Pilot study methods

Keana Richards

6/9/2020

546 participants were recruited on Amazon Mechanical Turk for a study on decision-making. Participants were eligible to participate if they passed all three comprehension check questions about the task they would be completing, were using a computer to complete the survey, identified their nationality as American, lived in the United States (to control for gender differences in competitiveness across cultures), did not respond “Other” for their gender, and finished the survey in its entirety. A total of 226 participants were excluded before analyzing the data based the above criteria: 9 were excluded because they did not indicate they were American or lived in the United States, 145 were excluded because they failed all 3 comprehension check questions, 2 were excluded for indicating “Other” for their gender, 39 were excluded for using a phone or tablet to complete the survey, and 34 were excluded for an incomplete survey (note: 2 participants were excluded based on more than one criteria). The final sample consisted of 320 participants (55.94% women), with an average age of 37.21 (*SD* = 11.56) years.

Participants were asked to complete three paid rounds of a one-minute “key-entry task,” where they matched numbers to their corresponding letters as quickly as possible. For instance, when presented with the number 12, they would have to enter “CR” into the corresponding text box, where the letter “C” was associated with 1 and “R” was associated with 2. To complete the task, participants were continuously presented a legend, which contained 5 numbers in one row and a randomly assigned alphabetical character in the subsequent row. Before participants entered the paid rounds, they were shown an example problem with the correct answer and then were required to pass 3 comprehension check questions, which were identical in structure to the questions asked during the paid rounds.

In rounds 1 and 2 of the task, participants were required to follow a piece-rate and tournament payment scheme, respectively [@Niederle2007]. Under a piece-rate payment scheme, participants earned a fixed reward ($.05) for each correct answer. In contrast, the tournament payment scheme offered a larger reward for each correct answer ($.10), but participants only earned a reward if their score was higher than a randomly assigned partner.

In round 3, participants had the option to choose between the two payment schemes, which served as the primary dependent variable. After completing the third round, participants completed a series of follow-up questions to be used for testing our secondary hypotheses and conducting exploratory analyses, including measures of confidence, risk aversion, and perceptions of which gender performed better on the task, all of which were incentivized. One of their responses to these measures was randomly selected and if the selected guess was correct, they received a bonus of $.10.

For the first confidence measure, participants indicated if they thought their score from round 2 was higher or lower than the person they competed against. Then, participants indicated their guess for which decile their score might fall into relative to all men who completed the task during round 2 in 10 percent increments (e.g., 0-10, 11-20). For instance, if a participant selected the option that represented the 51-60th decile, they thought their score was higher than 51 to 60 percent of male participants who completed round 2. They answered the same question relative to all female participants. For the measure of risk aversion, participants were asked “How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?” [@Dohmen2011b] on a 0 (Not at all willing to take risks) to 10 (very willing to take risks) scale.

Finally, we asked about participants’ perceptions of the effects of practice, if offered, on task performance (i.e., “Do you think your score would have improved if you practiced the task beforehand?”), along with participants’ hypothetical willingness to practice the task beforehand, if given the opportunity. We operationalized this measure as their response to the question “If you had the chance to practice the task, would you have taken that opportunity?” If they responded yes, they were asked how long they would have practiced (in minutes) if they were given unlimited time to practice.