Study 1 methods

Keana Richards

6/9/2020

## Participants

Participants were recruited on Amazon Mechanical Turk using the same screening criteria as all previous studies in Chapter 1. Like the last study of Chapter 1, 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 using the same exclusion criteria from before. These exclusions were applied for all main analyses reported in the results section.

The final dataset consists of 3980 participants (57.36% women), with an average age of 41.3 (*SD* = 13.2) years. Of the final sample, 75 participants (30.67% women) dropped out of the study before finishing and 192 participants were flagged by Qualtrics’ fraud detection software as suspicious based on the aforementioned criteria. We include analyses for the full sample in the appendix and all results are unchanged (INSERT DOUBLE CHECK).

## Procedures

As in Studies 1 and 2, participants included in the study were told they would be completing a multiplication task. Notably, we aimed to recruit a larger sample to grant us enough power to detect our anticipated interaction effects, and decided to shorten the task from two minutes to 30 seconds. Otherwise, the task used was identical to the ones used in previous studies.

Like the studies in Chapter 1, participants were first told about the rules for the multiplication task and were required to pass the same comprehension questions used in the previous studies before moving onto the main manipulation of payment scheme.

### Manipulation of payment scheme

Unlike previous studies participants were not able to choose a payment scheme. Instead, they were told about their random assignment to one of two payment schemes: the non-competitive piece-rate payment scheme or a competitive tournament payment scheme. Men and women were evenly assigned to both conditions. If they were assigned to the piece-rate payment scheme, they were paid $.10 per problem solved correctly. If they are assigned to the tournament payment scheme, they were be randomly matched with another participant that was also assigned to this payment scheme and received $.20 per problem if they solved more problems than the other participant. Otherwise, they received nothing.

Again, we checked that condition was assigned evenly across participants (control= 50.21%) and genders included in the study. Of the men who completed the study, 50.38% were assigned to the control condition and of the women who completed the study, 49.72% were assigned to the control condition. We also assessed condition-dependent attrition by identifying the number of participants that dropped out during/after learning about condition and find that a relatively small proportion of participants out of the total sample dropped out after learning about their respective condition (*N* = 42) and thus, are not concerned that condition-dependent attrition is driving any of the effects found in this study.

### Main dependent variables of interest: Measures of preparation and perceptions of relative preparation

Crucially, participants will be given the opportunity to prepare after they are informed of their payment scheme. The nature of the unlimited preparation was identical to that used in Study 3 of Chapter 1, where participants who chose to prepare were shown 10 multiplication problems that were created randomly by drawing from the pool of numbers used in the main multiplication task. Unlike the last study in Chapter 1, participants were not asked to explicitly indicate whether they would like to study the times tables. Instead, they were shown the times table right after the practice problems and told they could check their answers using the table as desired. By including the option to check their answers, we hoped to make the practice itself more useful by providing participants a way to receive feedback on their responses. At the bottom of each practice page, participants were asked if they would like to continue practicing multiplication problems, with the option to continue as many times as desired or opt out at any point. The amount of time (in seconds) participants spent on each practice page was also recorded. Thus, like the previous studies, we have multiple measures of preparation by design: 1) the actual decision to practice problems (before knowing what the practice entails), 2) the number of practice problems participants attempted (quantified as number of practice problems not left blank, irrespective of accuracy), 3) among participants who chose to practice problems, the amount of time they spent across all practice rounds they completed, and 4) the number of extra practice rounds participants completed after having completed the first round of practice. Again, the number of extra practice rounds serves as a way to quantify the number of times participants continue to practice after having seen what the practicing/studying looks like and having gone through it at least once. By encoding participants who both chose not to practice and those who chose not to continue practicing after the first round with zeroes in the dataset when creating this variable, we are able to separate out the effect of the choice to practice from the choice to continue practicing.

After completing the practicing/studying, participants guessed how much their amount of practicing for the task compared to all other participants that completed the task by indicating the decile of their practice relative to other participants. We also asked participants to indicate their anticipated decile when their amount of practicing was compared to that of all participants who identified as men and women, respectively.

We used these the decile questions to create perceived practice deviation variables as follows: self-rated decile - actual percentile based on number of practice problems completed. Therefore, negative values for this variable indicate a participant expected to have practice less, relative to other participants, than they actually did, and vice versa for positive values. A value of zero, therefore, indicates that a given participant was completely accurate in their guess of relative practicing.

### Main task and post-task measures

After practicing, participants completed the paid multiplication task, received feedback about their absolute (but not relative) performance, and completed many of the same follow-up questions used across Chapter 1, including risk attitudes, confidence, and perceptions of gender differences in preparation, competitiveness, and performance. One of the perceptions of gender differences questions deviated slightly from the previous studies, which was edited for the sake of clarity. Instead of asking participants to indicate “Do you think men or women in this study chose the tournament payment option more often?”,they were asked “If given the opportunity to choose between the two payment schemes (Piece Rate or Tournament), do you think men in this study would choose the piece rate or the tournament payment scheme more often?”, with the options to indicate: “Men would choose tournament more often than piece rate”, “Men would choose piece rate more often than tournament”, or “Men would choose each payment scheme equally”. This question was repeated with respect to women in the study.

We paid participants to answer the questions about their confidence and perceptions of gender differences correctly at the same rate as previous studies. Finally, they completed the same demographic questions from Chapter 1 and provided feedback on the study before being paid for their participation.