**User Documentation**

*Junyi Zhang, Rebecca Gu, Dejia Su*

*April 2018*

**Overview**

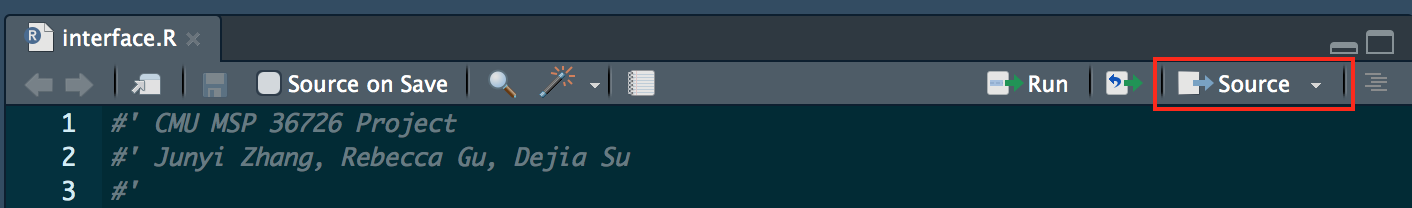
The Automation program serves for cleaning, combining and calculating alternation statistics from the data templates and raw data exported from the eye-tracking device. The Automation program can help researchers clean the data in an efficient and convenient way so that they can understand data better and spend time on valuable things instead of copying and pasting. The program is tested under both Mac OS and Windows OS.

**Installment**

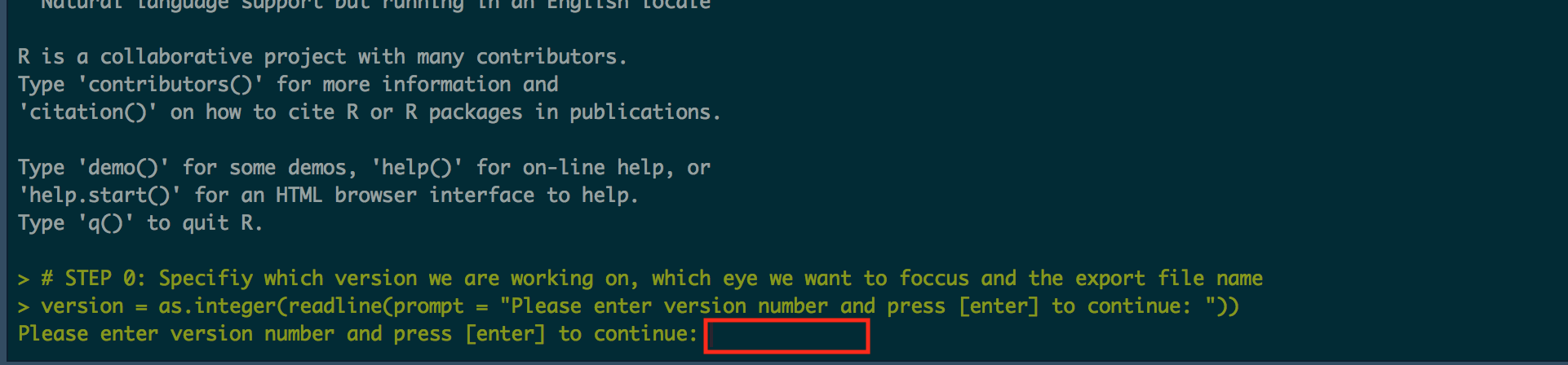
In order to use the Automation program, you need to have R and Rstudio ready on your computer. R is a free software environment for statistical computing and graphics while Rstudio is the IDE for R. R can be downloaded [here](https://www.r-project.org/), and Rstudio can be downloaded [here](https://www.rstudio.com/). Both softwares support Windows, Mac and Linux. After you download the both softwares, the Automation folder could be put anywhere you like. The folder contains two parts: one is interface.R, which is the R script you will interact with; the other one is functions folder, which contains all the necessary functions to perform the whole automation process. In general, you do not need to open the folder. Notice, interface.R and functions folder should always be put in the same working directory.

**How to Use**

1. Open the interface.R using Rstudio.
2. Source the whole file by clicking the [source] bottom.



1. The program will start asking questions regarding the data. You should enter the answer in console and hit enter after typing the answer.



The questions are listed as following:

* 1. Please enter version number and press [enter] to continue: The answer should be a number indicating the version, e.g. 1.
  2. Please enter eye we want to focus (left or right) and press [enter] to continue: The answer should be left or right. The specified eye will be used to calculate the alternation statistics in raw data.
  3. Please enter the export file name (without extension) and press [enter] to continue: The answer should be the desired export file name, e.g. version 1.
  4. Please enter last trial for the first condition and press [enter] to continue: The answer should be an integer indicating the last trial of the first condition. E.g., in version 1, trial 1 to 4 is standard condition while trial 5 to 12 is fully separated. The answer in this case should be 4.
  5. Please enter the first condition name and press [enter] to continue: The answer should be the name of the first condition in the version specified above. E.g., the first four trials in version 1 is standard. The answer should be standard.
  6. Please enter the second condition name and press [enter] to continue: The answer should be the name of the second condition in the version specified above. E.g. the second 8 trials in version 1 is fully separated. The answer could be separated.

1. In this step, you need to choose all the required datasets for automation process. The program will first specify the desired dataset name. After you hit enter, a dialogue box will pop, and then you can choose the desired dataset. Each dataset in this step must contain all the required variables. They are listed in the Appendix 1.
2. After all the datasets are chosen and loaded, the script will run automatically. Three csv files will be produced in the end.
   1. The long clean data named file\_name.csv. The file\_name is specified in step 3 part c. This file contains all the statistics in data templates and alternation statistics. Each row corresponds to one participant in one trial.
   2. The condition 1 summary data named version\_number\_cond1.csv. The version number and condition is specified in step 3 part a and e. This files contains average statistics for each participant in condition 1, where each row corresponds to one participant.
   3. The condition 2 summary data named version\_number\_con2.csv. The version number and condition is specified in step 3 part a and f. This files contains average statistics for each participant in condition w, where each row corresponds to one participant.

A detailed list of variables in each csv files are listed in Appendix 2.

**Remark**:

* The program may take a relatively long time to run for the first time since R need to download and install all the required package. The run time should reduce significantly afterwards.
* Some of warnings are okay such as a certain package is built under a different version of R. In general, it does not affect the usage of the program. Updating R periodically could resolve this issue.
* It is okay for datasets to contain variables that is not required. The program will ignore all the unrecognized variables. However, missing variables will cause an error and a message should be given.
* The alternation statistics only focus on the shift of AOI fixation of a specific eye. i.e. the shift of from fixation on text to fixation on image will count as one shift from text to image. However, the shift from fixation on text to blink does not count.

**Appendix 1 – Required variables in each dataset**

* AOI Blanks
  + Trial
  + Stimulus
  + Participant
  + AOI Name
  + Fixation Count
  + Fixation Time [ms]
  + Fixation Time [%]
* AOI Fixations
  + Trial
  + Stimulus
  + Participant
  + AOI Group
  + Fixation Count
  + Fixation Time [ms]
  + Fixation Time [%]
* Fixation, saccades and blink counts
  + Trial
  + Stimulus
  + Participant
  + Fixation Count
  + Fixation Duration Total [ms]
  + Saccade Count Saccade Duration Total [ms]
  + Saccade Amplitude Total [°]
  + Saccade Velocity Total [°/s]
  + Blink Count
* Trial Duration
  + Stimulus
  + Trial Duration [ms]
  + Participant
  + Tracking Ratio [%]

**Appendix 2 – List of Variables in output file**

* The long clean data

|  |  |
| --- | --- |
| Variable | Comment |
| Participant |  |
| Stimulus |  |
| Trial\_dur | Trial duration in ms |
| Tracking\_ratio |  |
| Trial |  |
| Blank\_fix\_count | Blank fixation count |
| Blank\_fix\_time | Blank fixation time |
| Blanks\_fix\_time\_pct | Blank fixation time percent |
| Image\_fix\_count | Image fixation count |
| Image\_fix\_time | Image fixation time |
| Image\_fix\_time\_pct | Image fixation time percent |
| Text\_fix\_count | Text fixation count |
| Text\_fix\_time | Text fixation time |
| Text\_fix\_time\_pct | Text fixation time percent |
| Fix\_count | Fixation count |
| Total\_fix\_dur | Total fixation duration |
| Saccade\_count |  |
| Total\_saccade\_dur | Total saccade duration |
| Total\_saccade\_amp | Total saccade amplitude |
| Total\_saccade\_velocity |  |
| Blink\_count |  |
| Condition |  |
| max\_text\_time | Single max text fixation duration |
| ws\_Text | Alternation from white space to text count |
| ws\_Image | Alternation from white space to image count |
| ws\_dash | Alternation from white space to dash count |
| Text\_Image | Alternation from text to image count |
| Text\_ws | Alternation from text to white space count |
| Text\_dash | Alternation from text to dash count |
| Image\_Text | Alternation from image to text count |
| Image\_ws | Alternation from image to white space count |
| Image\_dash | Alternation from image to dash count |
| dash\_Image | Alternation from dash to image count |
| dash\_Text | Alternation from dash to text count |
| dash\_ws | Alternation from dash to white space count |
| weighted\_text\_image | Weighted \* alternation from text to image count |
| weighted\_alt | Weighted \* alternation count |
| total\_alt | Total alternation count |
| Version |  |

\* Weighted means we put more weight on alternation at the beginning of each trial and put less weight on alternation at the end of trial. The alternation at time 0 will have weight of 1 and alternation at the end of trial will have weight of 0. The weight decrease at a constant rate. It means the alternations at the beginning of each trial have much more influence on reading pattern than the alternations at the end of each trial do.

* Condition Summary Data

|  |  |
| --- | --- |
|  |  |
| Variables | Comment |
| Participant |  |
| Avg\_trial\_dur | Average trial duration |
| Avg\_blank\_fix\_count | Average blank fixation count |
| Avg\_blank\_fix\_time | Average blank fixation time |
| Avg\_blank\_fix\_time\_pct | Average blank fixation time percent |
| Avg\_image\_fix\_count | Average image fixation count |
| Avg\_image\_fix\_time | Average image fixation time |
| Avg\_image\_fix\_time\_pct | Average image fixation time percent |
| Avg\_text\_fix\_count | Average text fixation count |
| Avg\_text\_fix\_time | Average text fixation time |
| Avg\_text\_fix\_time\_pct | Average text fixation time percent |
| Avg\_fix\_count | Average fixation count |
| Avg\_total\_fix\_dur | Average total fixation duration |
| Avg\_saccade\_count | Average saccade count |
| Avg\_total\_saccade\_dur | Average total saccade duration |
| Avg\_total\_saccade\_amp | Average total saccade amplitude |
| Avg\_total\_saccade\_velocity | Average total saccade velocity |
| Avg\_blink\_count | Average blink count |
| Avg\_ws\_Text | Average alternation from white space to text count |
| Avg\_ws\_Image | Average alternation from white space to image count |
| Avg\_ws\_dash | Average alternation from white space to dash count |
| Avg\_Text\_Image | Average alternation from text to image count |
| Avg\_Text\_ws | Average alternation from text to white space count |
| Avg\_Text\_dash | Average alternation from text to dash count |
| Avg\_Image\_Text | Average alternation from image to text count |
| Avg\_Image\_ws | Average alternation from image to white space count |
| Avg\_Image\_dash | Average alternation from image to dash count |
| Avg\_dash\_Image | Average alternation from dash to image count |
| Avg\_dash\_Text | Average alternation from dash to text count |
| Avg\_dash\_ws | Average alternation from dash to white space count |
| Avg\_total\_alt | Average total alternation count |
| Avg\_weighted\_alt | Average weighted alternation count |
| Version |  |