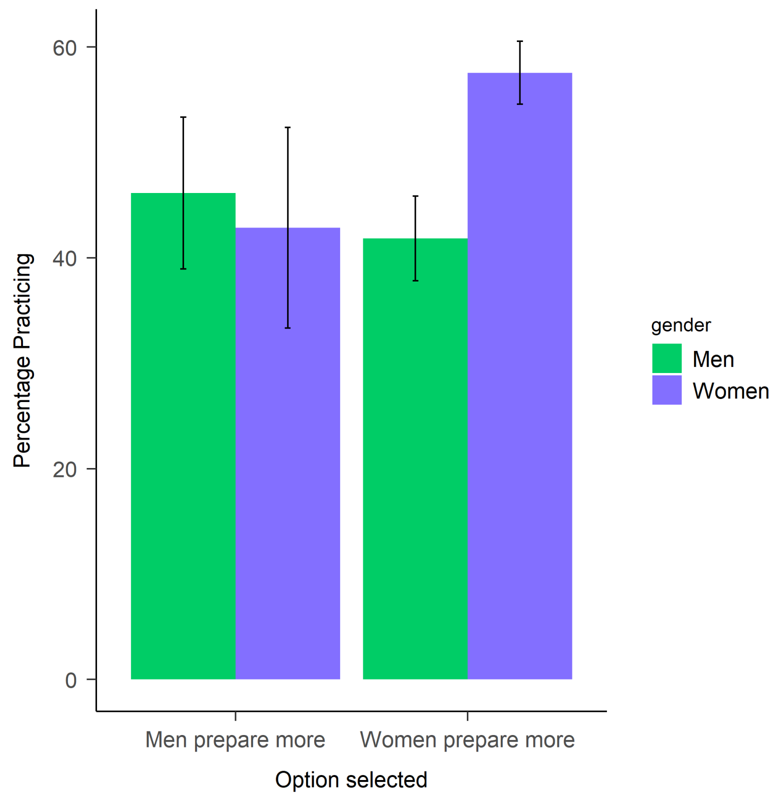

Figure 1.8: Proportion of men and women in Study 1 who chose to practice based on whether they thought men or women spend more time preparing for the multiplication task. Women who thought women prepare more for the multiplication task were especially likely to choose to practice. Error bars represent standard errors.

about preparation are aligned. This alignment suggests that gender stereotypes about practicing exist among the general population represented by our sample and may, at least for the task used in the current study, be accurate. Of course, these beliefs were elicited after participants themselves chose whether or not to practice. As such, their answers may have been influenced by knowledge of their own gender and behavior in the study and may not be reflective of strong gender stereotypes. Future work should explore participants' how beliefs may be shaped by the decision making-process and/or elicit beliefs before participants make any decisions to practice.

Perceptions of gender differences in preparation, performance, and competitiveness

Also, both men and women were significantly more likely than chance to correctly state that men would be more likely to choose to compete during the multiplication task, suggesting strong stereotypes about gender differences in competitiveness that align with observed behavior across the literature. The finding that participants believe women compete less suggests that they did not believe women prepare more because they were more likely to compete and used preparation as a way to increase the probability that they would earn more during the competition.

One possible explanation for participants' predictions is that they expected men to outperform women on the task, which would lead women to compensate by preparing more. Yet, they don't expect that: participants were equally likely to predict that women (vs. men) would perform better on the task, suggesting that participants did not have strong stereotypes about gender differences in performance on the multiplication task.

Effects of gender and perceptions on practicing

On top of the gender differences in practicing and perceptions of said gender differences, we explored whether these perceptions are aligned with individuals' practicing behavior. More specifically, we would expect that women may be especially likely to prepare for the task if they believe that women will prepare more than men, both in general on most tasks and specifically within the context of the multiplication task. We find support for this hypothesis across both questions about perceptions of gender differences in preparation. Though we cannot test this hypothesis causally with this specific set of data, these results would suggest that women's practicing behavior may have been influenced at least in part by beliefs about gender differences in preparation (though see Overall Discussion section for considerations of this interpretation).

Self-reported calculator use and perceptions of calculator use

There may be concern that participants used a calculator to answer the multiplication questions, which could affect the interpretation of the results if there is a gender difference in calculator use and/or calculator use is related to the choice to prepare. To address that potential concern, we included questions about participants' perceptions of and actual behaviors when it came to using calculators on the paid multiplication task. Based on their responses, it is unlikely that participants used calculators in the first place. We found that participants were unlikely to use calculators to complete the task and more importantly, there were no gender differences in the choice to use a calculator. Overall, we do not have evidence that gender differences in calculator use will be a confound when interpreting our results.

Summary

Overall, we find from the first study of Chapter 1 that our manipulation of the knowledge of the opportunity to prepare did not reduce the gender difference in competitiveness. We show that the gender difference in competitiveness itself is explained by gender differences in risk attitudes and task scores. Although the manipulation did not have the anticipated effects on gender differences in competitiveness, women chose to prepare more than men, which aligned with beliefs about gender differences in preparation. We also have evidence that women's preparation behaviors may in fact have been influenced by beliefs about gender differences in preparation. Finally, we show that participants were unlikely to be using calculators while completing the task, so it is unlikely that any of the effects found are confounded by participants' calculator use. Since we did not find evidence that the knowledge of the opportunity to prepare affected gender differences in competitiveness, the next study in Chapter 1 tests experimentally whether the actual experience of preparing through required preparation for a pre-specified number of rounds would affect the gender gap in competitiveness. We expect individuals may need to actually experience improvement in their performance directly for the preparation to affect one's willingness to compete.

1.3 Study 2

1.3.1 Methods

Participants

All study measures described below are publicly available on OSF both as a .pdf and .qsf. Participants were recruited on Amazon Mechanical Turk via CloudResearch using the same screening criteria as Study 1. For all studies in the dissertation after Study 1 of Chapter 1, we used CloudResearch to filter out participants that had already participated in any of the other studies in this dissertation. Therefore, only MTurkers who were naive to the design were included in the current and subsequent studies. Also, if participants had an identical IP address, MTurkID, and gender, we excluded their second response. The final sample consisted of 1088 participants (50.64% women), with an average age of 38.54 (SD = 12.5) years. 62 participants (51.61% women) dropped out of the study before finishing[3].

[3]Note: we use data from participants who dropped out when available

Demographics Across Conditions

Characteristic	Control condition, N = 536	Preparation condition, N = 535	**p-value**
Age	39 (13)	38 (12)	0.022
Unknown	23	22	
Gender			>0.9
Man	265 (49%)	264 (49%)	
Woman	271 (51%)	271 (51%)	
Race/ethnicity			0.4
Black/African-American	52 (10%)	63 (12%)	
East Asian	36 (7.0%)	27 (5.3%)	
Hispanic/Latino	25 (4.9%)	28 (5.5%)	
Middle Eastern	1 (0.2%)	1 (0.2%)	
Native American	5 (1.0%)	4 (0.8%)	
Pacific Islander	0 (0%)	3 (0.6%)	
Selected more than one option	22 (4.3%)	19 (3.7%)	
Selected other	10 (1.9%)	4 (0.8%)	
White/Caucasian	362 (71%)	364 (71%)	
Unknown	23	22	
Household Income			0.2
Less than $10,000	59 (12%)	74 (14%)	
$10,000 to $20,000	58 (11%)	59 (12%)	
$20,000 to $30,000	62 (12%)	76 (15%)	
$30,000 to $40,000	77 (15%)	74 (14%)	
$40,000 to $50,000	50 (9.7%)	57 (11%)	
$50,000 to $60,000	62 (12%)	50 (9.7%)	
$60,000 to $70,000	29 (5.7%)	40 (7.8%)	
$70,000 to $80,000	33 (6.4%)	28 (5.5%)	
$80,000 to $90,000	22 (4.3%)	14 (2.7%)	
$90,000 to $100,000	23 (4.5%)	14 (2.7%)	
$100,000 to $200,00	36 (7.0%)	23 (4.5%)	
Over $200,000	2 (0.4%)	4 (0.8%)	
Unknown	23	22	
Education			0.2
Less than a high school degree	0 (0%)	2 (0.7%)	
High School Diploma	66 (21%)	75 (25%)	
Vocational Training	15 (4.8%)	15 (5.0%)	
Some College	143 (46%)	142 (47%)	
Bachelor's degree	0 (0%)	0 (0%)	
Graduate Degree	89 (28%)	66 (22%)	
Unknown	223	235	

Table 1.8: Size of sample in Study 2 with corresponding percentage listed for gender, race, education, and household income, with p-values derived from Fisher's exact test. Mean with corresponding standard deviation listed for age, with p-values derived from Kruskal-Wallis test. If a participant did not respond to a given question, we list their response as 'Unknown'.

Demographics Based on Participant Gender

Characteristic	**Man**, N = 537	**Woman**, N = 551	**p-value**
Age	37 (12)	40 (13)	<0.001
Unknown	30	32	
Race/ethnicity			0.6
Black/African-American	59 (12%)	56 (11%)	
East Asian	38 (7.5%)	25 (4.8%)	
Hispanic/Latino	30 (5.9%)	23 (4.4%)	
Middle Eastern	1 (0.2%)	1 (0.2%)	
Native American	5 (1.0%)	4 (0.8%)	
Pacific Islander	2 (0.4%)	1 (0.2%)	
Selected more than one option	21 (4.1%)	20 (3.9%)	
Selected other	6 (1.2%)	8 (1.5%)	
White/Caucasian	345 (68%)	381 (73%)	
Unknown	30	32	
Household Income			0.007
Less than $10,000	61 (12%)	72 (14%)	
$10,000 to $20,000	55 (11%)	62 (12%)	
$20,000 to $30,000	49 (9.7%)	89 (17%)	
$30,000 to $40,000	80 (16%)	71 (14%)	
$40,000 to $50,000	55 (11%)	52 (10%)	
$50,000 to $60,000	61 (12%)	51 (9.8%)	
$60,000 to $70,000	38 (7.5%)	31 (6.0%)	
$70,000 to $80,000	27 (5.3%)	34 (6.6%)	
$80,000 to $90,000	24 (4.7%)	12 (2.3%)	
$90,000 to $100,000	21 (4.1%)	16 (3.1%)	
$100,000 to $200,00	35 (6.9%)	24 (4.6%)	
Over $200,000	1 (0.2%)	5 (1.0%)	
Unknown	30	32	
Education			>0.9
Less than a high school degree	1 (0.3%)	1 (0.3%)	
High School Diploma	68 (23%)	73 (23%)	
Vocational Training	17 (5.6%)	13 (4.2%)	
Some College	137 (46%)	148 (47%)	
Bachelor's degree	0 (0%)	0 (0%)	
Graduate Degree	78 (26%)	77 (25%)	
Unknown	236	239	

Table 1.9: Size of sample in Study 2 with corresponding percentage listed for race, education, and household income, with p-values derived from Fisher's exact test. Mean with corresponding standard deviation listed for age, with p-values derived from Kruskal-Wallis test. If a participant did not respond to a given question, we list their response as 'Unknown'.

Procedures

As in Study 1, participants included in the study were told they would be completing a two-minute multiplication task (identical to the one used in Study 1) and would be able to choose a payment scheme for their performance. The instructions and payment per question were identical to Study 1. After being told about the rules for the multiplication task and passing the same comprehension questions used in Study 1, participants were randomly assigned to either a preparation condition, where they were told they would complete several rounds of preparation before completing the multiplication task, or a control condition, where they were told they would complete several rounds of a counting task before continuing. Participants were randomly assigned to each condition, control= 50.05%. Of the men who completed the study, 50.09% were assigned to the control condition and of the women who completed the study, 50% were assigned to the control condition, $\chi^2(1, n = 1088) = 0.00$, $p > .999$.

The participants in the preparation condition completed 12 rounds of practice for each multiplication table (i.e., tables 1-12), with 6 problems per round. The problems for each round were selected at random. Participants in the control condition were asked to complete 5 questions where they counted the number of zeros in a 7x7 matrix of zeros and ones. After a 30-second break following completion of their respective tasks, all participants chose a payment scheme (i.e., piece-rate or tournament) for the multiplication task, where the order of presentation was counterbalanced. That is, half of the participants saw the tournament scheme presented as the first option and half saw the piece-rate payment scheme presented first.

After choosing a payment scheme, participants in both conditions had the option to spend (extra) time preparing for the multiplication task. Importantly, they were not told about the additional opportunity to prepare until after they made their decision to compete. Again, we had two measures of preparation behavior: the binary decision to practice and the number of extra practice rounds completed. If they chose to prepare, participants were given two minutes to complete a randomly selected set of problems from all 12 multiplication tables. Once they finished the first two-minute preparation round, participants could opt into 4 more rounds of preparation, each two minutes long, before they moved on to the paid portion of the study. Here, the practice rounds variable is encoded the same way for participants across conditions, where the choice to practice is encoded as 1 for that variable, and the variable increases incrementally (up to 5 given the design) thereafter ($M = 0.53$, $SD = 0.82$); 39.04% of participants opted for additional practice.

Then, participants completed the paid multiplication task for two minutes. We included many of the same follow-up questions as in Study 1, including risk aversion, confidence, and perceptions of gender differences in preparation, competitiveness, and performance. Like Study 1, participants were incentivized to answer the questions about their confidence and perceptions of gender differences correctly. Specifically, they were told that one of their responses to those questions would be randomly selected and if the selected option was correct, they would earn a bonus of $.10 on top of their guaranteed earnings. We also asked participants if they wished they had more time to prepare for the multiplication task, a binary variable (response options: "Yes" or "No") that we subsequently describe as participants' self-reported "feelings of preparedness". Thereafter, we included measures of their fatigue, field-specific ability beliefs, and interest in the multiplication task all on 1 (Strongly disagree) to 7 (Strongly agree) scales. For the fatigue scale, participants rated how fatigued and mentally exhausted they felt [**Milyavskaya2018**]. Participants indicated the degree to which they "enjoyed completing the multiplication task" for the interest scale [**Milyavskaya2018**]. Finally, to measure field-specific ability beliefs, we asked participants how much they perceived success in math depends on ability versus effort through six questions (e.g., "If you want to succeed in math, hard work alone just won't cut it; you need to have an innate gift or talent") [**Meyer2015**].

1.3.2 Results

Describing main variables of interest

We replicated the effect of gender on the choice to compete when gender is included as the only predictor in the logistic regression: 19.85% of men chose to compete compared to 13.91% of women, $b = -0.43$, 95% CI $[-0.76, -0.10]$, $z = -2.56$, $p = .010$. Like Study 1, the gender effect on competitiveness is no longer significant after adding the same control variables as before (i.e., risk attitudes, confidence, task scores, and the hypothesized interaction between gender and competition choice), where risk attitudes, $b = 0.34$, 95% CI $[0.26, 0.42]$, $z = 8.34$, $p < .001$, confidence, $b = 0.01$, 95% CI $[0.01, 0.02]$, $z = 3.04$, $p = .002$, and task scores, $b = 0.01$, 95% CI $[0.00, 0.02]$, $z = 3.10$, $p = .002$ appear to explain the gender differences in competitiveness (see Table 1.10).

		(1)			**(2)**			**(3)**	
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.25	0.20 – 0.31	**<0.001**	0.26	0.19 – 0.35	**<0.001**	0.01	0.00 – 0.02	**<0.001**
gender [Woman]	0.65	0.47 – 0.90	**0.010**	0.71	0.45 – 1.11	0.131	1.18	0.71 – 1.96	0.517
condition [pract]				0.89	0.58 – 1.37	0.597	0.93	0.58 – 1.48	0.753
gender [Woman] * condition [pract]				0.84	0.44 – 1.62	0.610	0.67	0.33 – 1.38	0.283
risk							1.41	1.30 – 1.53	**<0.001**
conf rank							1.01	1.01 – 1.02	**0.002**
task score							1.01	1.00 – 1.02	**0.002**
Observations	1051			1051			1026		
R^2 Tjur	0.006			0.008			0.139		

Table 1.10: All models are logistic regressions with choice to compete as the dependent variable, where man and control are the reference categories for participant gender and preparation condition, respectively. The gender difference in the choice to compete is not reduced by preparation condition, but is explained by risk attitudes, task scores, and confidence. p < .05 is considered significant and bolded.

Again, we find that gender predicts task scores when included by itself as a predictor, $b = -3.68$, 95% CI $[-6.27, -1.10]$, $t(1028) = -2.80$, $p = .005$. However, unlike Study 1, when other variables are included as predictors in the linear regression, we find that the effect of gender on task scores dissipates, $b = -0.46$, 95% CI $[-3.17, 2.24]$, $t(1020) = -0.34$, $p = .738$, suggesting that the other variables, such as competition choice, $b = 7.17$, 95% CI $[2.83, 11.52]$, $t(1020) = 3.24$, $p = .001$, risk attitudes, $b = -1.35$, 95% CI $[-1.85, -0.86]$, $t(1020) = -5.42$, $p < .001$, and confidence, $b = 0.35$, 95% CI $[0.30, 0.41]$, $t(1020) = 12.54$, $p < .001$, explained the gender difference in task scores in this study (see 1.12). In support of this possibility, we replicate the finding from the Study 1 of this chapter that gender predicts both risk attitudes, $b = -1.14$, 95% CI $[-1.46, -0.82]$, $t(1024) = -7.00$, $p < .001$ and confidence, $b = -9.76$, 95% CI $[-12.49, -7.02]$, $t(1027) = -6.99$, $p < .001$. See Table 1.11 for a summary of gender differences in the main variables of interest.

Variable	**Man**, N = 537[1]	**Woman**, N = 551[1]
task_score	48 (35, 62)	44 (32, 57)
comp_choice		
piecerate	416 (80%)	458 (86%)
tournament	103 (20%)	74 (14%)
pract_choice	185 (36%)	221 (42%)
risk	6.00 (3.00, 7.00)	4.00 (2.00, 6.00)
conf_rank	60 (50, 80)	50 (40, 70)

[1] Median (IQR); n (%)

Table 1.11: Gender differences in the main variables of interest, including: task scores, choice to compete, choice to practice, confidence, and risk attitudes. Medians are reported for task score, risk attitudes, and confidence, with IQRs in parentheses. For choice to practice and choice to compete, we report the number and percentage of participants that fall into each category for each respective gender.

	(1)			(2)			(3)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	50.04	48.20 – 51.87	**<0.001**	48.36	46.32 – 50.41	**<0.001**	34.48	30.30 – 38.67	**<0.001**
gender [Woman]	-3.68	-6.27 – -1.10	**0.005**	-2.24	-5.05 – 0.58	0.120	-0.46	-3.17 – 2.24	0.738
comp choice [tournament]				8.41	3.83 – 12.99	**<0.001**	7.17	2.83 – 11.52	**0.001**
gender [Woman] * comp choice [tournament]				-6.74	-13.74 – 0.27	0.059	-5.86	-12.42 – 0.70	0.080
conf rank							0.35	0.30 – 0.41	**<0.001**
risk							-1.35	-1.85 – -0.86	**<0.001**
Observations	1030			1030			1026		
R^2 / R^2 adjusted	0.008 / 0.007			0.020 / 0.017			0.158 / 0.154		

Table 1.12: All models are linear regressions with task score as the dependent variable, where man and piece-rate payment scheme are the reference categories for participant gender and competition choice, respectively. After controlling for risk attitudes, confidence, and competition choice, women no longer have lower scores on the multiplication task than men, p $<$ .05 is considered significant and bolded.

Effects of limited preparation condition on gender differences in choice to compete

We did not find evidence of an interaction between gender and preparation condition on the choice to compete in a logistic regression, $b = -0.17$, 95% CI $[-0.83, 0.48]$, $z = -0.51$, $p = .610$ (see Figure 1.9). Also, we did not find evidence of a significant effect of condition on the choice to compete as a sole predictor in a logistic regression, $b = -0.19$, 95% CI $[-0.52, 0.13]$, $z = -1.17$, $p = .243$.

Gender differences in preparation

Despite no evidence for the effect of requiring people to prepare (i.e., condition) on the choice to compete across participants, we replicate the effect of gender on the

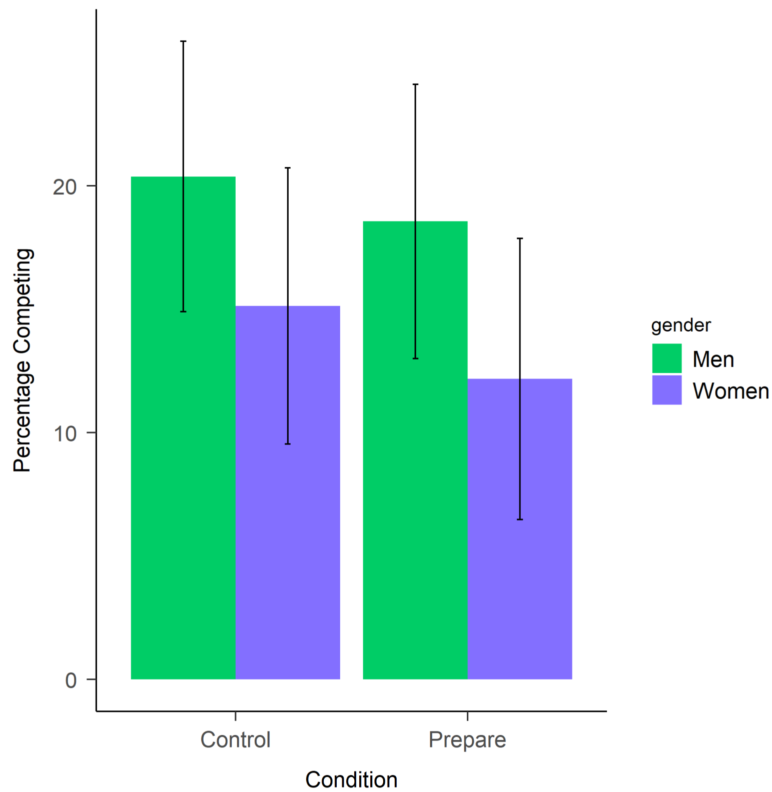

Figure 1.9: Proportion of men and women in Study 2 who chose to compete based on preparation condition. Limited preparation did not reduce the gender difference in competitiveness. Error bars represent standard errors.

choice to practice found in Study 1, where women were significantly more likely to opt-in to prepare for the task, even after being required to prepare in the preparation condition. 42.02% of women across conditions chose to practice for the multiplication task beyond what was required, relative to 35.99% of men, $b = 0.25$, 95% CI [0.00, 0.50], $z = 1.99$, $p = .047$ (see right panel of Figure 1.10). The gender effect holds even after controlling for the decision to compete and the interaction between gender and the decision to compete, $b = 0.31$, 95% CI [0.03, 0.59], $z = 2.17$, $p = .030$ (see left panel of Figure 1.10). Within the same model, we find that the choice to compete itself increases the likelihood a participant will practice before completing the paid task, $b = 0.83$, 95% CI [0.39, 1.27], $z = 3.70$, $p < .001$, but no evidence of an interaction between gender and payment scheme choice, $b = 0.04$, 95% CI [-0.63, 0.72], $z = 0.12$,

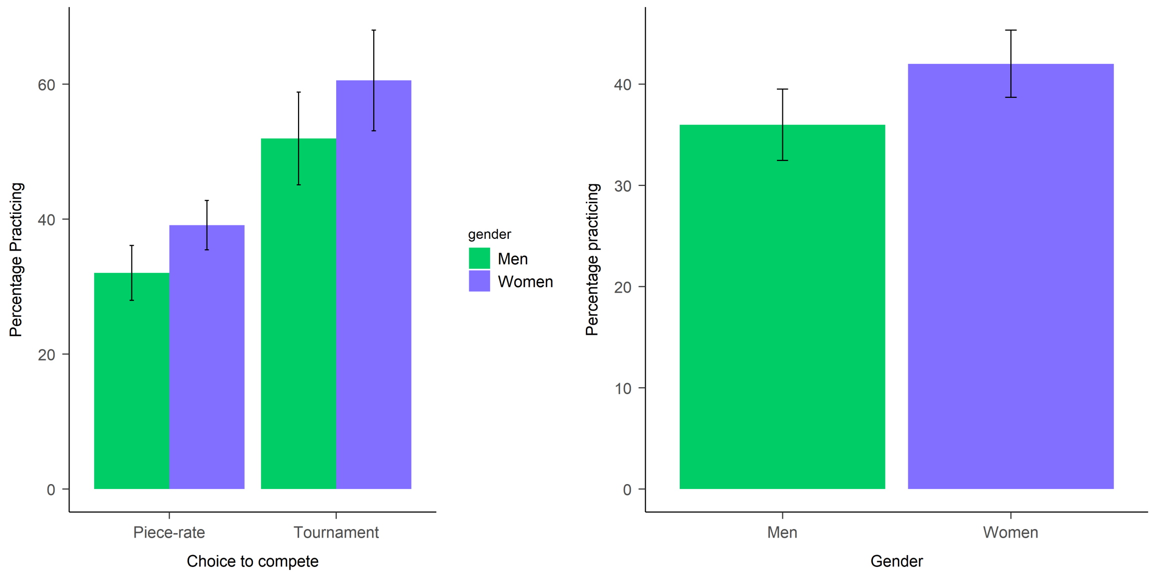

Figure 1.10: Right panel shows the proportion of men and women in Study 2 who chose to prepare. Left panel shows the proportion of men and women in Study 2 who chose to prepare based on choice to compete. Women choose to prepare more than men, regardless of their decision to compete. Error bars represent standard errors.

p = .906. To see if the gender effect is explained by other variables included in the study, we added confidence, risk attitudes, and task scores to the previous model, and find that gender still significantly predicts the choice to practice, $b = 0.41$, 95% CI [0.12, 0.71], z = 2.76, p = .006, over any effects of differences in risk attitudes, confidence, or task scores (see Table 1.13). We ran the same two-part hurdle model described in Study 1 with gender, competition choice, and the interaction between those variables predicting the number of practice rounds variable. Again, we do not find evidence of gender differences in the choice to continue preparing after the initial decision to prepare, b = -0.14, 95% CI [-0.53, 0.25], z = -0.71, p = 0.48.

	(1)			(2)			(3)		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.56	0.47 – 0.67	**<0.001**	0.47	0.38 – 0.58	**<0.001**	0.28	0.17 – 0.47	**<0.001**
gender [Woman]	1.29	1.00 – 1.66	**0.047**	1.36	1.03 – 1.80	**0.030**	1.51	1.13 – 2.03	**0.006**
comp choice [tournament]				2.29	1.48 – 3.57	**<0.001**	1.98	1.25 – 3.13	**0.003**
gender [Woman] * comp choice [tournament]				1.04	0.53 – 2.06	0.906	0.98	0.49 – 1.96	0.951
conf rank							1.00	0.99 – 1.01	0.898
risk							1.08	1.03 – 1.14	**0.004**
task score							1.00	1.00 – 1.01	0.498
Observations	1040			1040			1026		
R^2 Tjur	0.004			0.028			0.036		

Table 1.13: All models are logistic regressions with choice to prepare as the dependent variable, where man and piece-rate payment scheme are the reference categories for participant gender and competition choice, respectively. Women prepare more than men regardless of competition choice, task score, risk attitudes, or confidence. p < .05 is considered significant and bolded.

Perceptions of gender differences in preparation, performance, and competitiveness

Again, we find that these results align with participants' expectations, where they were significantly more likely to expect women (with 85.38% selecting women versus 14.62% selecting men) to choose to prepare more than men both in general, $\chi^2(1, n = 1088) = 513.72$, $p < .001$ (see Figure 1.14), and on the paid multiplication task, $\chi^2(1, n = 1088) = 394.33$, $p < .001$ (see Figure 1.11) (with 80.95% selecting women versus 19.05% selecting men), despite expecting men to choose to compete more often, $\chi^2(1, n = 1088) = 580.69$, $p < .001$ (see Figure 1.13) and expecting no gender differences in performance on the task, $\chi^2(1, n = 1088) = 0.51$, $p = .473$ (see Figure 1.12). See Table 1.14 for a summary of participants' responses to the questions about gender differences in preparation, performance, and competitiveness.

Characteristic	N = 1,088[1]
better_gender_guess	
Men	503 (49%)
Women	526 (51%)
perc_task_gender_pract	
Men	196 (19%)
Women	833 (81%)
perc_gender_comp	
Men	901 (88%)
Women	128 (12%)
perc_gen_gender_pract	
Men	150 (15%)
Women	876 (85%)

[1] n (%)

Table 1.14: Number and percentage of participants that selected each respective option when asked which gender would correctly solve more problems on the multiplication task, spend more time preparing for the multiplication task, choose the tournament payment scheme more often, and spend more time preparing on most tasks.

Effects of gender and perceptions on practicing

Like Study 1, we explored whether women who believed other women prepare more were especially likely to prepare. To that end, we ran a logistic regression with the choice to practice as the dependent variable and gender, beliefs about gender differences in preparation on most tasks, and the interaction between those two variables as the predictors. We replicate the interaction effect found in Study 1, such that women who said women generally prepare more on most tasks were especially likely to prepare, $b = 1.21$, 95% CI [0.42, 2.04], $z = 2.94$, $p = .003$ (see Figure 1.15). We ran the

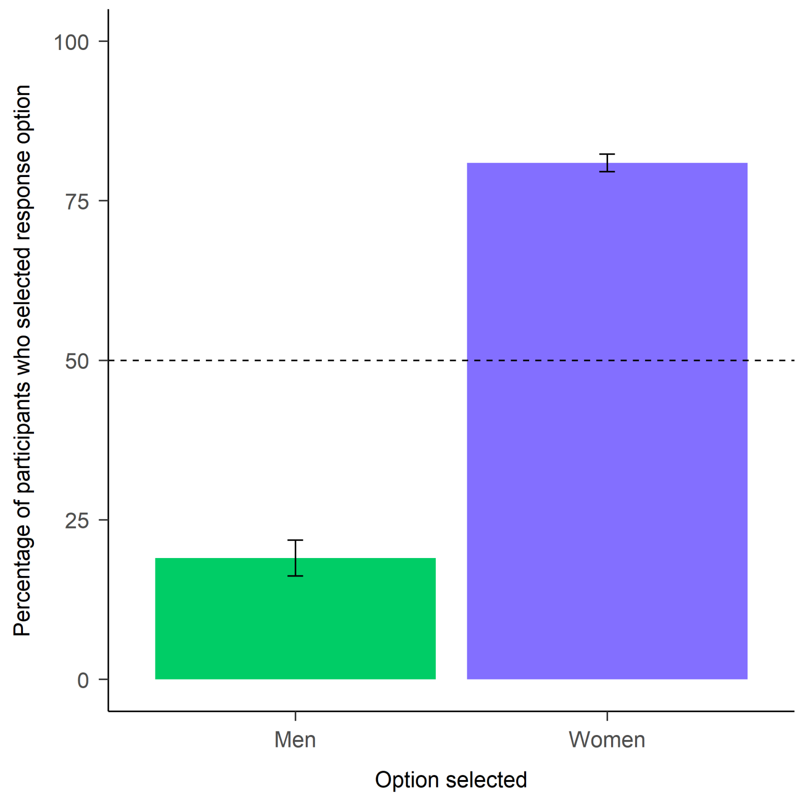

Figure 1.11: Proportion of participants that predicted women or men would spend more time preparing for the multiplication task. A significantly larger proportion of participants expected women to spend more time preparing for the multiplication task. Error bars represent standard errors.

same analysis with participants' beliefs about gender differences in preparation for the multiplication task, gender, and the interaction between the two as predictors instead, and replicate the interaction effect from the previous study, $b = 1.52$, 95% CI [0.86, 2.18], $z = 4.51$, $p < .001$, such that women who said women spent more time preparing for the multiplication task were especially likely to prepare (see Figure 1.16).

Post-manipulation measures

We also added several post-manipulation questions to tap into participants' experience of the multiplication task itself, feelings of preparedness, and general beliefs about the value of preparation to see if they may explain some of the observed effects. First, in a logistic regression with feelings of preparedness regressed upon condition

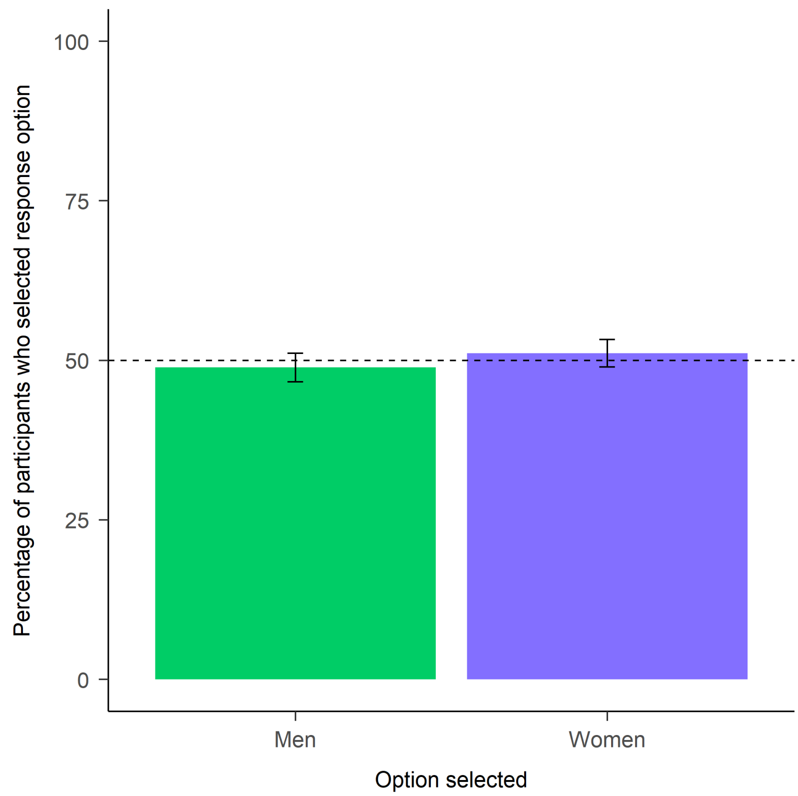

Figure 1.12: Proportion of participants that predicted women or men would correctly solve more problems on the multiplication task. There was no significant difference in the proportion of participants that expected women or men to perform better on the multiplication task. Error bars represent standard errors.

and the interaction between preparation choice and gender, only the choice to prepare predicted feelings of preparation, $b = 1.06$, 95% CI [0.56, 1.57], $z = 4.14$, $p < .001$.

We also tested whether the interaction between practice choice and condition, along with gender, predicted participants' interest in the multiplication and self-reported fatigue after completing the paid task. Our results suggest that participants who chose to prepare before the task reported feeling significantly more fatigued than those who did not choose to practice, $b = 0.31$, 95% CI [0.02, 0.59], $t(1021) = 2.08$, $p = .038$, and that participants in the preparation condition were significantly more fatigued than those who were assigned to the control condition, $b = 0.34$, 95% CI [0.08, 0.60], $t(1021) = 2.59$, $p = .010$. We do not find evidence that gender, $b = -0.02$, 95% CI [−0.22, 0.18], $t(1021) = -0.20$, $p = .839$, nor the interaction between condition and

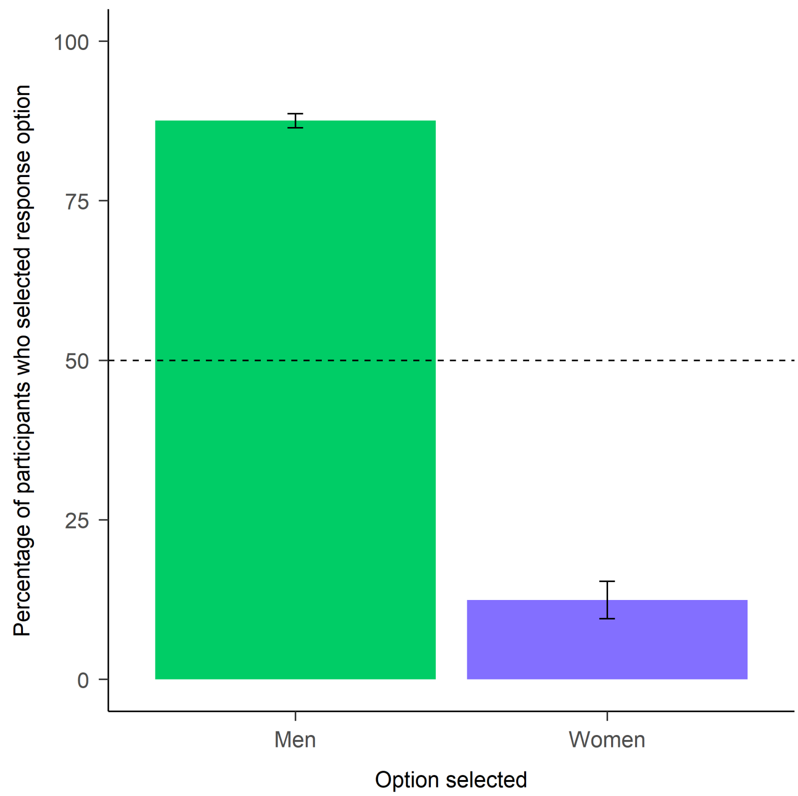

Figure 1.13: Proportion of participants that predicted women or men would choose to compete. A significantly larger proportion of participants expected men to be more likely to choose to compete. Error bars represent standard errors.

practice choice, $b = 0.17$, 95% CI $[-0.25, 0.59]$, $t(1021) = 0.78$, $p = .433$, predicted self-reported fatigue. When gender is included as a sole predictor of self-reported fatigue, we again find no evidence that there are significant gender differences in fatigue, $b = 0.01$, 95% CI $[-0.20, 0.21]$, $t(1024) = 0.08$, $p = .940$. Also, we do not have evidence that gender predicted field-specific ability beliefs, $b = -0.09$, 95% CI $[-0.23, 0.05]$, $t(1024) = -1.30$, $p = .192$, contrary to previous work [**Leslie2015**]. Finally, we find that women report being significantly less interested in the task, $b = -0.27$, 95% CI $[-0.45, -0.08]$, $t(1021) = -2.88$, $p = .004$, even though participants who chose to prepare tend to be significantly more interested in the task, $b = 0.34$, 95% CI $[0.08, 0.60]$, $t(1021) = 2.61$, $p = .009$.

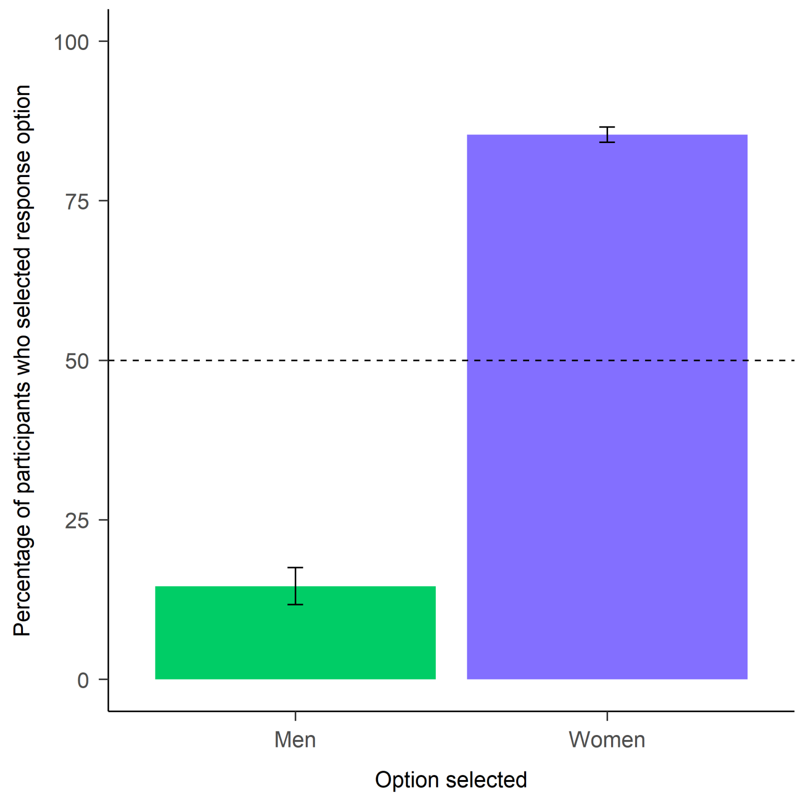

Figure 1.14: Proportion of participants that predicted women or men would spend more time preparing on most tasks. A significantly larger proportion of participants expected women to spend more time preparing on most tasks. Error bars represent standard errors.

1.3.3 Discussion

Summary of main experimental results

There is no evidence that the limited opportunity to prepare affected participants' choice to compete, nor did we find evidence of the anticipated interaction effect between gender and condition. Thus, it appears that gender differences in competitiveness are unaffected by having had the opportunity to prepare for a limited, pre-specified amount of time.

In fact, the gender difference in competitiveness itself, like in Study 1, does not hold when risk attitudes, confidence, and task scores are included in the model, suggesting the difference was explained by those variables. In this way, our study deviates from previous research suggesting that competitiveness is a stand-alone trait [**Niederle2007**],

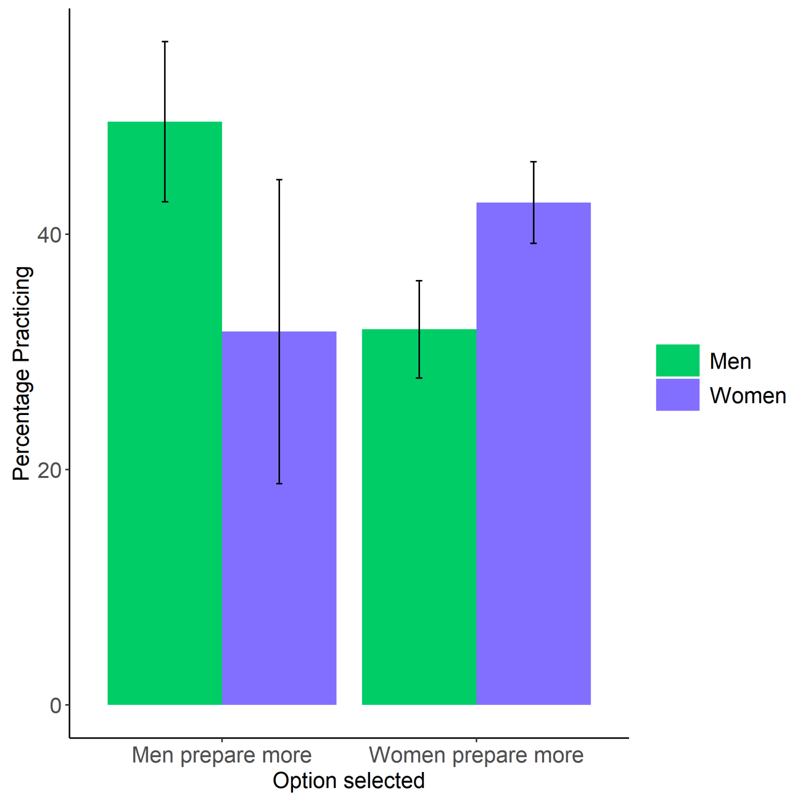

Figure 1.15: Proportion of men and women in Study 2 who chose to practice based on whether they thought men or women spend more time preparing on most tasks. Women who thought women generally prepare more were especially likely to choose to practice. Error bars represent standard errors.

and is more in line with recent evidence that competitiveness is explained by other factors (e.g., confidence or risk attitudes) [**Gillen2019**, **Veldhuizen2017**]. In line with this argument, we replicate the finding both within the broader literature and in the previous study that women are more risk averse and less confident than men. We discuss these findings across the studies of Chapter 1 in light of broader literature on whether competitiveness is a stand-alone trait further in the overall discussion section.

Contrary to Study 1, the gender differences in task scores observed when gender is included as the only predictor appear to be explained by differences in risk attitudes and confidence, since the effect of gender on task score is no longer significant when those predictors are included in the model. Thus, based on these first two sets of results, it is not entirely clear yet if there are gender differences in performance on

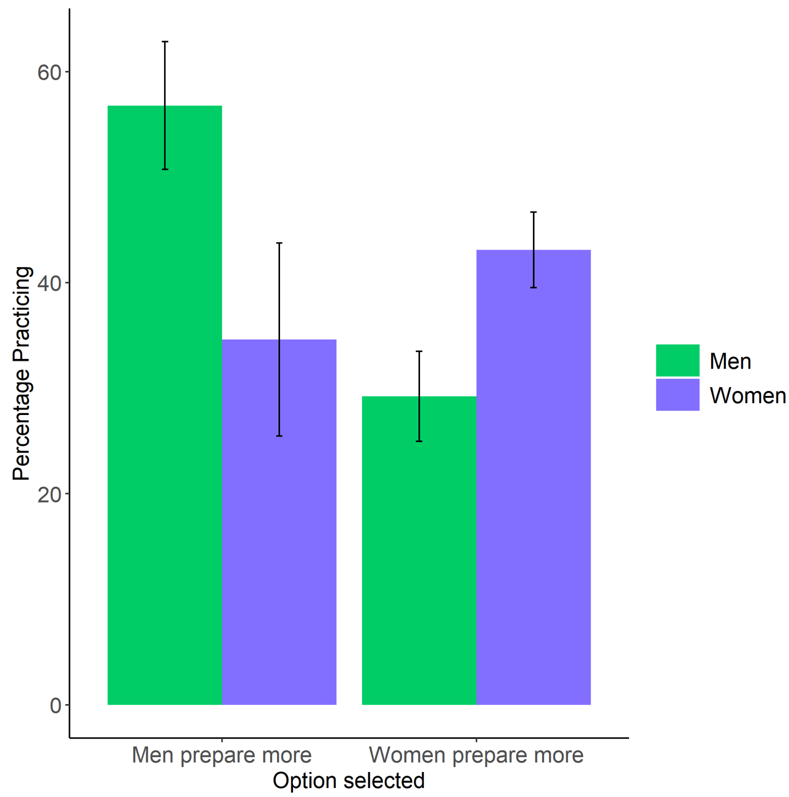

Figure 1.16: Proportion of men and women in Study 2 who chose to practice based on whether they thought men or women spend more time preparing for the multiplication task. Women who thought women prepare more for the multiplication task were especially likely to choose to practice. Error bars represent standard errors.

the multiplication task used in the study.

Gender differences in preparation and gender stereotypes

Even though we do not have evidence within the context of this study that there is a gender difference in competitiveness or performance on the multiplication task when controlling for other variables, women still chose to prepare more than men - regardless of which payment scheme they ultimately selected, their risk attitudes, confidence, or task scores. At the same time, we have evidence that the choice to compete in general is related to a participants' choice to prepare, such that individuals who chose the competitive payment scheme chose to prepare at higher rates than those who chose the non-competitive payment scheme. We replicate the effects found in Study 1 when asking participants about their perceptions of these gender differences in performance, competitiveness, and preparation, such that participants were 1) significantly more likely to say that women prepare more than men both in general and on the multiplication task, despite 2) competing less and 3) performing just as well as men. Thus, we have further evidence based on participants' responses to the incentivized questions about their beliefs that there are gender stereotypes about preparation, which may in turn affect gender differences in the choice to prepare (see Section **??** for a review of the literature showing how gender stereotypes affect subsequent behavior).

On top of the gender differences in practicing and perceptions of said gender differences, we again explored whether these perceptions aligned with individuals' practicing behaviors. We find evidence that women were especially likely to prepare for the task if they believed that women prepare more than men, both in general on most tasks and specifically within the context of the multiplication task. Again, we highlight the importance of future research in testing the direction of these effects, since we cannot attest to whether perceptions lead to behavior or whether behavior leads to perceptions directly using this dataset.

Post-manipulation measures

Finally, we explore the new post-manipulation measures that are specific to this study. First, we find that feelings of preparedness are related to the choice to prepare. In light of the other findings, this result implies that feelings of preparedness do not derive from the required amount of preparation that one completes (that is, random assignment to condition within the context of this study), but instead, the choice to spend extra time preparing over and above what is required[4]. Additionally, the fact that gender does not predict feelings of preparedness, despite women choosing to

[4]Though it is worth noting that the effect of the choice to prepare may reflect the effects of variables that we are not measuring, while the random assignment to condition is estimating the causal effect of preparation on feelings of preparedness. It would be important to consider possible confounds for the effect of the choice to prepare and try to adequately capture those in future research

prepare more, has potentially important implications if replicated. That is, women may prepare more than men, yet still feel as though they are less prepared relative to others. We explore this possibility in the study of Chapter 2. These results would suggest that women's greater decisions to prepare may only contribute to feelings of preparedness that are not significantly different than men's, if they contribute at all. However, it is unclear whether the choice to prepare may be related to other variables not captured in this study that predict feelings of preparedness. Thus, future studies should explore women's and men's feelings of preparedness both before and after the choice to prepare as a way to try to address the question of what is driving the effect.

On top of the question about feelings of preparedness after completing the practice and task itself, we also asked participants to indicate their feelings about the multiplication task itself, and more specifically, how fatigued they felt after completing the task and how interested they were in the task. With regards to the question about fatigue, both participants in the limited preparation condition and participants who chose to prepare reported significantly higher levels of fatigue than participants who did not choose to prepare and were not required to. Yet, women did not report feeling significantly more fatigued than men, despite, again, choosing to prepare at significantly higher rates. Thus, women's higher rates of preparation did not seem to have negative effects on this measure attempting to capture potential negative effects of preparing. When testing effects of gender, condition, and practice choice on interest, we find that women were significantly less interested in the task, while participants who chose to prepare tended to be more interested in the task. Finally, we show that gender does not predict field-specific ability beliefs - in other words, men and women's emphasis on the importance of hard work for success (at least within the domain of math), do not appear to be significantly different in this study. This is the only study in the dissertation that includes the measures of feelings of preparedness, fatigue, and interest, and as such, replication of these results would be important to confirm they are not false positives [**Coffman2015**].

Summary

Overall, we find from Study 2 of Chapter 1 that our manipulation of the limited opportunity to prepare did not reduce the gender difference in competitiveness. We show that the gender difference in competitiveness itself is explained by gender differences in risk attitudes, task scores, and confidence. Although the manipulation did not have the anticipated effects on gender differences in competitiveness, we replicate the effect that women chose to prepare more than men, which aligned with beliefs about gender differences in preparation. We also replicate the interaction effect between gender and beliefs about gender differences in preparation on the choice to prepare, again suggesting that women's preparation behaviors may have been influenced by beliefs about gender differences in preparation.

Finally, we explored participants' experiences while completing the multiplication task through questions about their feelings of preparedness, fatigue, and interest. We find that the limited preparation condition did not predict feelings of preparedness, but instead participants who chose to prepare were significantly more likely to report feeling more prepared for the multiplication task. For the measure of fatigue, participants who chose to prepare were significantly more likely to report feeling fatigued after completing the task. Finally, women were significantly less interested in the task, while participants who chose to prepare tended to be more interested in the task. Since this is the only study across all studies within this dissertation that employs these measures, and thus, we have not been able to replicate these effects, we encourage future research to ensure these effects are indeed replicable.

Since we did not find evidence that the limited opportunity to prepare affected gender differences in competitiveness, the last study in Chapter 1 tests experimentally whether having the unlimited opportunity to prepare would affect the gender gap in competitiveness. It is possible that participants in the limited preparation condition may have felt like they were not getting any additional preparation above and beyond what other participants were receiving because they could not choose how much they prepared, and thus, they did not feel like the preparation would have helped their performance substantially more than it would have helped other participants. If this possible explanation for the null effect in Study 2 is correct, we should expect to see an effect of unlimited preparation on reducing the gender difference in competitiveness.

1.4 Study 3

1.4.1 Methods

Participants

All study measures described below are publicly available on OSF both as a .pdf and .qsf. Participants were recruited on Amazon Mechanical Turk using the same screening criteria as Studies 1 and 2. Unlike Study 2, where we filtered out second responses based on identical IP addresses, we used Qualtrics' fraud detection software to filter out responses that were suspicious either because they were likely 1) bots and/or 2) duplicate responses. For all main analyses, we excluded participants who had 1) Q_RecaptchaScore less than .5 (indicating the respondent is likely a bot) 2) Q_RelevantIDDuplicate equal to 1 (indicating the response is likely a duplicate) 3) Q_RelevantIDDuplicateScore greater than or equal to 75 (indicating the response is likely a duplicate) or 4) Q_RelevantIDFraudScore greater than or equal to 30 (indicating the response is likely fraudulent and a bot).

The final dataset consists of 1072 participants (47.2% women), with an average age of 39.33 (SD = 11.88) years. Of the final sample, 45 participants (44.44% women) dropped out of the study before finishing (again, we use their data when available) and 67 participants were flagged by Qualtrics' fraud detection software as suspicious based on the aforementioned criteria.

Demographics Across Conditions			
Characteristic	Control condition, N = 528	Preparation condition, N = 531	**p-value**
Age	39 (12)	40 (12)	0.10
Unknown	13	9	
Gender			>0.9
Man	281 (53%)	281 (53%)	
Woman	247 (47%)	250 (47%)	
Education			0.9
No formal education	0 (0%)	0 (0%)	
Less than high school	3 (0.6%)	3 (0.6%)	
High school graduate (diploma)	57 (11%)	49 (9.4%)	
High school graduate (GED)	14 (2.7%)	13 (2.5%)	
Some college (1-4 years, no degree)	122 (24%)	112 (21%)	
Associate's degree (including occupational or academic degrees)	50 (9.7%)	58 (11%)	
Bachelor's degree (BA, BS, etc)	199 (39%)	207 (40%)	
Master's degree (MA, MS, MENG, MSW, etc)	58 (11%)	70 (13%)	
Professional school degree (MD, DDC, JD, etc)	7 (1.4%)	4 (0.8%)	
Doctorate degree (PhD, EdD, etc)	5 (1.0%)	6 (1.1%)	
Unknown	13	9	

Table 1.15: Size of sample in Study 3 with corresponding percentage listed for gender and education, with p-values derived from Fisher's exact test. Mean with corresponding standard deviation listed for age, with p-values derived from Kruskal-Wallis test. If a participant did not respond to a given question, we list their response as 'Unknown'. Note: we did not include questions about race/ethnicity nor income in this study.

Demographics Based on Participant Gender			
Characteristic	**Man**, N = 566	**Woman**, N = 506	**p-value**
Age	39 (12)	40 (12)	0.10
Unknown	14	21	
Education			0.041
No formal education	0 (0%)	0 (0%)	
Less than high school	4 (0.7%)	2 (0.4%)	
High school graduate (diploma)	60 (11%)	46 (9.5%)	
High school graduate (GED)	14 (2.5%)	13 (2.7%)	
Some college (1-4 years, no degree)	125 (23%)	109 (22%)	
Associate's degree (including occupational or academic degrees)	46 (8.3%)	62 (13%)	
Bachelor's degree (BA, BS, etc)	230 (42%)	176 (36%)	
Master's degree (MA, MS, MENG, MSW, etc)	57 (10%)	71 (15%)	
Professional school degree (MD, DDC, JD, etc)	7 (1.3%)	4 (0.8%)	
Doctorate degree (PhD, EdD, etc)	9 (1.6%)	2 (0.4%)	
Unknown	14	21	

Table 1.16: Size of sample in Study 3 with corresponding percentage listed for education, with p-values derived from Fisher's exact test. Mean with corresponding standard deviation listed for age, with p-values derived from Kruskal-Wallis test. If a participant did not respond to a given question, we list their response as 'Unknown'. Note: we did not include questions about race/ethnicity nor income in this study.

Procedures

As in Studies 1 and 2, participants included in the study were told they would be completing a two-minute multiplication task (identical to the ones used in previous studies) and would be able to choose a payment scheme for their performance. After being told about the rules for the multiplication task and passing the same comprehension questions used in the previous studies, participants were assigned to either an unlimited preparation condition, where they could complete as many practice multiplication problems as they wanted, with the option to opt out of the practice at any time before moving on to the multiplication task, or a control condition, where they were told they could complete as many rounds of a subtraction exercise as they wanted before the multiplication task. An equal number of participants were randomly assigned to both conditions (control= 49.86%), with no significant difference in representation of men and women across conditions, $\chi^2(1, n = 1072) = 0.00$, $p = .971$, confirming there was random assignment to conditions based on gender. Participants across both conditions were given the option to study the multiplication (preparation condition) or subtraction (control condition) tables for as long as they wanted. We measured both the decision to study the respective table within each condition, along with the amount of time that participants who chose to study the tables spent on that page. Next, participants within each condition were given the option to complete problems from the tables within their respective condition. Problems were created by randomly drawing pairs of numbers from 1 to 12 and asking participants to multiply them together (preparation condition) or subtract them (control condition). Participants across both conditions were able to complete 10 problems at a time before being prompted to indicate whether they would like to continue completing problems. Like the previous studies, we measured the decision to practice multiplication problems (or in the case of the control condition, complete subtraction problems), along with the number of rounds of practice participants completed. Here, the number of rounds of practice was encoded as follows: participants who did not choose to practice had a value of zero for this variable, participants who chose to practice had a value of one for this variable, and thereafter the variable increased incrementally in correspondence with each round of practice participants completed. This variable had a mean of 0.33, with an *SD* of 1.03.

Afterwards, all participants chose a payment scheme for the multiplication task, described and counterbalanced in the same way as in Studies 1 and 2. Then, participants completed the paid multiplication task for two minutes. We included many of the same follow-up questions as in Studies 1 and 2, including measures of risk attitudes, confidence, perceptions of gender differences in preparation on the multiplication task (for participants in the preparation condition), perceptions of gender differences in preparation in general, competitiveness, and performance. Notably, we added in an additional response option for all of the questions about perceptions of gender differences in behavior, such that participants in this study could choose between saying men were more likely to engage in the specified behavior, women were more likely to engage in the specified behavior, or that there would be no gender differences in the specified behavior. Like before, participants were incentivized to answer the questions about their confidence and perceptions of gender differences correctly, and were paid at the same rate as Studies 1 and 2. Participants also completed a manipulation check, where they were told about the two conditions, and were asked which of the conditions they thought was more helpful in boosting scores on the paid multiplication task. Finally, they completed some demographic questions and provided feedback on the study before being paid for their participation.

1.4.2 Results

Describing main variables of interest

We start by exploring gender differences across the main variables of interest to compare the characteristics of our sample to previous samples in the literature and provide context for the subsequent analyses. First, we replicated the effect from the previous studies of gender on the choice to compete when gender is included as the only predictor in the logistic regression: 20.22% of men chose to compete compared to 8.85% of women, $b = -0.96$, 95% CI [−1.34, −0.59], $z = -5.01$, $p < .001$. However, like the previous two studies, when running regressions including control variables (i.e., task score, risk attitudes, confidence, and the interaction between gender and condition), we find that the effect of gender on the choice to compete is not significant, $b = -0.15$, 95% CI [−0.69, 0.37], $z = -0.56$, $p = .573$, suggesting the effect is explained fully by risk attitudes, $b = 0.33$, 95% CI [0.25, 0.41], $z = 7.79$, $p < .001$, task score, $b = 0.02$, 95% CI [0.01, 0.03], $z = 3.84$, $p < .001$, and confidence, $b = 0.01$, 95% CI [0.00, 0.02],

$z = 2.17$, $p = .030$ (See Table 1.17). On the other hand, we replicate effects from the literature of gender on both confidence, $b = -14.06$, 95% CI $[-16.73, -11.39]$, $t(1035) = -10.35$, $p < .001$, and risk attitudes, $b = -1.02$, 95% CI $[-1.33, -0.71]$, $t(1035) = -6.48$, $p < .001$, where women tend to be less confident with regards to their performance on the task and generally more risk averse relative to men.

	(1)			(2)			(3)		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.25	0.21 – 0.31	**<0.001**	0.27	0.20 – 0.36	**<0.001**	0.01	0.00 – 0.02	**<0.001**
gender [Woman]	0.38	0.26 – 0.55	**<0.001**	0.51	0.31 – 0.81	**0.006**	0.86	0.50 – 1.45	0.573
condition [pract]				0.87	0.58 – 1.32	0.526	0.81	0.52 – 1.27	0.359
gender [Woman] * condition [pract]				0.51	0.23 – 1.10	0.089	0.55	0.24 – 1.24	0.156
task score							1.02	1.01 – 1.03	**<0.001**
risk							1.39	1.28 – 1.51	**<0.001**
conf rank							1.01	1.00 – 1.02	**0.030**
Observations	1040			1040			1037		
R^2 Tjur	0.025			0.030			0.156		

Table 1.17: All models are logistic regressions with choice to compete as the dependent variable, where man and control are the reference categories for participant gender and preparation condition, respectively. The gender difference in the choice to compete is not reduced by preparation condition, but is explained by risk attitudes, confidence, and task scores. $p < .05$ is considered significant and bolded.

Another important consideration when interpreting any main effects found in this study is whether there are gender differences in task scores. When including gender by itself as a predictor of performance, we find that women have significantly lower task scores, $b = -7.60$, 95% CI $[-10.07, -5.12]$, $t(1035) = -6.01$, $p < .001$. However, when controlling for competition choice, confidence, and risk attitudes, the effect of gender is not significant, $b = -1.35$, 95% CI $[-3.79, 1.09]$, $t(1031) = -1.09$, $p = .278$, suggesting that these variables explain the gender difference in task scores. Specifically, we find that both confidence, $b = 0.43$, 95% CI $[0.38, 0.49]$, $t(1031) = 16.84$, $p < .001$, and competition choice, $b = 8.23$, 95% CI [4.42, 12.03], $t(1031) = 4.24$, $p < .001$, positively predict task scores (that is, those who are more confident and chose to compete tend to have higher task scores), while risk attitudes negatively predict task scores (that is, individuals who are more risk seeking tend to have lower scores), $b = -1.14$, 95% CI $[-1.58, -0.69]$, $t(1031) = -4.98$, $p < .001$ (See Table 1.19). See Table 1.18 for a summary of gender differences in the main variables of interest.

Variable	**Man**, N = 566[1]	**Woman**, N = 506[1]
task_score	48 (35, 65)	41 (31, 53)
comp_choice		
piecerate	442 (80%)	443 (91%)
tournament	112 (20%)	43 (8.8%)
practice_problems_binary	107 (19%)	104 (21%)
risk	5.00 (3.00, 7.00)	4.00 (2.00, 6.00)
conf_rank	60 (50, 80)	50 (30, 60)

[1] Median (IQR); n (%)

Table 1.18: Gender differences in the main variables of interest, including: task scores, choice to compete, choice to practice (among the full dataset), confidence, and risk attitudes. Medians are reported for task score, risk attitudes, and confidence, with IQRs in parentheses. For choice to practice and choice to compete, we report the number and percentage of participants that fall into each category for each respective gender.

		(1)			**(2)**			**(3)**	
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	51.57	49.87 – 53.26	**<0.001**	49.31	47.44 – 51.18	**<0.001**	29.30	25.55 – 33.06	**<0.001**
gender [Woman]	-7.60	-10.07 – -5.12	**<0.001**	-5.84	-8.48 – -3.20	**<0.001**	-1.35	-3.79 – 1.09	0.278
comp choice [tournament]				11.21	7.04 – 15.38	**<0.001**	8.23	4.42 – 12.03	**<0.001**
gender [Woman] * comp choice [tournament]				-5.59	-13.13 – 1.95	0.146	-4.15	-10.82 – 2.51	0.222
conf rank							0.43	0.38 – 0.49	**<0.001**
risk							-1.14	-1.58 – -0.69	**<0.001**
Observations	1037			1037			1037		
R^2 / R^2 adjusted	0.034 / 0.033			0.062 / 0.059			0.269 / 0.265		

Table 1.19: All models are linear regressions with task score as the dependent variable, where man and piece-rate payment scheme are the reference categories for participant gender and competition choice, respectively. After controlling for risk attitudes, confidence, and competition choice, women no longer have lower scores on the multiplication task than men, p < .05 is considered significant and bolded.

Effects of unlimited preparation condition on gender differences in choice to compete

We do not find evidence of a significant effect of preparing with multiplication or subtraction problems (task relevant vs. irrelevant preparation condition) on the choice to compete across all participants, $b = -0.33$, 95% CI $[-0.68, 0.01]$, $z = -1.88$, $p =$ .061 when included as a single predictor of the choice to compete in a logistic regression. In a subsequent logistic regression adding in gender and the interaction between gender and condition as predictors, we find that gender is the only significant predictor of the choice to compete, $b = -0.68$, 95% CI $[-1.18, -0.21]$, $z = -2.77$, $p = .006$ and no evidence of the expected interaction effect between gender and condition on the choice to compete, $b = -0.68$, 95% CI $[-1.48, 0.09]$, $z = -1.70$, $p = .089$ (see Figure 1.17).[5]

Gender differences in preparation

Our next set of analyses focused on the effects of gender on decisions to practice. Thus, all subsequent analyses focus on the subset of participants that were assigned to the unlimited preparation condition that were actually given the opportunity to practice beforehand (N = 571).

We do not replicate the effect found in both previous studies of gender on the decision to practice multiplication problems, neither in the model where gender is included by itself as a sole predictor of the choice to practice, $b = 0.11$, 95% CI $[-0.27, 0.50]$, $z = 0.58$, $p = .565$ (see right panel of Figure 1.18), nor in tandem with the choice to compete and the interaction between gender and the choice to compete

[5]The interaction between gender and preparation condition on the choice to compete is still not significant among the full dataset (i.e., after including participants that were flagged by Qualtrics' fraud detection software): $b = -0.54$, 95% CI $[-1.27, 0.18]$, $z = -1.45$, $p = .148$

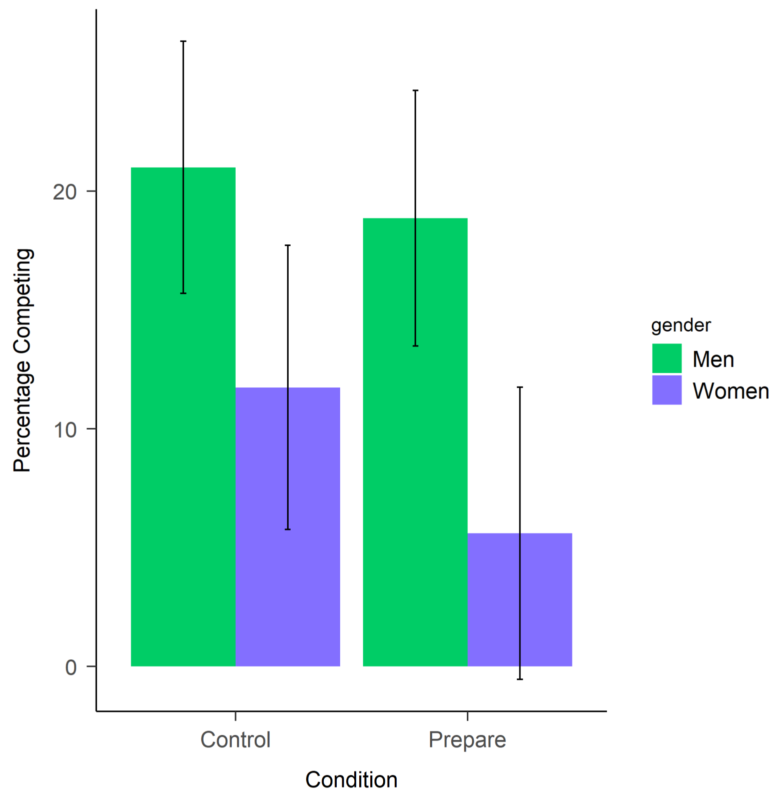

Figure 1.17: Proportion of men and women in Study 3 who chose to compete based on preparation condition. Unlimited preparation did not reduce the gender difference in competitiveness. Error bars represent standard errors.

as predictors, $b = 0.34$, 95% CI $[-0.09, 0.77]$, $z = 1.55$, $p = .121$ (see left panel of Figure 1.18).[6] Though the effect is not significant, the proportion of women that chose to practice was still higher than the proportion of men that chose to practice (26.87% of women chose to prepare via practice, relative to 24.5% of men) among participants within the unlimited practice condition. Thus, the direction of the effects are not in contrast with previous studies. In line with the previous studies, we do not find an interaction between gender and choice to compete on the choice to practice, $b = -1.22$, 95% CI $[-2.84, 0.15]$, $z = -1.64$, $p = .101$. Instead we find that, like previous studies, the choice to compete is positively related to the choice to practice,

[6]The effect of gender on the choice to prepare is still not significant among the full dataset (i.e., after including participants that were flagged by Qualtrics' fraud detection software): $b = 0.12$, 95% CI $[-0.25, 0.50]$, $z = 0.64$, $p = .519$

$b = 0.88$, 95% CI [0.23, 1.51], $z = 2.70$, $p = .007$. In adding confidence and risk attitudes as predictors to the model with the interaction effect on top of the main effects, we do not find evidence that either of those predictors are significantly related to the choice to practice either (See Table 1.20). Like the previous two studies in Chapter 1, we also ran a two-part hurdle model with gender, competition choice, and the interaction between those variables predicting the number of practice rounds variable (among participants in the practice condition) and again do not find evidence of gender differences in the choice to continue preparing after the initial decision to prepare, , $b = 0.42$, 95% CI [-0.31, 1.15], $z = 1.14$, $p = 0.25$.

Unlike the previous studies, this study separated the decision to study tables and amount of time studying tables from the decision to practice and number of problems completed, so we had the novel opportunity to explore questions about gender differences in studying here. We do not find evidence that there are gender differences in the decision to study the multiplication tables, $b = 0.09$, 95% CI [-0.28, 0.47], $z = 0.49$, $p = .624$. However, among participants who did choose to study the multiplication tables (N = 234; 40.98% of participants in the unlimited preparation condition), we find that women studied for more time (in seconds) than men on average, $M_{women} = 33.01$, $SD_{women} = 72.74$; $M_{men} = 18.75$, $SD_{men} = 25.24$, $b = 21.94$, 95% CI [2.65, 41.23], $t(211) = 2.24$, $p = .026$.[7]

[7]However, we note that the SD for studying time among women is exceptionally high relative to that of men, so the effect may be driven by select outliers pulling the mean

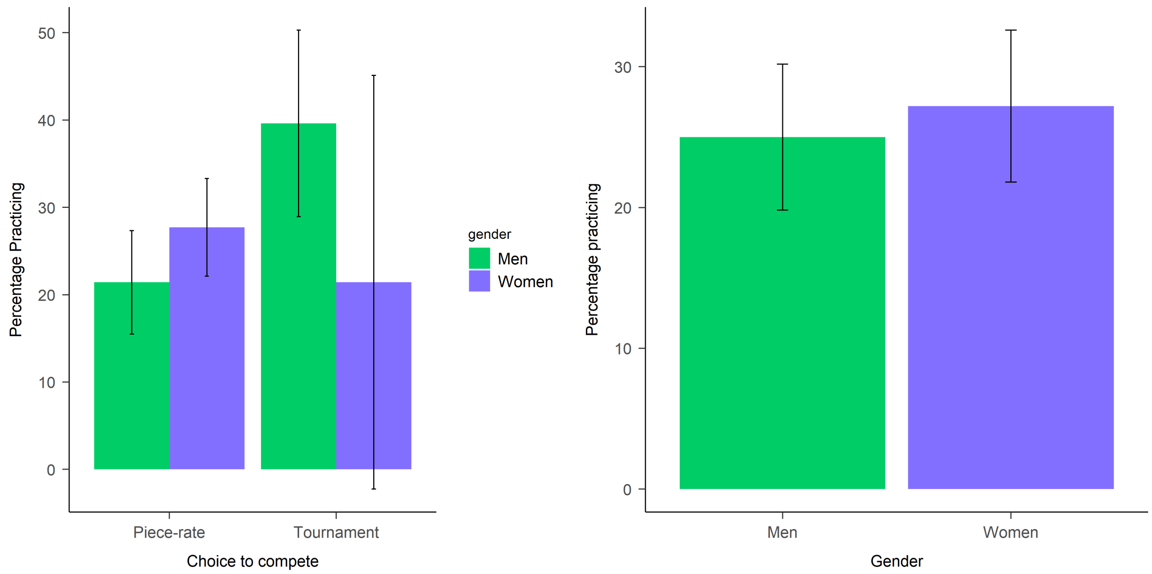

Figure 1.18: Right panel shows the proportion of men and women in Study 3 who chose to prepare. Left panel shows the proportion of men and women in Study 3 who chose to prepare based on choice to compete. Error bars represent standard errors.

	(1)			(2)			(3)		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.24	0.19 – 0.29	**<0.001**	0.27	0.20 – 0.37	**<0.001**	0.24	0.13 – 0.43	**<0.001**
gender [Woman]	1.13	0.83 – 1.52	0.442	1.41	0.92 – 2.17	0.121	1.29	0.91 – 1.84	0.153
comp choice [tournament]				2.41	1.26 – 4.53	**0.007**	1.90	1.13 – 3.14	**0.013**
gender [Woman] * comp choice [tournament]				0.30	0.06 – 1.16	0.101	0.36	0.13 – 0.92	**0.041**
conf rank							1.00	0.99 – 1.00	0.278
task score							1.00	0.99 – 1.01	0.753
risk							1.04	0.97 – 1.11	0.264
Observations	1057			522			1037		
R^2 Tjur	0.001			0.015			0.010		

Table 1.20: All models are logistic regressions with choice to prepare as the dependent variable, where man and piece-rate payment scheme are the reference categories for participant gender and competition choice, respectively. We do not find evidence that women prepare more than men. $p < .05$ is considered significant and bolded.

Perceptions of gender differences in preparation, performance, and competitiveness

Like all studies before, for each question about perceptions of gender differences, we run a chi-square goodness of fit test with the null hypothesis that participants' will choose each option at a similar rate. Since participants were given the option in this study to select one of three response options, rather than two options like the first two studies, we first perform a chi-square goodness of fit test with all response options to see if they are all equally likely. If the test with all three response options was significant, we then performed more targeted chi-square goodness of fit tests with pairs of response options within a given question to test which specific pairs of response options are significantly different. See Table 1.21 for a summary of participants' responses to the questions about gender differences in preparation, performance, and competitiveness.

When asked to predict gender differences in preparation for the multiplication task,[8] we find that participants' responding is not evenly distributed across the response options, $\chi^2(2, n = 571) = 192.97$, $p < .001$. In performing more targeted analyses, we find that participants were significantly more likely to choose "women" (56.7%) than "men" (7.28%), $\chi^2(1, n = 571) = 199.29$, $p < .001$, or "no difference" (36.02%), $\chi^2(1, n = 571) = 24.10$, $p < .001$ in response to this question (see Figure 1.19).

This result replicates when participants are asked about gender differences in general tendencies to prepare, $\chi^2(2, n = 1072) = 358.38$, $p < .001$ (see Figure 1.22), such that, across all participants, a significantly higher proportion of participants said women prepare more in general (59.5%) than the proportion of participants that said men prepare more (12.34%), $\chi^2(1, n = 1072) = 320.97$, $p < .001$, or there are no gender differences in general tendencies to prepare (28.16%), $\chi^2(1, n = 1072) = 116.20$, $p < .001$.

Though participants consistently expected women to be more likely to prepare than men, they did not expect that there would be a gender difference in performance,

[8]Note: this question was only asked among participants in the preparation condition

$\chi^2(2, n = 1072) = 77.76$, $p < .001$ (see Figure 1.20), where participants were significantly more likely to indicate "no difference" (46.09%) compared to "men" (28.64%), $\chi^2(1, n = 1072) = 42.27$, $p < .001$, or "women" (25.27%), $\chi^2(1, n = 1072) = 63.05$, $p < .001$, in response to this question.

Despite thinking that there would be no gender difference in performance and expecting women to prepare for the task more, participants consistently expected men to be more likely to compete (78.01%), $\chi^2(2, n = 1072) = 961.58$, $p < .001$ (see Figure 1.21), rather than expecting women to compete more (4.05%), $\chi^2(1, n = 1072) = 691.29$, $p < .001$, or expecting no difference in willingness to compete across genders, (17.94%), $\chi^2(1, n = 1072) = 390.08$, $p < .001$. We discuss these findings about participants' beliefs in light of the actual study results in the discussion section.

Characteristic	N = 1,072[1]
better_gender_guess	
Men	297 (29%)
No difference	478 (46%)
Women	262 (25%)
perc_task_gender_pract	
Men	38 (7.3%)
No difference	188 (36%)
Women	296 (57%)
perc_gender_comp	
Men	809 (78%)
No difference	186 (18%)
Women	42 (4.1%)
perc_gen_gender_pract	
Men	128 (12%)
No difference	292 (28%)
Women	617 (59%)

[1] n (%)

Table 1.21: Number and percentage of participants that selected each respective option when asked whether men or women would correctly solve more problems on the multiplication task, spend more time preparing for the multiplication task, choose the tournament payment scheme more often, and spend more time preparing on most tasks. Participants were also given the option in this study to indicate there would be no gender difference for any of the variables.

Effects of gender and perceptions on practicing

Like the previous two studies, we explored whether women who believed other women prepare more were especially likely to prepare. To that end, we ran the same

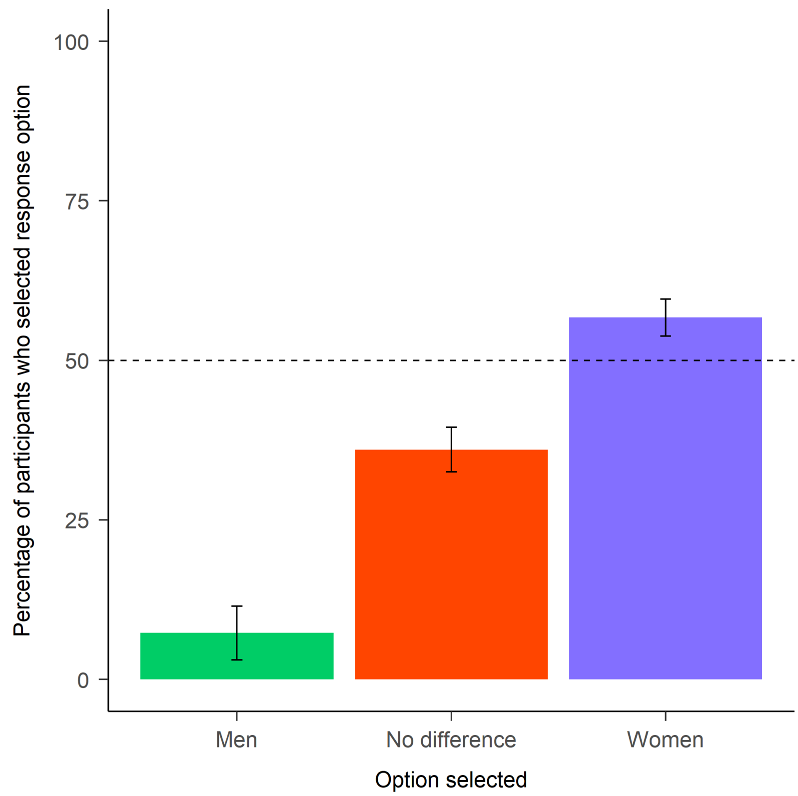

Figure 1.19: Proportion of participants that predicted women would spend more time preparing for the multiplication task, men would spend more time preparing for the multiplication task, or that there would be no gender differences in preparation for the task. A significantly larger proportion of participants expected women to spend more time preparing for the multiplication task. Error bars represent standard errors.

logistic regression with the choice to practice as the dependent variable and gender, beliefs about gender differences in preparation on most tasks, and the interaction between those two variables as the predictors. Notably, the predictors representing participants' beliefs in these models had three possible values: women, men, and no difference. Like the previous studies, we focus on the interaction effect between gender and participants' selecting women as the gender that spends more time preparing while controlling for all other effects in the model. We do not replicate the interaction effect found in the previous studies, such that women who said women generally prepare more on most tasks were not significantly more likely to prepare, $b = 0.42$, 95% CI $[-0.91, 1.82]$, $z = 0.60$, $p = .545$ (see Figure 1.23). We ran the same analysis with

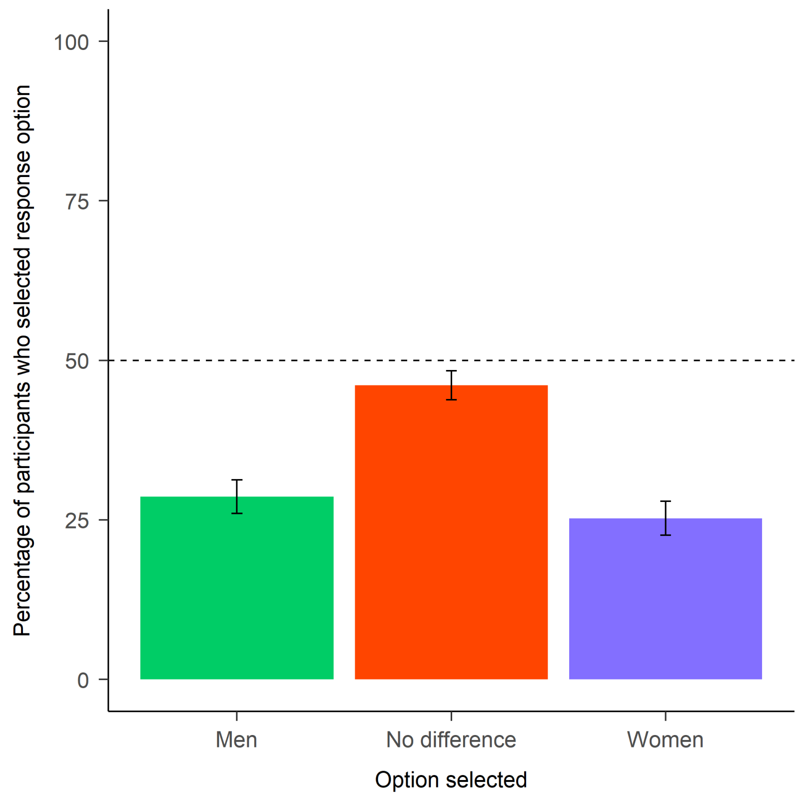

Figure 1.20: Proportion of participants that predicted women would correctly solve more problems on the multiplication task, men would correctly solve more problems on the multiplication task, or that there would be no gender difference in performance on the multiplication task. A significantly larger proportion of participants expected there to be no gender difference in performance on the multiplication task. Error bars represent standard errors.

participants' beliefs about gender differences in preparation for the multiplication task, gender, and the interaction between the two as predictors instead, and replicate the interaction effect from the previous studies, $b = 2.40$, 95% CI [0.95, 3.99], $z = 3.13$, $p = .002$. However, in this case, the interaction effect appears to be driven by men being especially likely to prepare when they think men spent more time preparing for the multiplication task (see Figure 1.24).

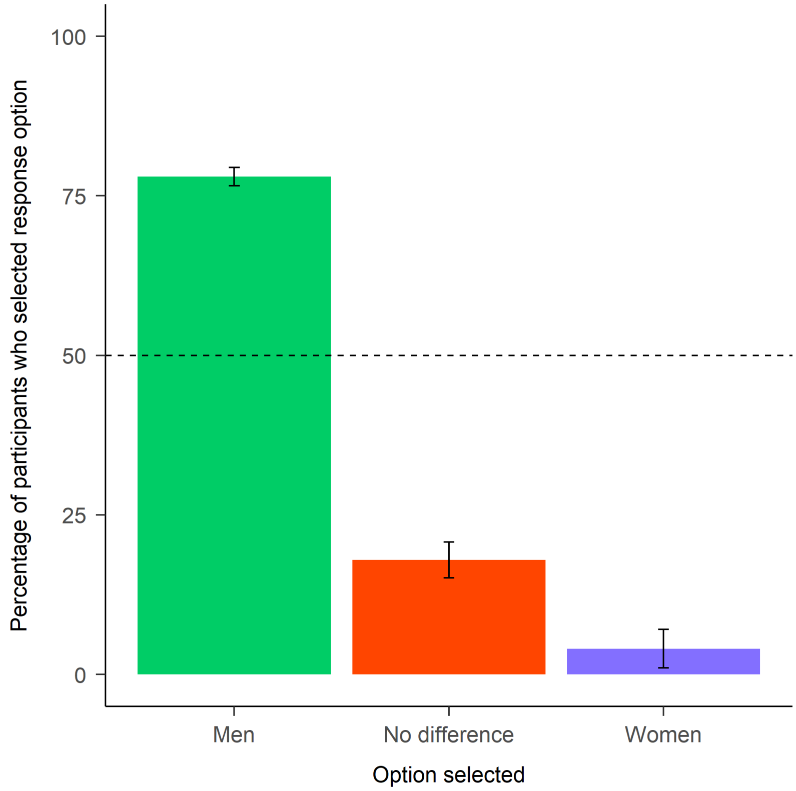

Figure 1.21: Proportion of participants that predicted women would choose the tournament payment scheme more often, men would choose the tournament payment scheme more often, or there would be no gender differences in the choice to compete. A significantly larger proportion of participants expected men to be more likely to choose to compete. Error bars represent standard errors.

Validating the perceived utility of preparation condition over control condition

Since we did not find any effect of the unlimited preparation condition, and this study uses a novel control task that has not been used in previous studies, we ran additional analyses to test whether participants actually felt like the preparation condition was more useful for performance on the paid multiplication task than the control condition. It is possible we did not observe an effect of condition on the choice to compete, nor an interaction effect between gender and condition, because participants simply did not think the preparation task would be more helpful for improving their performance relative to the control condition.

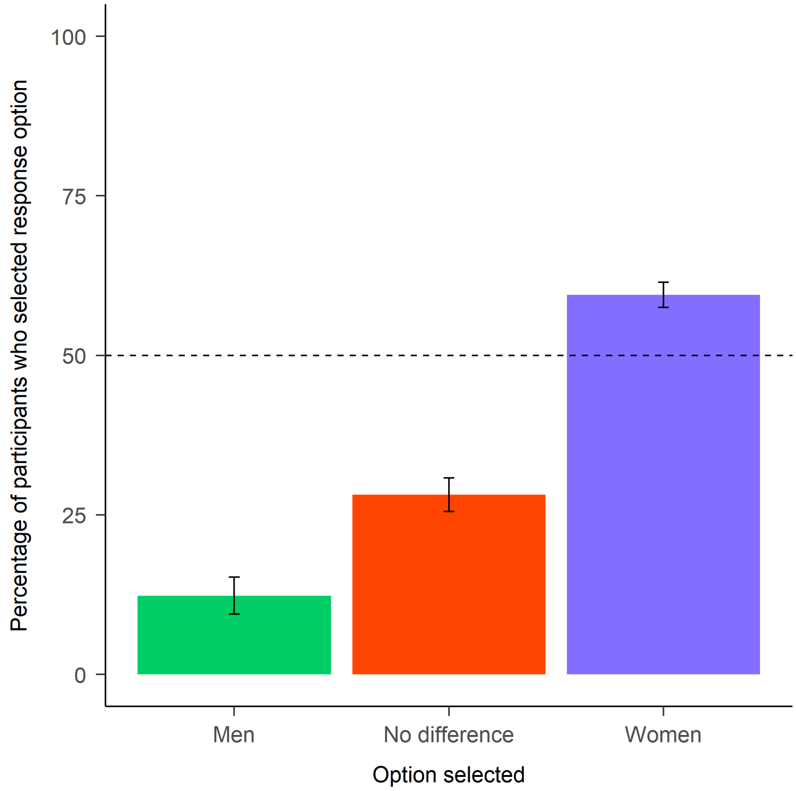

Figure 1.22: Proportion of participants that predicted women prepare more in general, men prepare more in general, or that there are no gender differences in preparation in general. A significantly larger proportion of participants expected women prepare more in general. Error bars represent standard errors.

To explore this null effect further, we looked into how much participants were choosing to practice across conditions, which may provide insight into why there was no difference in the choice to compete across conditions. It is possible that participants in the control condition decided to complete subtraction problems at similar rates as participants in the preparation condition decided to complete multiplication problems, and if so, this may have led them to compete at similar rates because the subtraction tables felt easier than the multiplication tables, and therefore boosted their confidence or reduced perceptions of risk. Contrary to this possibility, we find that participants in the practice condition tended to choose to study, $b = 0.69$, 95% CI [0.43, 0.95], $z = 5.20$, $p < .001$, and practice, $b = 0.78$, 95% CI [0.47, 1.10], $z = 4.89$, $p < .001$, the multiplication tables at significantly higher rates relative to participants in the

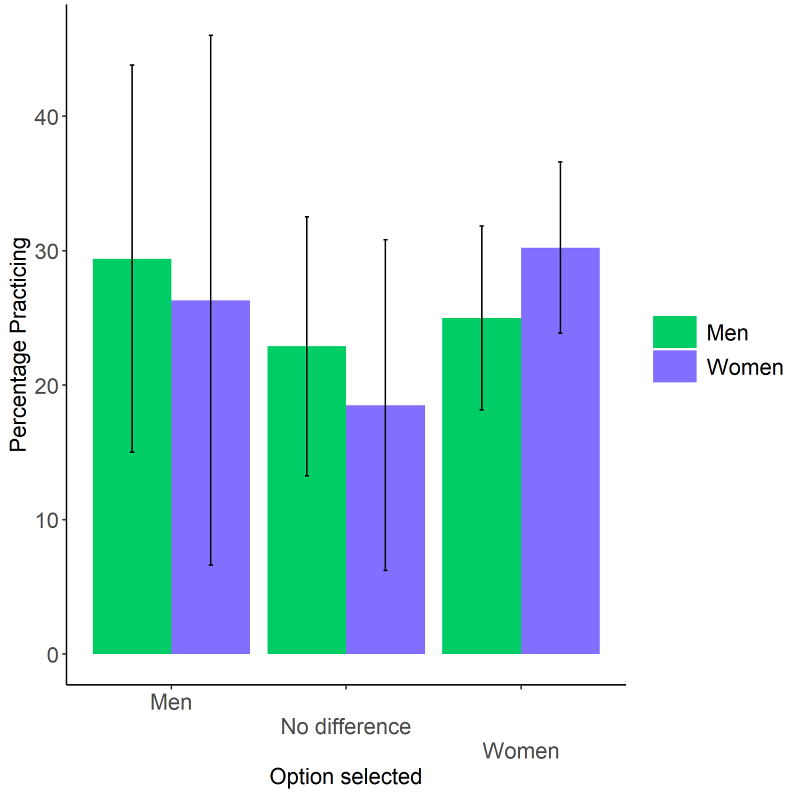

Figure 1.23: Proportion of men and women in Study 3 who chose to practice based on whether they thought men or women spend more time preparing on most tasks. In this study, participants also had the option to say there was no gender difference in preparation. It is worth noting that a small sample of participants (N = 136 participants chose to prepare across the entire study) are represented in this graph. Error bars represent standard errors.

control condition where they completed subtraction problems.

In further evidence of the perceived utility of the preparation condition for improving performance on the paid multiplication task, participants across both conditions tended to believe when asked in the manipulation check that practicing multiplication problems would be more likely to improve performance on the paid multiplication task than practicing subtraction problems, $\chi^2(1, n = 1072) = 548.50$, $p < .001$.

1.4.3 Discussion

In this final test of the effects of the opportunity to prepare on the choice to compete, we explored whether having unlimited time to prepare would reduce the gender gap in

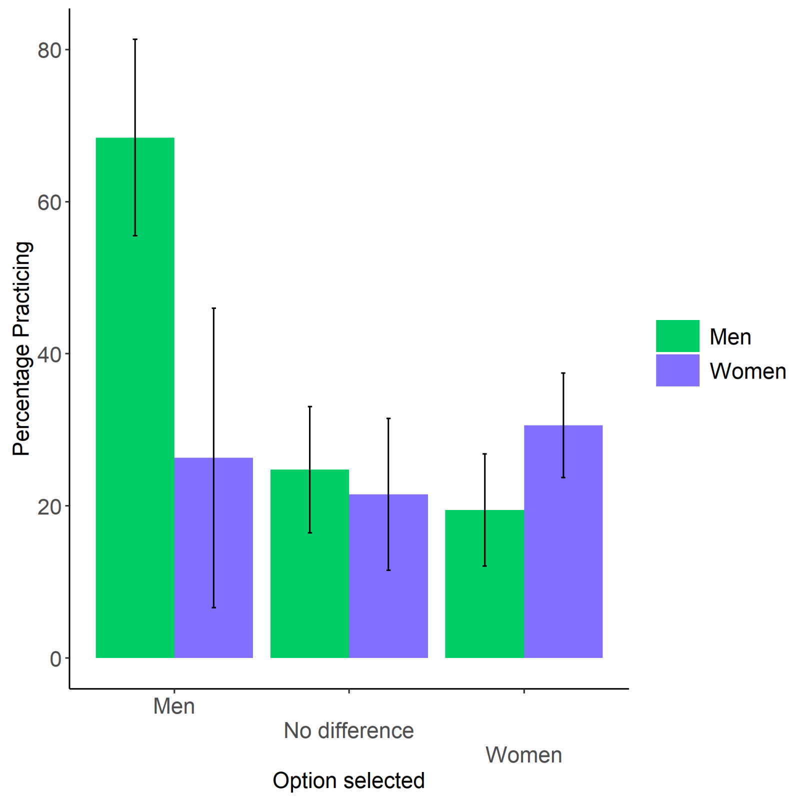

Figure 1.24: Proportion of men and women in Study 3 who chose to practice based on whether they thought men or women spend more time preparing for the multiplication task. In this study, participants also had the option to say there was no gender difference in preparation. It is worth noting that a small sample of participants (N = 136 participants chose to prepare across the entire study) are represented in this graph. Error bars represent standard errors.

competitiveness. First, we test whether there is a gender difference in competitiveness, which we do find again within the context of this design. That is, women do indeed compete less than men when gender is included as a predictor by itself in the model - but the fact that the effect disappears when controlling for risk attitudes, confidence, and task scores provides further evidence that competitiveness may not be a stand-alone trait. We also replicate the effects of gender on risk attitudes and confidence. In this study, we do not find evidence of a gender difference in task score when controlling for confidence and risk attitudes, like Study 2. Thus, Study 1 is the only study thus far in the dissertation that has found a gender difference in task scores that does not appear to be explained by gender differences in risk attitudes or confidence.

Summary of main experimental results

Next, we explored the main experimental results for the intervention attempting to reduce the gender difference. We hypothesized that being assigned to the unlimited preparation condition that closely resembled the paid task would lead participants to choose to compete more often than participants who completed the control subtraction task that was not relevant to the paid task. In support of this assumption, when asked which condition they thought would be most helpful for improving performance on the multiplication task during the manipulation check, participants resoundingly chose the preparation condition over the control condition. Participants' responses on the manipulation check aligned with their choices to prepare across conditions, such that participants in the unlimited preparation condition were significantly more likely to choose to study and practice relative to participants in the control condition. Overall, we have evidence that the subtraction task was an appropriate choice for the control condition, as participants thought that it would be less likely to help their performance on the paid task relative to the multiplication task. Despite both behavioral and self-report evidence that participants thought the unlimited preparation condition would be more helpful for performance, we did not find evidence that condition assignment affected participants' choices to compete. Yet again, contrary to our hypothesis of an interaction between gender and condition, the effect of the unlimited opportunity to prepare before completing the paid task did not differ by gender.

Gender differences in preparation

Like the previous studies, the gender difference in preparation was also an important part of our analyses. Yet, in this study we do not find evidence of a gender difference in the choice to prepare. There are a few possible reasons we did not replicate the effect of gender on the choice to prepare. The discrepancy could be explained by the lower number of participants in the preparation condition relative to previous studies given the random assignment to condition. For this reason, we may have had less power in estimating the effects of gender on the choice to prepare. It is also possible that differences in the design of this study relative to the previous two studies drove the differences, which will be further explored in the overall discussion section.

In this study, we changed the nature of preparation - such that we had the novel opportunity to explore different ways of measuring preparation, including the choice to study multiplication tables and the amount of time a participant spends studying multiplication tables. With these new measures, we find that, among participants who chose to study the tables, women spend more time on the page (presumably studying, although we cannot directly attest to that, given the online nature of the experiment), though they did not choose to study at significantly higher rates than men. Thus, we have mixed evidence based on these new measures that women prepare more than men.

Perceptions of gender differences in preparation, performance, and competitiveness

Like the previous two studies in this dissertation, we explored participants' perceptions of gender differences to see whether there tends to be consistency in beliefs about gender differences in performance on the multiplication task, along with gender differences in the choice to prepare and compete, and if so, whether these beliefs aligned with the actual results found. Importantly, participants were again incentivized to guess the study results correctly to reduce the likelihood that participants would respond in a socially desirable way.

Notably, we added a third response option to the measures in this study, such that participants could indicate that they did not expect a gender difference in a given behavior. The third response option served as a test of robustness, since one could argue that the two-response option measures used in previous studies force participants to choose between indicating men or women when they would have otherwise indicated there was no gender difference in a given behavior.

Yet, contrary to that argument, the results mirror results found in previous studies, suggesting that there are consistent beliefs about gender differences in competitiveness and preparation, despite the consistency in beliefs about a *lack* of gender differences in performance.

Like the previous studies, we also explored whether these perceptions are aligned with individuals' practicing behavior. However, we do not find evidence that women's preparation behaviors were related to their beliefs about gender differences in preparation. It is entirely possible that the smaller proportion of participants that even had the opportunity to prepare in this study (per random assignment of preparation condition) did not provide sufficient power to test the interaction effect. In fact, out of the sample of participants in the preparation condition, only a total of 136 participants agreed to prepare, which may be too small of a sample to detect the interaction effect.

Summary

Overall, we find from Study 3 of Chapter 1 that our manipulation of the unlimited opportunity to prepare did not reduce the gender difference in competitiveness. Like Studies 1 and 2, we show that the gender difference in competitiveness itself is explained by gender differences in risk attitudes and confidence. However, we do not replicate the effect that women chose to prepare more than men. Even still, participants believed that women prepared more than men both in general and on the multiplication task, even after we added a third response option where participants could indicate there were no gender differences in preparation. We do not find evidence that women's preparation behaviors are related to their perceptions of gender differences in preparation, though we argue that we may have not had enough power for testing the interaction effect in the current study.