Raw Reaction Time Example

Christopher L. Aberson

Last Knitted: 2023-09-18

### Vignette Setup:

### Project/Data Title:

Continuous lexical decision task: classification of Dutch words as either actual words or nonwords

Data provided by: Tom Heyman

### Project/Data Description:

Data come from a study reported in Heyman, De Deyne, Hutchison, & Storms (2015, Behavior Research Methods; henceforth HDHS). More specifically, the study involved a continuous lexical decision task intended to measure (item-level) semantic priming effects (i.e., Experiment 3 of HDHS). It is similar to the SPAML set-up (see <https://osf.io/q4fjy/>), but with fewer items and participants. The study had several goals, but principally we wanted to examine how a different/new paradigm called the speeded word fragment completion task would compare against a more common, well-established paradigm like lexical decision in terms of semantic priming (i.e., magnitude of the effect, reliability of item-level priming, cross-task correlation of item-level priming effects, etc.). Experiment 3 only involved a continuous lexical decision task, so the datafile contains no data from the speeded word fragment completion task.

### Methods Description:

Participants were 40 students from the University of Leuven, Belgium (10 men, 30 women, mean age 20 years). A total of 576 pairs were used in a continuous lexical decision task (so participants do not perceive them as pairs): 144 word–word pairs, 144 word–pseudoword pairs, 144 pseudoword–word pairs, and 144 pseudoword–pseudoword pairs. Of the 144 word-word pairs, 72 were fillers and 72 were critical pairs, half of which were related, the other half unrelated (this was counterbalanced across participants). The dataset only contains data for the critical pairs. Participants were informed that they would see a letter string on each trial and that they had to indicate whether the letter string formed an existing Dutch word or not by pressing the arrow keys. Half of the participants had to press the left arrow for word and the right arrow for nonword, and vice versa for the other half.

### Data Location:

<https://osf.io/frxpd/>

The example dataset also includes *R* scripts at this location that used Accuracy in Parameter Estimation in a different fashion.

HDHS<- read.csv("HDHSAIPE.txt", sep="")  
str(HDHS)

## 'data.frame': 2880 obs. of 8 variables:  
## $ RT : num 0.52 0.453 0.467 0.534 0.573 ...  
## $ zRT : num -0.303 -0.492 -0.453 -0.265 -0.153 ...  
## $ Pp : int 1 1 1 1 1 1 1 1 1 1 ...  
## $ Type : chr "R" "R" "R" "R" ...  
## $ Prime : chr "hengst" "matrak" "eland" "erwt" ...  
## $ Target : chr "veulen" "wapen" "gewei" "wortel" ...  
## $ accTarget: int 1 1 1 1 1 1 1 1 1 0 ...  
## $ accPrime : int 1 1 1 1 0 1 1 1 1 0 ...

### Date Published:

2022-02-04

### Dataset Citation:

Heyman, T. (2022, February 4). Dataset AIPE. Retrieved from osf.io/frxpd [based on Heyman, T., De Deyne, S., Hutchison, K. A., & Storms, G. (2015). Using the speeded word fragment completion task to examine semantic priming. Behavior Research Methods, 47(2), 580-606.]

### Keywords:

Semantic priming; continuous lexical decision task

### Use License:

CC-By Attribution 4.0 International

### Geographic Description - City/State/Country of Participants:

Belgium

### Column Metadata:

metadata <- import("HDHSMeta.txt")  
  
flextable(metadata) %>% autofit()

| Variable Name | Variable Description | Type (numeric, character, logical, etc.) |
| --- | --- | --- |
| RT | Response time to the target in seconds | Numeric |
| zRT | Z-transformed target response times per participant | Numeric |
| Pp | Participant identifier (1 to 40) | Integer |
| Type | Whether target was preceded by a related prime (<93>R<94>) or an unrelated prime (<93>U<94>) | Character |
| Prime | Prime stimulus (in Dutch) | Character |
| Target | Target stimulus (in Dutch) | Character |
| accTarget | Whether response to target was correct (1) or not (0) | Integer |
| accPrime | Whether response to the preceding prime was correct (1) or not (0) | Integer |

### AIPE Analysis:

# pick only correct answers  
HDHScorrect <- HDHS[HDHS$accTarget==1,]   
summary\_stats <- HDHScorrect %>% #data frame  
 select(RT, Target) %>% #pick the columns  
 group\_by(Target) %>% #put together the stimuli  
 summarize(SES = sd(RT)/sqrt(length(RT)), samplesize = length(RT)) #create SE and the sample size for below  
##give descriptives of the SEs  
describe(summary\_stats$SES)

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 72 0.05 0.05 0.03 0.04 0.02 0.02 0.31 0.29 3.63 13.19 0.01

##figure out the original sample sizes (not really necessary as all Targets were seen by 40 participants)  
original\_SS <- HDHS %>% #data frame  
 count(Target) #count up the sample size  
##add the original sample size to the data frame  
summary\_stats <- merge(summary\_stats, original\_SS, by = "Target")  
##original sample size average  
describe(summary\_stats$n)

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 72 40 0 40 40 0 40 40 0 NaN NaN 0

##reduced sample size  
describe(summary\_stats$samplesize)

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 72 38.12 3.09 39 38.83 1.48 22 40 18 -3.29 12.08 0.36

##percent retained  
describe(summary\_stats$samplesize/summary\_stats$n)

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 72 0.95 0.08 0.98 0.97 0.04 0.55 1 0.45 -3.29 12.08 0.01

flextable(head(HDHScorrect)) %>% autofit()

| RT | zRT | Pp | Type | Prime | Target | accTarget | accPrime |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0.5202093 | -0.3026836 | 1 | R | hengst | veulen | 1 | 1 |
| 0.4532606 | -0.4918172 | 1 | R | matrak | wapen | 1 | 1 |
| 0.4670391 | -0.4528923 | 1 | R | eland | gewei | 1 | 1 |
| 0.5335296 | -0.2650529 | 1 | R | erwt | wortel | 1 | 1 |
| 0.5732744 | -0.1527717 | 1 | R | ijzel | glad | 1 | 0 |
| 0.3870141 | -0.6789673 | 1 | R | sauna | warm | 1 | 1 |

#### Stopping Rule

What the usual standard error for the data that could be considered for our stopping rule?

SE <- tapply(HDHScorrect$RT, HDHScorrect$Target, function (x) { sd(x)/sqrt(length(x)) })  
min(SE)

## [1] 0.01511263

max(SE)

## [1] 0.3058638

cutoff <- quantile(SE, probs = .4)  
cutoff

## 40%   
## 0.03113963

The items have a range of 0.0151126 to 0.3058638. We could use the 40% decile SE = 0.0311396 as our critical value for our stopping rule, as suggested by the manuscript analysis. We could also have a set SE to a specific target if we do not believe we have representative pilot data in this example. You should also consider the scale when estimating these values (i.e., millisecond data has more room to vary than other smaller scales).

#### Minimum Sample Size

To estimate minimum sample size, we should figure out what number of participants it would take to achieve 80% of the SEs for items below our critical score of 0.0311396?

# sequence of sample sizes to try  
samplesize\_values <- seq(20, 500, 5)  
  
# create a blank table for us to save the values in   
sim\_table <- matrix(NA,   
 nrow = length(samplesize\_values),   
 ncol = length(unique(HDHS$Target)))  
  
# make it a data frame  
sim\_table <- as.data.frame(sim\_table)  
  
# add a place for sample size values   
sim\_table$sample\_size <- NA  
  
# loop over sample sizes  
for (i in 1:length(samplesize\_values)){  
   
 # temp dataframe that samples and summarizes  
 temp <- HDHScorrect %>%   
 group\_by(Target) %>%   
 sample\_n(samplesize\_values[i], replace = T) %>%   
 summarize(se = sd(RT)/sqrt(length(RT)))   
   
 colnames(sim\_table)[1:length(unique(HDHScorrect$Target))] <- temp$Target  
 sim\_table[i, 1:length(unique(HDHScorrect$Target))] <- temp$se  
 sim\_table[i, "sample\_size"] <- samplesize\_values[i]  
 }  
  
final\_sample <-   
 sim\_table %>%   
 pivot\_longer(cols = -c(sample\_size)) %>%   
# rename(item = variable, se = value) %>%   
 group\_by(sample\_size) %>%   
 summarize(percent\_below = sum(value <= cutoff)/length(unique(HDHScorrect$Target))) %>%   
 ungroup()  
  
flextable(final\_sample %>% head()) %>% autofit()

| sample\_size | percent\_below |
| --- | --- |
| 20 | 0.3888889 |
| 25 | 0.4027778 |
| 30 | 0.4305556 |
| 35 | 0.4861111 |
| 40 | 0.5416667 |
| 45 | 0.5972222 |

#### Expected Data Loss

The original data found that 4.6875 percent of the data were unusable.

#### Minimum Sample Size

# use semanticprimer cutoff function for prop variance  
cutoff <- calculate\_cutoff(population = HDHScorrect,   
 grouping\_items = "Target",  
 score = "RT",  
 minimum = as.numeric(min(HDHScorrect$RT)),  
 maximum = as.numeric(max(HDHScorrect$RT)))  
# showing how this is the same as the person calculated version versus semanticprimeR's function  
cutoff$cutoff

## 40%   
## 0.03113963

final\_table <- calculate\_correction(  
 proportion\_summary = final\_sample,  
 pilot\_sample\_size = HDHScorrect %>% group\_by(Target) %>%   
 summarize(sample\_size = n()) %>%   
 ungroup() %>% summarize(avg\_sample = mean(sample\_size)) %>%   
 pull(avg\_sample),  
 proportion\_variability = cutoff$prop\_var  
 )  
  
flextable(final\_table) %>%   
 autofit()

| percent\_below | sample\_size | corrected\_sample\_size |
| --- | --- | --- |
| 81.94444 | 80 | 72.49787 |
| 88.88889 | 120 | 103.12059 |
| 90.27778 | 155 | 126.20112 |

Based on these simulations, we can decide our minimum sample size is likely close to 72 and 75.375 including information about data loss.

#### Maximum Sample Size

In this example, we could set our maximum sample size for 90% power (as defined as 90% of items below our criterion), which would equate to 126 and 131.90625 with the expected data loss. The final table does not include 95% of items below our criterion, even after estimating 500 participants. An investigation of the table indicates that it levels off at 93-94%.

#### Final Sample Size

In any estimate of sample size, you should also consider the potential for missing data and/or unusable data due to any other exclusion criteria in your study (i.e., attention checks, speeding, getting the answer right, etc.). In this study, we likely expect all participants to see all items, and therefore, we could expect to use the minimum sample size as our final sample size, the point at which all items reach our SE criterion, or the maximum sample size. Note that maximum sample sizes can also be defined by time, money, or other means.