## portfolio 1. part 1. Visual Search experiment

[https://github.com/kurmie/Portfolio-1-Eye-tracking/blob/master/DataAnalysis.Rmd]

*Introduction*

The focus of this experiment was to replicate a study aimed to evaluate the effect of underlying top-down processing on visual foraging. The use of top-down processing depended on task structure and goals. Stimuli were divided into two tasks: target counting task and hidden target searching task.

Our hypotheses exist on two different planes:

Conceptual: Visual search patterns are affected by task structure and goals (top-down)

Operational 1: Duration of fixations differs depending on a type of a task.

Operational 2: Amplitude of saccades differs depending on a type of a task.

Therefore, the observed phenomena were used to evaluate if visual search patterns are affected by the type of processing underlying it.

1. *Methods*

Data collection

The experiment took place in a lab room placed in a basement, so the lighting is assumed to be non-dependent on daylight and consistent across the participants. Every participant was seated approximately 30 cm in front of a 20-inch flat panel monitor. An Eye Link 1000 head mounted eye tracker was used to record monocular eye positions and pupil size data at 1000 Hz sampling frequency. Prior to exposing the participants to the experimental design, the eye-tracker was calibrated for every participant using the in-built nine-point automated calibration procedure, that was repeated if necessary. The validation procedure needed to report average errors below 1 and maximal error below 1.5 to successfully end the calibration procedure. Further the experimental paradigm was run as a PsychoPy implementation and time stamps for the initiation of stimulus exposure were recorded.

Participants

Participants of the experiment were looked at to evaluate if there are potential issues that might affect the quality of data, such as eyeglasses and lenses or presence of mascara. However, neither of factors performed as exclusion criteria. In total, there were six participants of the experiment, four of which were female. One of the participants wore glasses. All of the participants were Cognitive Science students in age range between 20 to 30 years old. The validation procedure reported that calibration process was finished with average errors below 1 for every participant (on average 0.87).

Data pre-processing

The recorded eye-tracking data (velocities, pupil size, and x and y coordinates) were automatically pre-processed using the in-built DataViewer software. Artifacts were removed. Eye-blinks, saccades and fixations were identified. The validation data were used to model measurement error and assess if any participants should be excluded. The data was high-pass filtered at a 100 s cut off to counter calibration drift. Systematic bias in fixation estimates on fixation crosses was estimated and positions were accordingly adjusted at every trial.

Data analysis

To test the operational hypothesis, the following measures were included in the data analysis:

|  |  |  |
| --- | --- | --- |
| Measure | Role | Reasoning |
| Duration of a fixation | Outcome/dependent variable | It was selected as one of the main characteristics of the visual search pattern |
| Amplitude of a saccade | Outcome/dependent variable №2 | It was selected as the second main characteristic of the visual search pattern |
| Type of the task | Main predictor variable | Manipulation of this variable is the most straightforward manipulation of the outcome according to the hypothesis:  It displays the consequences of top-down inference on visual scene processing. |
| Number of the trial | Predictor variable | It was used to control for the effect of familiarity and/or fatigue based on amount of time participants spent in the experimental setting |
| Number of the fixation | Predictor variable | It was used to test if variation in the outcome variable could be explained by the progress of visual processing: the duration of early fixations could be different from later fixations on the same stimulus |
| Participant ID | Random effect | It was used to control for random effects of individual variation across participants |

Different combinations of predictor measures were considered in the number of models. The models for the first outcome variable, duration of fixations, are listed in the table below:

|  |  |
| --- | --- |
| № of a model | Formula |
| 1 | Fixation Duration = β0i + β1i Type of the task + (1+Type of the task | Participant ID) +ε |
| 2 | Fixation Duration = β0i + β1i Type of the task + β2i Number of the trial+ (1+Type of the task+ Number of the trial | Participant ID) +ε |
| 3 | Fixation Duration = β0i + β1i Type of the task + β2i Number of the fixation + (1+Type of the task+ Number of the fixation | Participant ID) +ε |
| 4 | Fixation Duration = β0i + β1i Type of the task + β2i Number of the trial + β3i Number of the fixation + (1+Type of the task+ Number of the trial+ Number of the fixation | Participant ID) +ε |
| 5 | Fixation Duration = β0i + β1i Type of the task \* Number of the trial+ (1+Type of the task \* Number of the trial | Participant ID) +ε |
| 6 | Fixation Duration = β0i + β1i Type of the task \* Number of the fixation + (1+Type of the task \* Number of the fixation | Participant ID) +ε |

The models for the second outcome variable, amplitude of saccades, are listed in the table below:

|  |  |
| --- | --- |
| № of a model | Formula |
| 7 | Saccade Amplitude = β0i + β1i Type of the task + (1+Type of the task | Participant ID) +ε |
| 8 | Saccade Amplitude = β0i + β1i Type of the task + β2i Number of the trial+ (1+Type of the task+ Number of the trial | Participant ID) +ε |
| 9 | Saccade Amplitude = β0i + β1i Type of the task \* Number of the trial+ (1+Type of the task\* Number of the trial | Participant ID) +ε |

The mixed effects models were selected as the most appropriate type of a model due to the presence of repeated measures per participant in the data and in order to control for random effects. The structure of the random effects was consistent in all models, allowing random intercepts for participants and random slopes for the predictors included in the model. Two interaction models (5 and 6) were created for the first outcome variable. In these models it was assumed that the effects of predictors influence each other: the effect of the number of a trial has an influence on the effect of the type of the task (model 5), and the effect of the number of a fixation had an influence on the effect of the type of the task (model 6). One interaction model (9) of the same structure as model 5 was created for the saccade amplitude outcome variable.

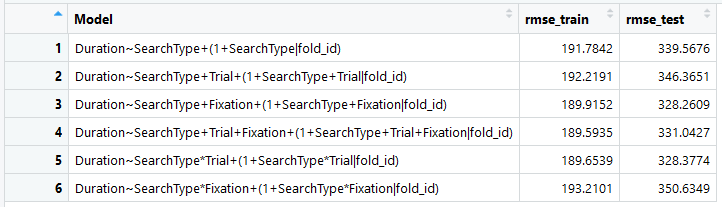
A 3-fold cross-validation model selection procedure was employed to identify which combination of parameters produces the best model. Cross-validation was stratified at the participant level and balanced across conditions. Out-of-sample error operationalized as root mean square error was used as the selection criterion. The model with the least root mean square error was selected as the best model and was ran on the full dataset to optimize parameter estimation.

The data analysis was performed using the statistical software of R studio (RStudio Team, 2015) on R (R Core Team, 2017). The data from the eye-tracker was cleaned up using packages readr version 1.1.1 (Wickham, Hester &amp;amp; Francois; 2017), and data.table version 1.10.4-3 (Dowle, Srinivasan &amp;amp; 2017). The generalized linear mixed effect models were designed by using packages lme4 (Bates, Maechler, Bolker &amp;amp; Walker, 2015), lmerTest (Kuznetsova, Brockhoff &amp;amp; Christensen, 2016). The cross-validation procedure was performed by using the package caret version 6.0-79 (Kuhn, 2018), and root mean square error was calculated by using the package Metrics version 0.1.3 (Hamner, Frasco, LeDell; 2017).

1. *Results*

Fixation duration models

The 3-fold cross-validation selection process yielded the following model performance results:

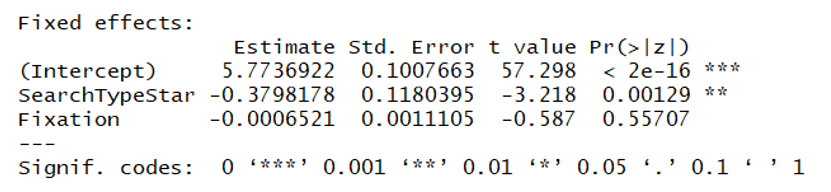


As can be seen in the results, the model 3 showed the lowest out-of-sample error. The parameters of the best model included the type of the task and the number of a fixation, with an average root mean square error of 328.26 when performed on the test data.

The formula of the selected model is presented below.

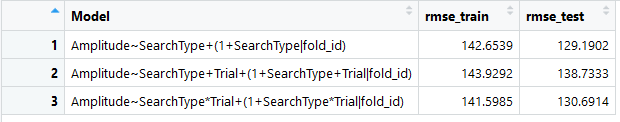
*Fixation Duration = β0i + β1i \* Type of the task + β2i \* Number of the fixation + (1+Type of the task+ Number of the fixation | Participant ID) +ε*

Performing the task of finding a hidden target had a significant negative effect on the duration of fixations on a stimulus (β = -0.3798, se = 0.1180, t-value = -3.218, p<0.05). In other words, when the participant was engaged in the search task, the fixations were 380ms shorter than when engaged in the counting task. The number of the fixation had a non-significant negative effect on its duration (β = -0.0006, se = 0.0011, t-value = -0.587, p>0.05), which means that the progress of the stimulus processing and familiarity with the stimulus did not affect how fast participant was fixating on it.



Saccade amplitude models

The 3-fold cross-validation selection process yielded the following model performance results for the saccade amplitude models:

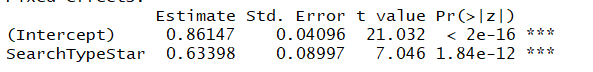


As can be seen in the results, model 7 showed the lowest out-of-sample error. The parameters of the best model included the only predictor of the type of the task, with an average root mean square error of 129.19 when performed on the test data.

The formula of the best selected model is presented below.

*Amplitude of a saccade = β0i + β1i \* Type of the task + (1+Type of the task | Participant ID) +ε*

Performing the task of finding a hidden target had a significant positive effect on the amplitude of saccades performed while processing the stimulus (β = 0.63398, se = 0.08997, t-value = 7.046, p<0.05). In other words, when the participant was engaged in the search task, the saccades were 0.63mm longer than when engaged in the counting task.



Visual representation of data

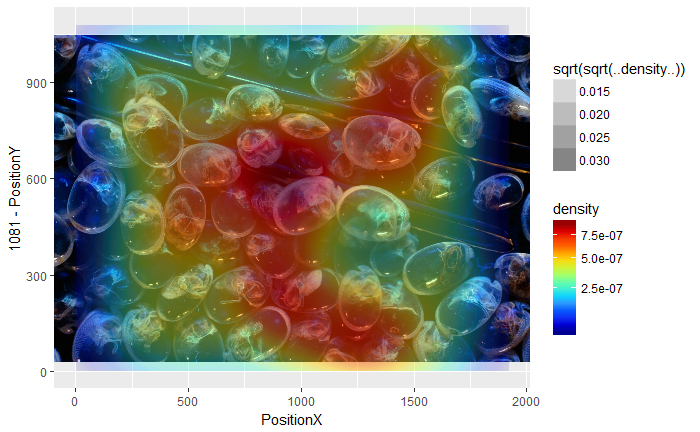


Figure 1. The heatmap of density of fixations for the target search task

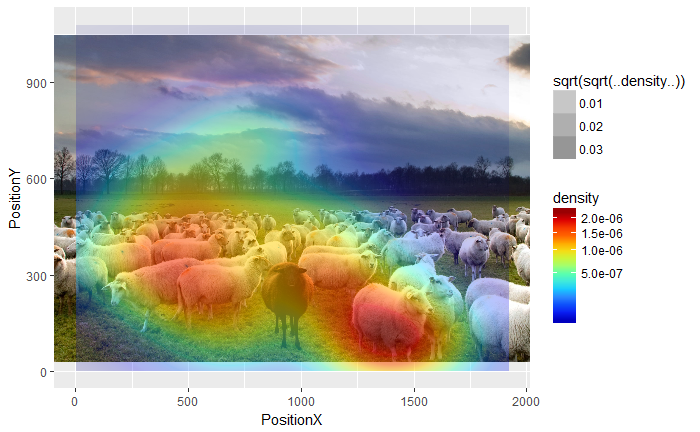


Figure 2. The heatmap of density of fixations for counting task

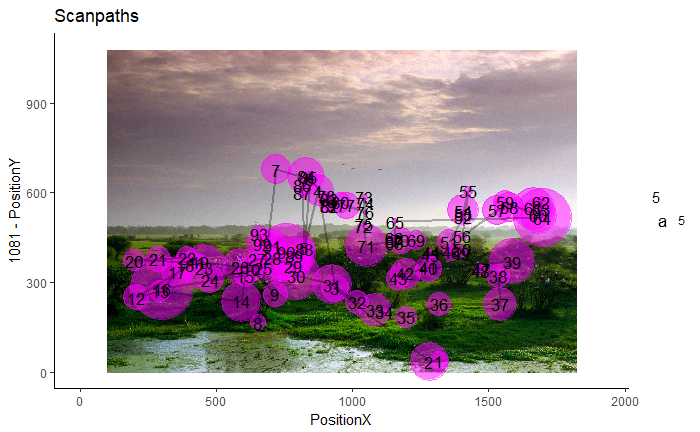


Figure . The scanpath visualization

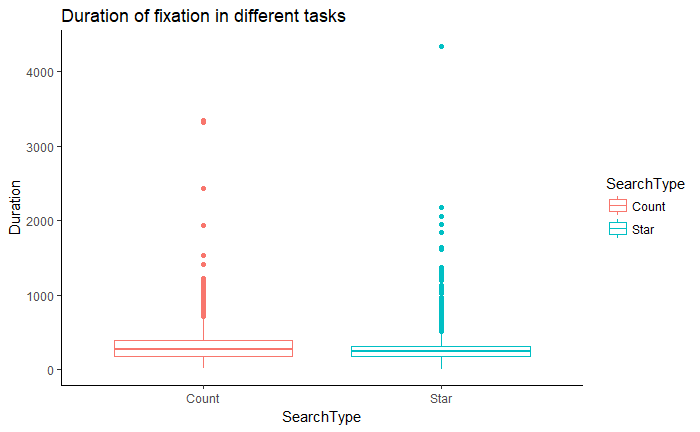


Figure 4. The boxplot for comparison of duration of fixation between task conditions

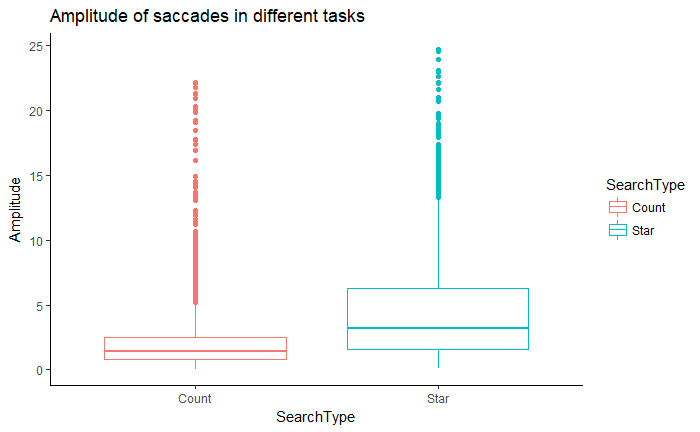


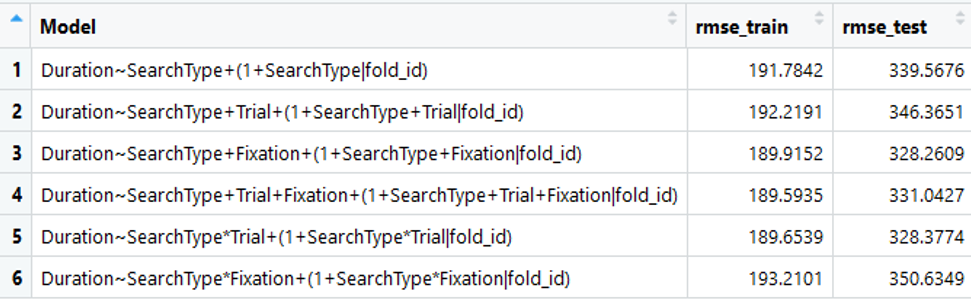
Figure 5. The boxplot for comparison of amplitude of saccades between task conditions.

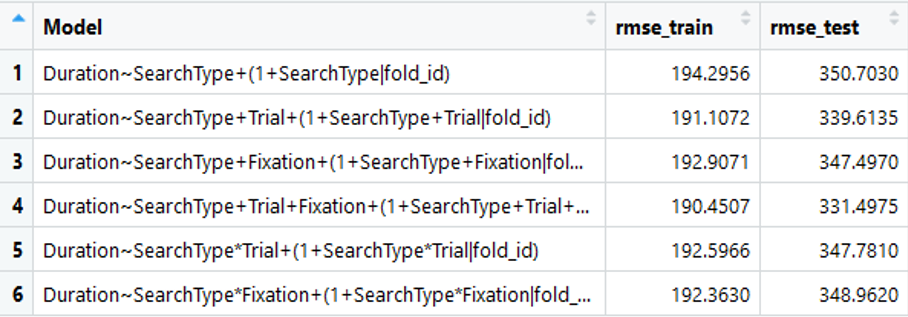
1. *Discussion*

The goal of the experiment was to analyze the differences in visual search pattern depending on behavioral goal of the task. Two variables were chosen to describe the visual search pattern: duration of fixations and amplitude of saccades. The differences in the measures, if found, would confirm that the visual search pattern can be altered under the top-down influence, defined by a type of a given task.

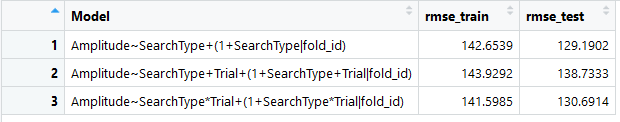
The results from the analysis of the fixation duration supported the first operational hypothesis. The results from the analysis of the saccade amplitude supported the second operational hypothesis. Therefore, the conceptual hypothesis that visual search patterns can be altered depending on a task structure and goals found support in this experimental setting and given the type of the data analysis used. The findings suggest that while performing the visual foraging task, the fixations are shorter, and saccades are longer, than while performing the target counting task.

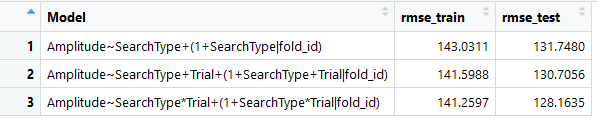
The validity of the findings should be questioned due to serious limitations in terms of quantity of participants and methodological issues. Namely, the measure of root mean square error was not the best criterion for estimating which model produced the best results. After multiple attempts of running the cross-validation process, I found out that the relative rmse measures could not be replicated – the best model differed every time the cross-validation was repeated. For instance, the alternative cross-validation results are presented below:





The alternative results for the second set of models:





As a general recommendation for the further research, a different method for the best model estimation should be used, and more participants should be involved to give the research enough power.