Basic Research Methods: eye tracking tutorial

MagdaLena Matyjek

10/01/2020

Hands-on exercise

Task:

- a passive-viewing task
- Three conditions: positive, negative and neutral pictures from IAPS (International Affective Picture System)
- 3000 ms presentation, 500 ms ITI
- 30 x each condition (no repeating items)
- time: ~5 min

Sample:

- you!

Procedure

1. Positioning
2. Calibration (+ re-calibration)
3. Data collection
4. Data pre-processing
5. Data visualisation
6. Data interpretation

Hands-on time!

Now we collect data from volunteers. Meanwhile, form groups of 5-10 people and prepare the experimental design and predictions by considering the following points:

1. **Experimental design**:

   - conditions (3: pos, neg, neu)
   - no of repetitions of stimuli
   - sample size + statistical power
   - type of statistical analysis

2. **Research questions / hypotheses**

   - backward reasoning: what could be the research question asked in this study?
   - hypotheses and predictions: based on what you already know about eye tracking, pupillometry, and processing of emotional and neutral images in humans, propose your hypotheses and predictions for our results.

Raw data

```
# show the first 4 rows
head(rd,4)
```

```
##                   timestamp                     time
## 1                       MSG 2020-01-07 10:17:43.26 1061923
## 2 2020-01-07 10:17:43.023                1061923567
## 3 2020-01-07 10:17:43.040                1061923584
## 4 2020-01-07 10:17:43.056                1061923600
##               state     rawx     rawy     avgx     avgy
## 1 start_recording       NA       NA       NA       NA
## 2                 7 1063.299 719.7043 1036.940 699.1066 26
## 3                 7 1040.149 710.0520 1037.394 698.9050 25
## 4                 7 1051.085 682.4178 1038.015 699.0016 25
##      Lrawy    Lavgx    Lavgy  Lpsize Lpupilx Lpupily
## 1       NA       NA       NA       NA       NA       NA
## 2 757.5413 1050.300 710.9240 24.6516  0.3618  0.5723 108
## 3 739.2700 1050.397 712.3792 24.8139  0.3617  0.5719 102
## 4 715.2422 1050.301 712.5284 24.4666  0.3614  0.5715 105
```

Structured raw data

```
# show selected columns
data <- rd[,c(1,3:5,13,20)]
# ...and the first 5 rows
head(data,5)
```

```
##                   timestamp              fix      rawx      r
## 1 2020-01-07 10:17:43.259 start_recording        NA
## 2 2020-01-07 10:17:43.023             True 1063.299 719.7
## 3 2020-01-07 10:17:43.039             True 1040.149 710.0
## 4 2020-01-07 10:17:43.055             True 1051.085 682.4
## 5 2020-01-07 10:17:43.072             True 1130.159 803.0
##    Rpsize
## 1      NA
## 2 27.8472
## 3 26.6744
## 4 27.4394
## 5 31.6633
```

Pre-processing

Steps of pre-processing...

Pre-processed data

```
##     code  t_ms   sample cond trial_no segmentName size_fc
## 1   1111 4.361 24.52612  NEG        1    baseline -0.0616
## 2   1111 4.362 24.52464  NEG        1    baseline -0.0631
## 3   1111 4.363 24.52319  NEG        1    baseline -0.0646
## 4   1111 4.364 24.52176  NEG        1    baseline -0.0660
## 5   1111 4.365 24.52036  NEG        1    baseline -0.0674
## 6   1111 4.366 24.51898  NEG        1    baseline -0.0688
## 7   1111 4.367 24.51763  NEG        1    baseline -0.0701
## 8   1111 4.368 24.51632  NEG        1    baseline -0.0714
## 9   1111 4.369 24.51503  NEG        1    baseline -0.0727
## 10 1111 4.370 24.51377  NEG        1    baseline -0.0740
## 11 1111 4.371 24.51255  NEG        1    baseline -0.0752
## 12 1111 4.372 24.51137  NEG        1    baseline -0.0764
## 13 1111 4.373 24.51022  NEG        1    baseline -0.0775
## 14 1111 4.374 24.50911  NEG        1    baseline -0.0786
## 15 1111 4.375 24.50803  NEG        1    baseline -0.0797
## 16 1111 4.376 24.50700  NEG        1    baseline -0.0808
## 17 1111 4.377 24.50600  NEG        1    baseline -0.0818
```

Basic descriptives

Sample size:

```
## [1] 1
```

Mean, max, and min (relative to baseline):

```
## [1] 0.53
```

```
## [1] 4.35
```

```
## [1] -3.15
```

Question: What do negative values mean?

Aggregate data

Question: What is data aggregation?

For plotting, we need the time course, but for statistics we have to aggregate the data over conditions and a time window. Here we will use 1 - 3 secs.

Question: Why won't we use the data < 1 sec?

```
# For plotting:
data_P_av_trials <- aggregate(list(data$size_fdb),by=list(c
colnames(data_P_av_trials) <- c("cond","t","trial","size")

data_P_av <- aggregate(list(data$size_fdb),by=list(data$cor
colnames(data_P_av) <- c("cond","t","size")

# For stats:
data2 <- data[data$t_fdb >= 1 & data$t_fdb <= 3,]
data_av_trials <- aggregate(list(data2$size_fdb),by=list(da
colnames(data_av_trials) <- c("cond","trial","size")
```

Plotting the data - visual inspection

Let's plot the averaged data over participants for each condition (the grey bands are the 95% confidence intervals):

```
## Warning: package 'ggplot2' was built under R version 3.5

## `geom_smooth()` using method = 'gam' and formula 'y ~ s(
```