Study Information

Hypotheses

H*: One-stage tournaments are behaviorally different from multistage setups

Statistical test: Mann-Whitney U-test

Expected pattern of results: Subjects exert higher effort in two-stage tournament treatments than in the one-stage tournament

Design Plan

Study type

Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

Blinding

For studies that involve human subjects, they will not know the treatment group to which they have been assigned.

Is there any additional blinding in this study?

N/A

Study design

The study will be a between-subjects, two-treatment design. In the one-stage (OS) treatment, four participants take part in a single-stage contest; in the two-stage (TS) treatment, four participants compete in a 2-stage elimination contest, where the two winners of the first stage move to a second stage (the losers do not compete in the second stage).

The main unit of analysis is the effort chosen by a participant in the contest. In the OS treatment, it is their single effort choice. In the TS treatment, it is the effort choice in the first stage.

The experimental design is between-subjects, which means participants will only be exposed to one treatment.

The original paper included the instructions (translated from German) in its online appendix, and the authors have provided the software used in the original study, as well as additional experimental materials (a cost table and control questions).

No files selected

Randomization

We will employ simple randomization to treatment. We will advertise sessions for the study using our recruitment software hroot, which selects a random sample of volunteers and sends invitation email with dates and times of sessions. We will not condition the randomization to any observable characteristic, as this was not done in the original study.

Sampling Plan

Existing Data

Registration prior to creation of data

Explanation of existing data

This project is being pre-registered prior to any new data collection.

Data collection procedures

Participants will be registered volunteers with the FEELE lab at the University of Exeter. We will advertise sessions for the study using our recruitment software hroot, which selects a random sample of volunteers and sends invitation email with dates and times of sessions. We will not use any exclusion criteria for recruitment for two reasons: first, the original study did not report any exclusion criteria ("All participants were students from various fields of study", p.158); second, participants in FEELE have not taken part in a study of this kind, so there are no concerns regarding previous experience.

Recruitment emails will be sent two to three days prior to the study taking place to maximize show-up rates.

Should the number of participants who turn up to a session not be divisible by 4, I will randomly select the largest feasible multiple of 4 participants to take part and the remainder will receive £5 for turning up on time, as per the FEELE lab rules. Also, the minimum session size will be 8. Should fewer than 8 participants turn up, I will not run the session, as anonymity of decisions is compromised, and compensate those who turn up with £5.

Participants will be paid £5 for taking part, plus an additional payment which will be a function of their payoff in the study itself. I estimate the additional payment will be £11.75, based on the average payment reported in the paper (€11.75). The £5 is standard in the FEELE lab (and is similar in magnitude to the €4 reported in the study).

Participants will be invited to the FEELE lab and sit at individual booths. They will read the instructions, and respond to control questions. They will then do the experiment itself. They will enter their decisions on a computer screen (z-tree). I will not collect the additional questionnaires reported in the paper on risk attitudes and competitiveness, as this was not provided by the authors and it is not part of the pre-registered hypothesis. Since the

questionnaire content was not pre-announced in the original instructions (and will not be so in the replication study), this will not have any impact on decisions. I will only collect standard demographic information (age, sex and field of study).

The "Material and Methods" component of the OSF page for this project includes screenshots of the experimental interface, the instructions, comprehension check questions and the cost table.

No files selected

Sample size

The target analytic sample size is 312 participants (156 per treatment).

Sample size rationale

Power calculations were conducted in accordance with the guidelines of the Social Sciences Replication Project (SSRP). However, instead of carrying out two rounds of data collection to achieve the stage 1 sample size target followed by the stage 2 sample size target as in Phase 1 of SCORE, in Phase 2 only one data collection effort will be carried out to achieve a single target analytic sample size.

All replications will aim to achieve the target analytic sample size, which is determined by a power analysis achieving 90% power to detect 50% of the original effect size (formally referred to as the Stage 2 sample size). However, if this sample size is not attainable for the replicating lab, the sample size will instead be calculated by achieving 90% power to detect 75% of the original effect size (formally referred to as the Stage 1 sample size).

For this study, the stage 1 analytic sample size (i.e., 312 participants) will be collected.

Documentation of the power analysis conducted for this replication is available at the following view-only link: https://osf.io/9erqg/?view_only=5e8ac0ecd6b34e45be595de097be83aa.

Stopping rule

The planned sample size is 312 participants. After achieving this sample, sampling will stop and planned analyses will be run.

The original study did not report a stopping rule. Given that a fraction (15-20%) of participants at the FEELE lab who sign up to a session does not attend, and session size may vary depending on sign-ups (session size may vary between 8 and 20) it is possible that I will collect a larger sample than the pre-registered value. Therefore the stopping rule I will use will be to stop when the total sample is greater than or equal to the pre-registered value in each treatment. I will not discard any observations collected in excess of the original target sample, since power analysis states a minimum sample to observe a desired effect size.

Variables

Manipulated variables

The experimental manipulation will be the number of stages in the tournament. In the OS treatment, subjects will compete in a one-stage tournament. In the TS treatment, they will compete in a two-stage tournament. I will generate a dichotomous variable for the OS treatment (1 if OS, 0 otherwise).

No files selected

Measured variables

The main measured variables are participant effort (an integer between 0 and 125 in the OS treatment, and two such values in the TS treatment for two participants in each group); beliefs about effort levels of other participants (integers between 0 and 125) and socio-economic variables: age (integer >18), gender (1 if male, 0 if female, 2 if non-binary, 3 if decline to answer) and field of study (categorical).

No files selected

Indices

N/A

No files selected

Analysis Plan

Statistical models

For the purposes of SCORE, to test H* and attempt to replicate the H* statistical evidence (if relevant), we will take individual effort in stage 1 and conduct a Mann-Whitney test of equality of distribution of effort in stage 1 between OS and TS treatments.

No files selected

Transformations

N/A

Inference criteria

The criterion for a successful replication attempt for the SCORE project is a statistically significant effect (alpha = .05, two tailed) in the same pattern as the original study on the focal statistical evidence (H*). That is, a larger mean effort in stage 1 of the TS treatment than the mean effort in the OS treatment.

Data exclusion

We will not exclude outliers, as the original paper did not propose any rule for exclusion.

Missing data

Other than a participant abandoning the study (in which case the whole group would be lost), there are no reasons to expect missing data. Decisions in the experiment are forced responses.

Exploratory analysis

N/A

Other

Other

There are three main deviations from the original study. First, the experiment will be conducted in English, rather than German. Secondly, the subject pool will be made up of University of Exeter students, as opposed to University of Bonn. Finally, the payment to participants will be made electronically through PayPal or bank transfer. None of these deviations should impact results.

I will make public all the experimental materials, including software, instructions and data through the project's OSF page.