

# WCST data analysis - Step 1

## Overview

We used this script for the following purposes:

- to gather data from all the raw data files
- to merge data into a few data frames
- to calculate variables from the data
- to output data into excel spreadsheets for manual rating of the free text answers
- to exclude participants
- to save data frames

## Read data sources

To run this script, copy-paste the following files to “datadir”, and change “datadir” to your path to that folder:

- prolific\_export\_5f917030b1ac5a05a2123cac.csv: this contains the demographic data of participants from Prolific
- all the result files from the server (one for each participant): these contain the data from the game (file names: condition\_ID\_Results.txt)
- vpn.txt from the server: this contains data from the pre-game questionnaire
- survey.txt from the server: this contains data from the post-game questionnaire
- time.txt from the server: this contains the timestamps of different actions from the participants during the experiment (e.g., started game, filled out survey, etc.). This is not used for data analysis, but can be useful to check manually if something is not clear

We have manually made some modifications to these files to ensure that the data is read correctly. In the vpn.txt file:

- Participant 306: accidentally copied an URL next to their ID, we deleted the URL
- Participant 509: wrote “4o” for age, we corrected it to 40
- Participant 894: copy-pasted a random text to the ID field, we deleted it

In the survey.txt file:

- We deleted empty lines, where participants hit enter in a free text field for the following participants: 188, 237, 298, 363, 434, 529, 563, 582, 646, 763, 790
- Participant 593 wrote “10!” for difficulty, we corrected it to 10
- Participant 790, when describing the rule, the html number > appeared instead of >; we changed it back, so the rule reads “1>2>3>4.”
- Participant 886 wrote “d” for the level of difficulty field - we deleted this entry

We collected survey data separately for a few participants for whom the server froze before they could submit their survey. We made a google form with the survey questions and we asked participants to fill it out after they indicated their problem via the Prolific messaging system. This usually happened within a week of them completing the experiment. One participant indicated that they do not remember the experiment well - we excluded their data. Some participants proactively sent us their survey answers via messaging within the Prolific platform before we could send them the link for the google form. For these participants, we copy-pasted their answers from their message to the Google form.

## Merge data

We merged the data from the two survey files and the prolific file into a data frame, called “participantdata”.

## Calculate variables

### Variables per move

For each move participants made during the game we had the following variables from the server:

- Trial number
- Source card
- Target card
- Match: correct or incorrect move
- Total time: the time elapsed between the source card appearing and the source card being dropped
- Move time: the time elapsed between dragging and dropping the source card
- Time stamp

Apart from these, we calculated a few more variables for each move:

- Think time: total time - move time
- Matching shape: whether the move conformed to the shape rule
- Matching color: whether the move conformed to the color rule
- Matching number: whether the move conformed to the number rule

### Variables per participant

We also calculated some descriptive variables for each participant based on their move data:

- Number of moves
- Task time (min): time spent with the task (only the card sorting game)
- Mean total time (ms): mean trial time
- Mean move time (ms): mean drag and drop time
- Mean think time (ms): mean thinking time

We merged these variables with the main data frame (participantdata).

## Exclusion criteria

We excluded entries where the participant:

- Played the game more than once, even if they got assigned a different condition after the first time (we still included their first attempt though): -195
- Did not fill out the first data sheet: -1
- Did not start the game after filling out the data sheet: -76
- Refreshed the screen during the game or went back to the instructions page after starting the game: -31
- Who did not fill out the final survey: -98
- Indicated that they were colorblind: -4
- Revoked their consent to use their data: -4

Those participants who did not fill out the survey could be those, who

- quit the experiment during the game before time out, or
- left their computer and never pressed the “Continue” button after time out, or
- who finished their game, but who could not submit the survey, because the server froze.

Our participant funnel looked like this:

- The number of experimental sessions started: 958
- The number of participants who properly finished the experiment and surveys without refreshing the screen or going back in their browser: 557
- From these, the number of participants who were not colorblind: 553
- From these, the number of participants who did not revoke their consent to use their data: 549

All-in-all, we excluded 409 initialized experimental sessions and we were left with 549 participants.

## Stopping criteria

### Stopping rule

We run the experiment until we had enough (*78 participants in each condition*) after automatically evaluating them based on the exclusion criteria described above. First, the conditions were assigned randomly, then as we got closer to the desired number of participants, we only let a certain number of participants to enter each condition to make up for the missing number of participants. This process was not exact, so we ended up with a few extra participants in some conditions. We stopped the experiment when each condition had at least 78 participants after the exclusions.

Condition	Number of participants	Aha rate of solvers	Failure rate	Avg task time	Avg number of moves
wlin	79	0.823	0.000	1.595	31.835
wlout	78	0.932	0.064	4.286	68.756
wnol	78	0.980	0.359	8.985	166.500
wnolfs	78	0.884	0.115	6.606	124.962
wnola	78	0.900	0.487	10.360	236.103
moonsq	79	0.761	0.101	4.659	95.620
wonly	79	0.582	0.000	0.840	19.278

### To do on Prolific

Approve the participation and pay the following participants (to be copied to Prolific -> Approve by upload):

Batch approve list

Manually reject participation on Prolific:

- If the same ID shows “quit” and then “played twice”, approve anyway. These are probably cases where the game froze and then the participant restarted the experiment.
- Check Messages too, because sometimes participants explain their technical difficulties there.
- Check comments in the survey too, because sometimes it is obvious from their comments that they finished the experiment, even though they have missing data.

prolific\_ID   prolific\_action

### Costs

We estimated that the median time for completing the experiment would be 8 minutes. The suggested hourly rate is 7.5 GBP on the Prolific website. From this, we calculated the fee for the task to be 1 GBP (independently of how long it took for the participant to finish). Together with 33% service fee and 20% VAT of the service fee, each approved participant costs around 1.396 GBP.

The time-out limit was calculated by the Prolific website to be 39 minutes for the whole experiment including the surveys (for the game, our own time-out was 15 minutes).

826 participants initiated the experiment on the Prolific website. From this:

- 111 participants returned the task
- 17 participants timed-out
- We excluded 31 participants for various reasons (see above)
- We approved and paid the remaining 667 participants
- From these, we could use the data of 549 participants

Sometimes we lost data because of server error. Other times, the screen of participants froze (probably also server error). In these cases, we cannot use the data of these participants, but we still have to pay them, because it was not their fault.

Costs for the experiment should be: 931.132

The pilot experiments cost 73.23 GBP all together:

- Pilot 1: 33.60 GBP
- Pilot 2: 16.80 GBP
- Pilot 3: 14.01 GBP
- Pilot 4: 8.82 GBP

## Save data

### Data files

- We saved all dataframes that should be used in Step 2 as "`__Step1__output.RData`".
- We saved free text answers for manual rating as "`__free_text_answers.xlsx`".

### Participant plots

We can save plots for each participant, representing their strategies.