results

Keana Richards

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 @ref(tab:tab-comp-choice-study4)). 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: 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 @ref(tab:tab-task-scores-study4)). See Table @ref(tab:summary-table-gender-study4) 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 2: 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		

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

Table 3: 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.

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 @ref(fig:s300)).[1]

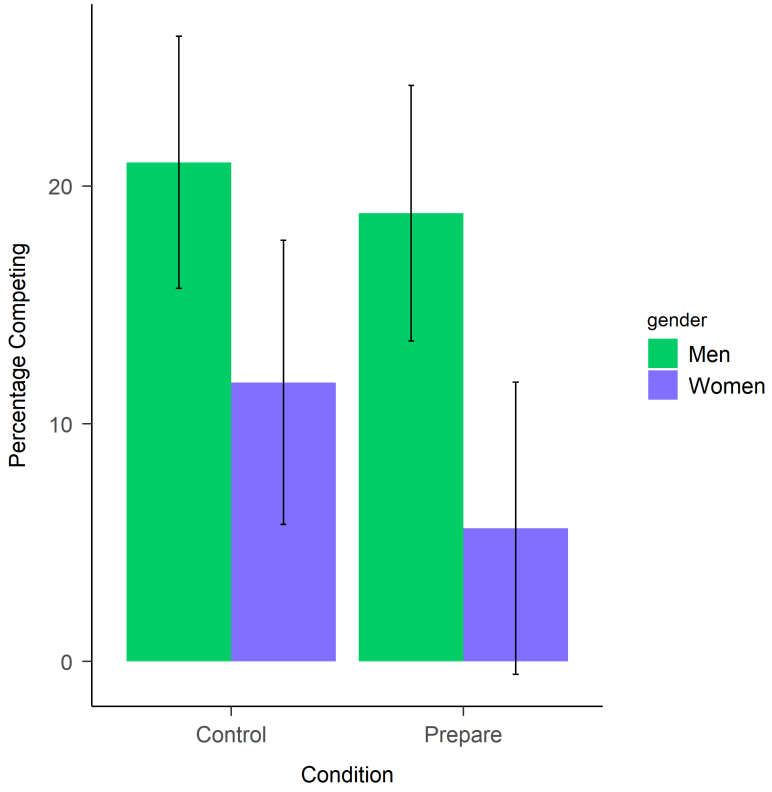

Figure 1: Proportion of men and women in Study 3 who chose to compete based on preparation condition. Unlimited preparation did not reduce the gender difference in competitiveness. Error bars represent standard errors.

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,

[1]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], $z = -1.45$, $p = .148$

$b = 0.11$, 95% CI [−0.27, 0.50], $z = 0.58$, $p = .565$ (see right panel of Figure @ref(fig:panel-study4)), nor in tandem with the choice to compete and the interaction between gender and the choice to compete as predictors, $b = 0.34$, 95% CI [−0.09, 0.77], $z = 1.55$, $p = .121$ (see left panel of Figure @ref(fig:panel-study4)).[2] 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, $b = 0.88$, 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 @ref(tab:tab-pract-choice-study4)). 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], $t(211) = 2.24$, $p = .026$.[3]

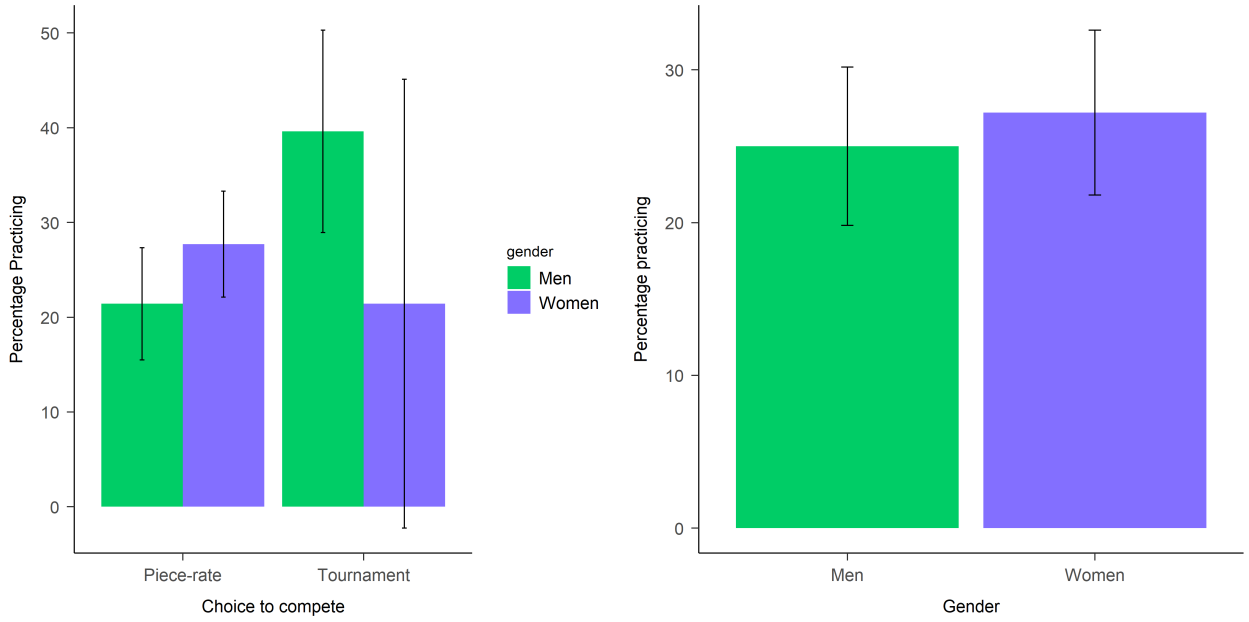

Figure 2: Right panel shows the proportion of men and women in Study 3 who chose to prepare. Left panel shows the proportion of men and women in Study 3 who chose to prepare based on choice to compete. Error bars represent standard errors.

[2]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], $z = 0.64$, $p = .519$

[3]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

	(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 4: 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 @ref(tab:summary-table-beliefs-study4) 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,[4] 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 @ref(fig:s303)).

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 @ref(fig:s306)), 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, $\chi^2(2, n = 1072) = 77.76$, $p < .001$ (see Figure @ref(fig:s304)), 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.

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

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 @ref(fig:s305)), 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 5: 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.

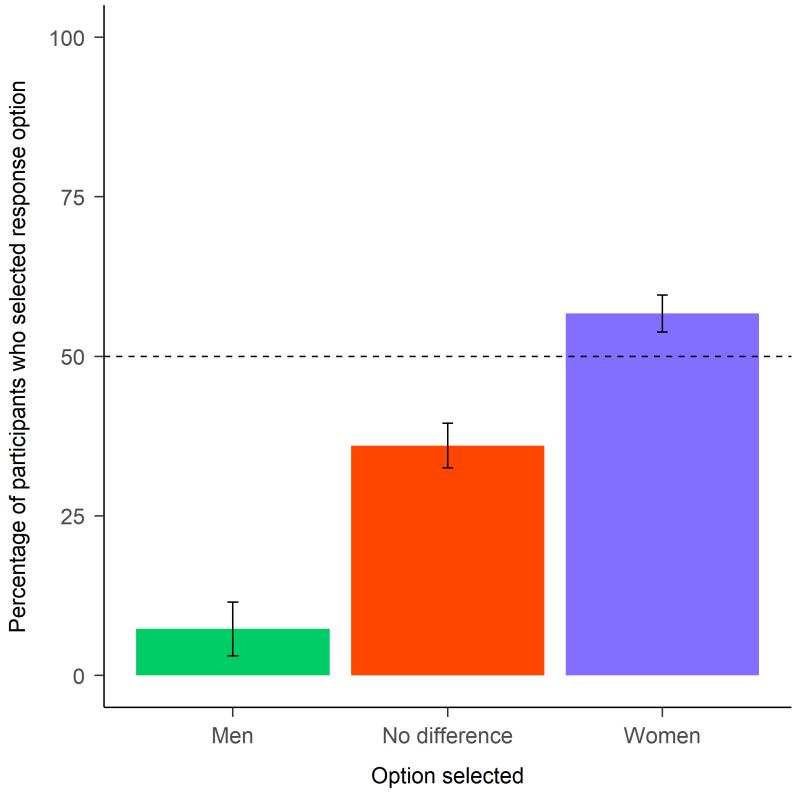

Figure 3: Proportion of participants that predicted women would spend more time preparing for the multiplication task, men would spend more time preparing for the multiplication task, or that there would be no gender differences in preparation for the task. A significantly larger proportion of participants expected women to spend more time preparing for the multiplication task. Error bars represent standard errors.

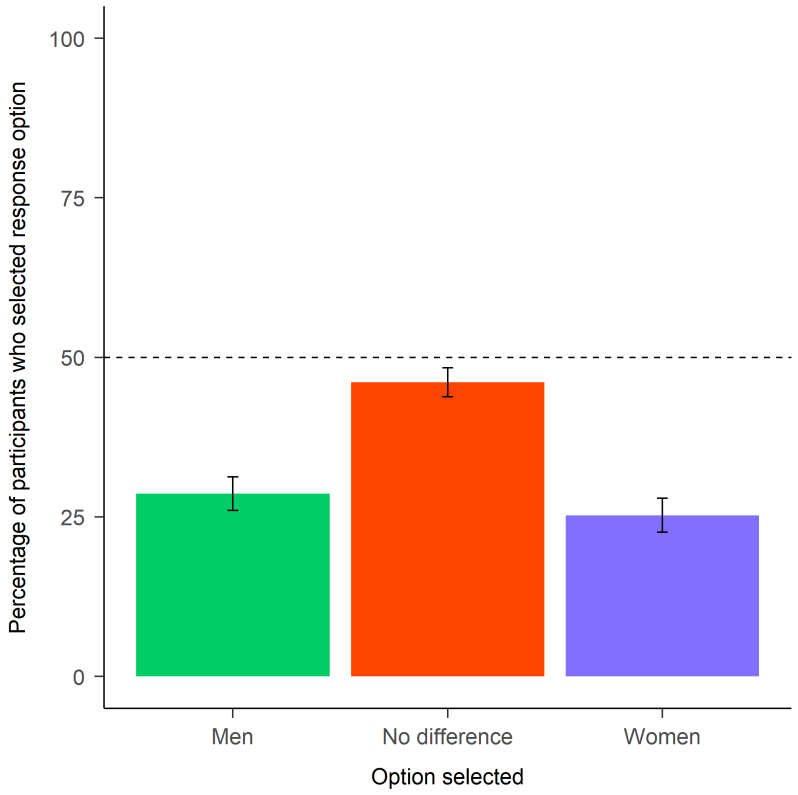

Figure 4: Proportion of participants that predicted women would correctly solve more problems on the multiplication task, men would correctly solve more problems on the multiplication task, or that there would be no gender difference in performance on the multiplication task. A significantly larger proportion of participants expected there to be no gender difference in performance on the multiplication task. Error bars represent standard errors.

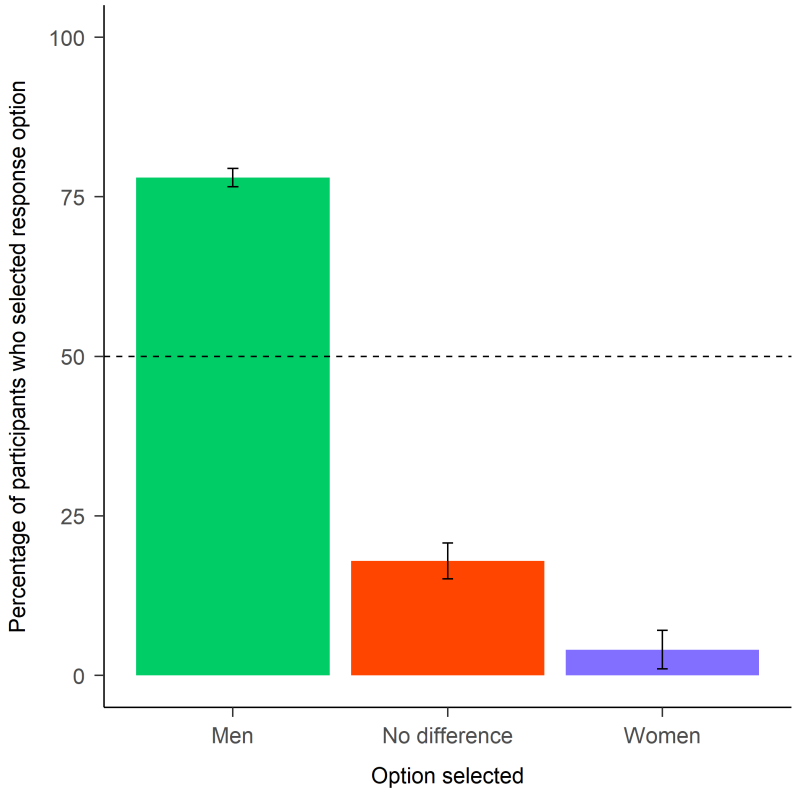

Figure 5: Proportion of participants that predicted women would choose the tournament payment scheme more often, men would choose the tournament payment scheme more often, or there would be no gender differences in the choice to compete. A significantly larger proportion of participants expected men to be more likely to choose to compete. Error bars represent standard errors.

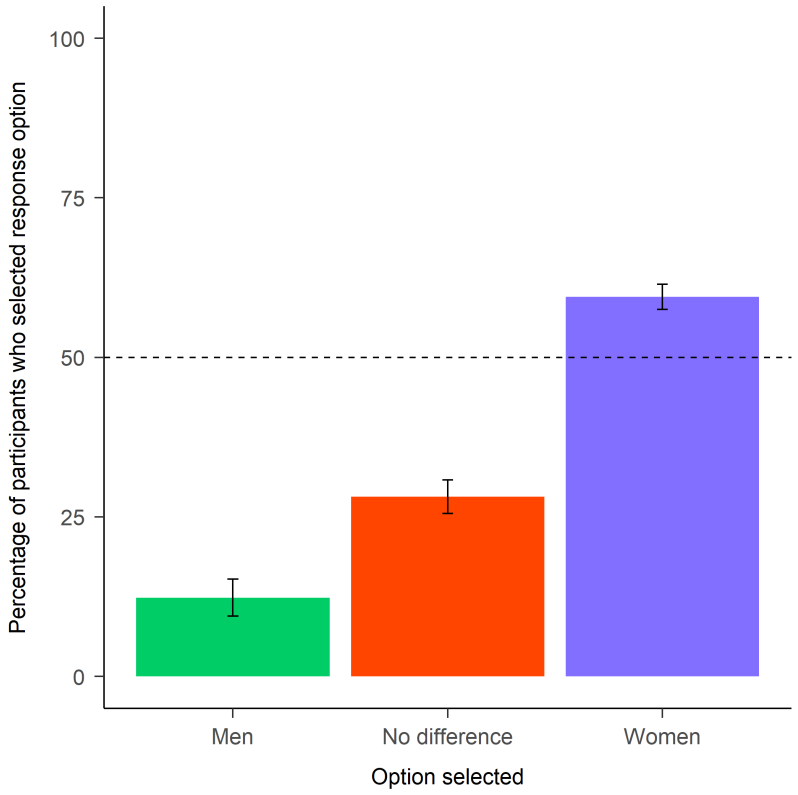

Figure 6: Proportion of participants that predicted women prepare more in general, men prepare more in general, or that there are no gender differences in preparation in general. A significantly larger proportion of participants expected women prepare more in general. Error bars represent standard errors.

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 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 @ref(fig:pract-choice-by-gender-and-perc-gen-prep-bar-study4)). We ran the 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 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 @ref(fig:pract-choice-by-gender-and-perc-task-prep-bar-study4)).

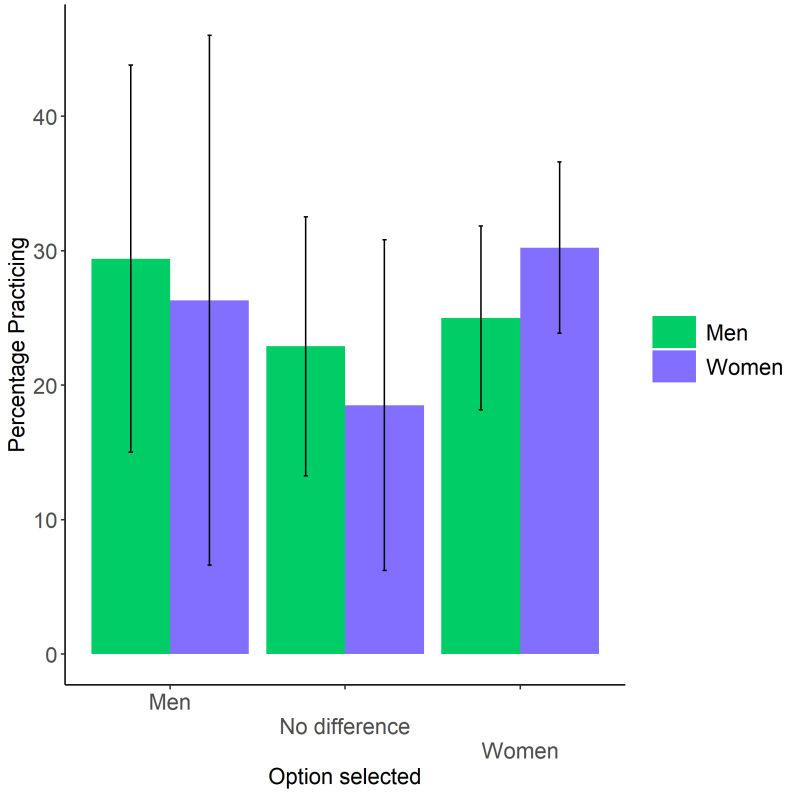

Figure 7: Proportion of men and women in Study 3 who chose to practice based on whether they thought men or women spend more time preparing on most tasks. In this study, participants also had the option to say there was no gender difference in preparation. It is worth noting that a small sample of participants (N = 136 participants chose to prepare across the entire study) are represented in this graph. Error bars represent standard errors.

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

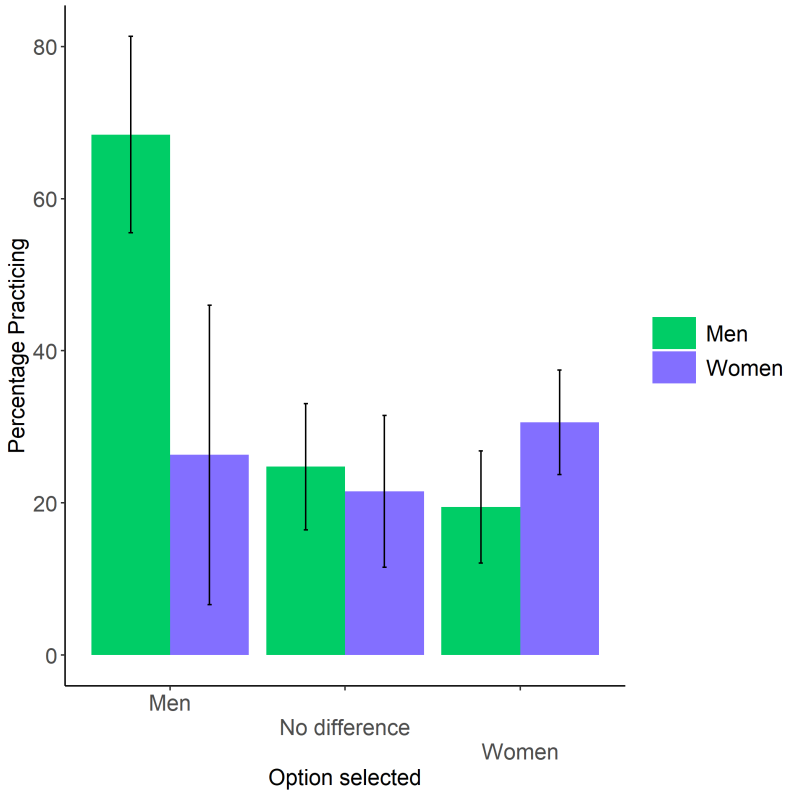

Figure 8: Proportion of men and women in Study 3 who chose to practice based on whether they thought men or women spend more time preparing for the multiplication task. In this study, participants also had the option to say there was no gender difference in preparation. It is worth noting that a small sample of participants (N = 136 participants chose to prepare across the entire study) are represented in this graph. Error bars represent standard errors.

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.

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 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, $\chi^2(1, n = 1072) = 548.50$, $p < .001$.