WCST data analysis - Step 2

# Manipulate data

## Read data frames and manually rated data

We have loaded data frames saved by the Step 1 script, and merged the manually categorized free text answers about the rule with participantdata\_inc.

## Exclude participants

We excluded extra participants if there were more than 78 in a condition.

## Modify data

Diffifulty of tasks was supposed to be a number 1-10, but it was also a free text answer, and some participants went outside the recommended range (e.g.: difficulty = 50). We capped these numbers to 10.

# Descriptive statistics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Condition | Number of participants | Aha rate of solvers | Failure rate | Avg task time | Avg number of moves |
| LettersIn | 78 | 14 | 0.000 | 0.821 | 1.605 |
| LettersOut | 78 | 5 | 0.064 | 0.932 | 4.286 |
| NoLetters | 78 | 1 | 0.359 | 0.980 | 8.985 |
| FixedSequence | 78 | 8 | 0.115 | 0.884 | 6.606 |
| AmbiguousCards | 78 | 4 | 0.487 | 0.900 | 10.360 |
| Moon | 78 | 17 | 0.103 | 0.757 | 4.672 |
| LettersOnly | 78 | 33 | 0.000 | 0.577 | 0.846 |

# Statistical tests

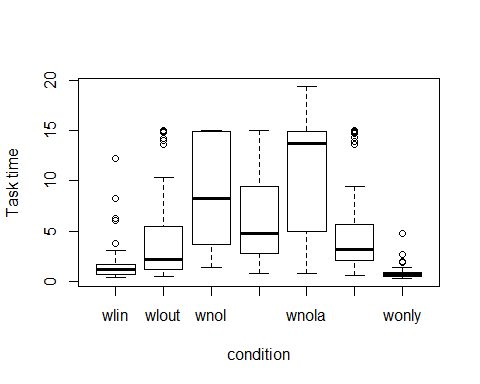
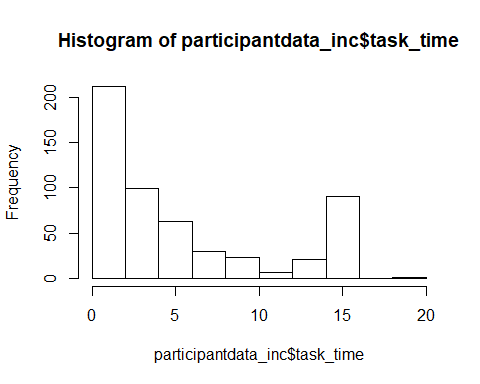
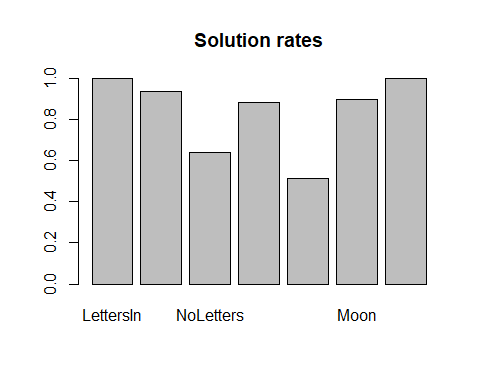
## Difficulty of the task

A task is more difficult in one condition than in another condition, if less participants are able to solve it, or if it takes longer for participants to solve it (even if solution rates are the same). This is why, we compared the difficulty of the task

* first, by Fisher’s exact test on the number of solvers (pairwise comparisons)
* then, if the Fisher’s exact test was not significant, we compared task time by two-sample Wilcoxon test (same as the Mann-Whitney test) or by
* ANOVA with Greenhouse-Geisser correction: robust also for non-normally distributed data

The histogram of task times shows that the data is not normally distributed, so we had to use nonparametric tests.

We also used binary logistic regression.



## Aha feelings

We used Fisher’s exact tests to make pairwise comparison between conditions regarding the aha-ratings of solvers (whether they reported having aha-feelings or not).

# Results - Experiment 1

## Difficulty of the task - Experiment 1

We analyzed the contingency table containing the number of solvers and non-solvers in pairs of conditions with Fisher’s exact test. A p<0.05 means that the row/column association is statistically significant:

WLIN-WLOUT: p = 0.0585073  
WLOUT-WNOL: p = 8.205745210^{-6}

Wilcoxon test:  
WLIN-WLOUT: p = 5.054639710^{-7}

## Aha feelings - Experiment 1

We analyzed the contingency table containing the number of solvers who reported Aha-feelings and number of solvers who did not report Aha-feelings in pairs of conditions with Fisher’s exact tests. A p<0.05 means that the row/column association is statistically significant:

WLIN > WONLY: p = 0.001533  
WLOUT > WONLY: p = 4.403268610^{-7}  
WNOL > WONLY: p = 7.236559310^{-8}

WLOUT > WLIN: p = 0.0500146  
WNOL > WLOUT: p = 0.3993585

# Results - Experiment 2

## Difficulty of the task - Experiment 2

MOONSQ < WNOLA: Removing distracting visual cues makes finding the sequence rule easier WNOLFS < WNOLA: The sequence rule can be used in both conditions but the exclusion rule can only be used in the WNOLFS condition, which might help.

We analyzed the contingency table containing the number of solvers and non-solvers in pairs of conditions with Fisher’s exact test. A p<0.05 means that the row/column association is statistically significant:

WNOLFS-WNOLA: p = 5.552353710^{-7}  
MOONSQ-WNOLA: p = 1.641139810^{-7}

## Aha feelings - Experiment 2

WNOLFS > WONLY: p = 3.757604210^{-5} WNOLA > WONLY: p = 2.990651610^{-4} MOONSQ > WONLY: p = 0.0241923

# Other ideas

## Language

## Restructuring

Sliding average for move time

## Rules used (matching rule for moves)

## Correct rule

Does it correlate with problem difficulty or aha

## Aha-rate of non solvers

Correlates with difficulty? Number of false ahas vs subjective difficulty or solution time

## Goal of the experiment