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 taks 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 | 0.821 | 0.000 | 1.605 | 31.962 |
| LettersOut | 78 | 0.932 | 0.064 | 4.286 | 68.756 |
| NoLetters | 78 | 0.980 | 0.359 | 8.985 | 166.500 |
| FixedSequence | 78 | 0.884 | 0.115 | 6.606 | 124.962 |
| AmbiguousCards | 78 | 0.900 | 0.487 | 10.360 | 236.103 |
| Moon | 78 | 0.757 | 0.103 | 4.672 | 94.923 |
| LettersOnly | 78 | 0.577 | 0.000 | 0.846 | 19.295 |

# Statistical data analysis: Experiment 1

## Difficulty of the task

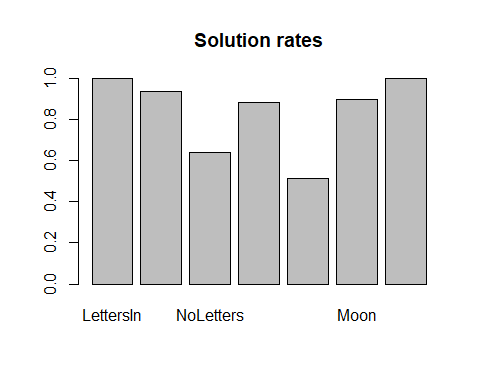
We compared the difficulty of the task in different conditions

* first by a Fisher’s exact test on the number of solvers
* if the Fisher’s exact test test is not significant, we will compare task time, because it is possible that the number of solvers does not differ in the two conditions but still, if task time is significantly lower in one of the conditions that means that those who solved it, solved it faster. For comparing task time, we will use ANOVA if the data is normally distributed, or a two-sample Wilcoxon test (same as the Mann-Whitney test) if it is not.

### Solution rate

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

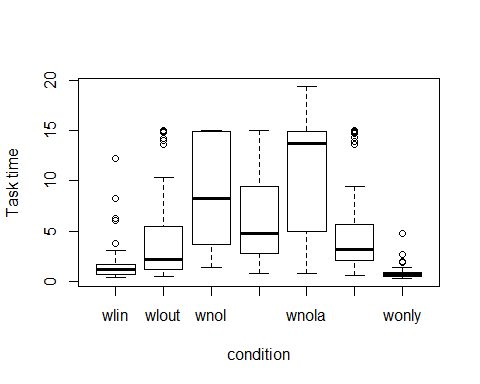
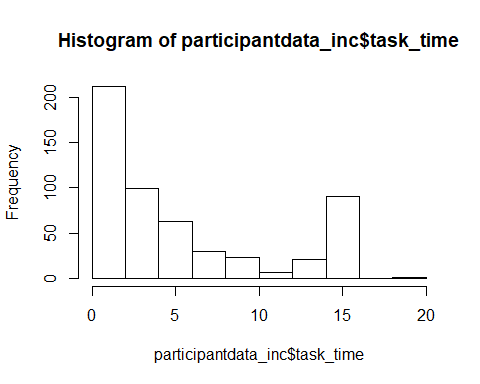
Experiment 1: WLIN < WLOUT < WNOL



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

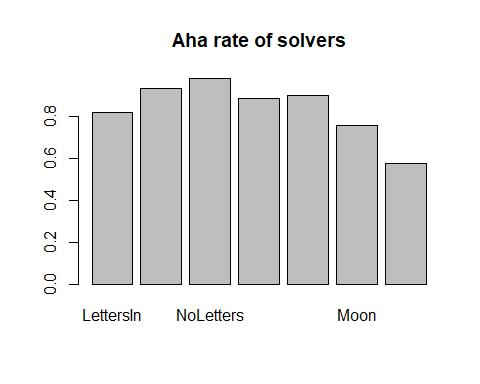
### Solution time

We checked whether the data was normally distributed with Kolmogorov-Smirnoff test:

 If the data is normally distributed, we use ANOVA, if it is not, we use Wilcoxon.

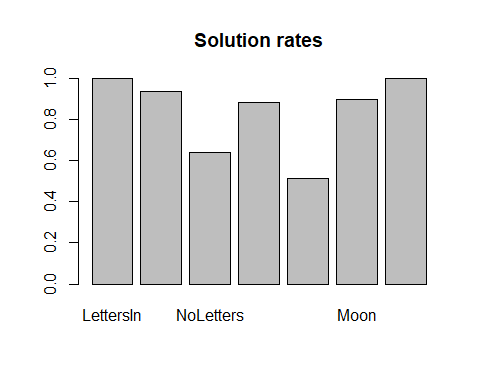
### Binary logistic regression

## Aha feelings: Fisher’s exact test



# Statistical data analysis: 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.



##   
## Fisher's Exact Test for Count Data  
##   
## data: partic\_groups2[c("wnolfs", "wnola"), c("Nbof\_solvers", "Nbof\_nonsolvers")]  
## p-value = 5.552e-07  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
## 3.029108 18.713246  
## sample estimates:  
## odds ratio   
## 7.185509

##   
## Fisher's Exact Test for Count Data  
##   
## data: partic\_groups2[c("moonsq", "wnola"), c("Nbof\_solvers", "Nbof\_nonsolvers")]  
## p-value = 1.641e-07  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
## 3.352528 22.396910  
## sample estimates:  
## odds ratio   
## 8.192766

## Other

### Language

### Restructuring

Sliding average for move time

### Rules used