

# KH2017 analyses (Behavior Research Methods, 2017)

*Pascal J. Kieslich & Felix Henninger*

## Reference

Kieslich, P. J., & Henninger, F. (2017). Mousetrap: An integrated, open-source mouse-tracking package. *Behavior Research Methods*, 49(5), 1652-1667. <https://doi.org/10.3758/s13428-017-0900-z>

## General preparations

### Load libraries

```
library(mousetrap)
library(ggplot2)
library(dplyr)
```

### Custom ggplot2 theme

```
theme_set(theme_classic()+
  theme(
    axis.line = element_line(colour = "black"),
    axis.ticks = element_line(colour = "black"),
    axis.text = element_text(colour = "black"),
    panel.border = element_rect(colour = "black", fill=NA)
  ))

options(width=90)
```

## Data preparation

### Read in merged raw data

```
# Read in raw data from csv.gz file
# (commented out as the raw data are now included in the mousetrap R package)
# raw_data <- read.csv("merged_data/raw_data.csv.gz", stringsAsFactor=FALSE)

# Instead, get raw data directly from data stored in mousetrap package
raw_data <- KH2017_raw
```

### Filter trials

```
# Look at percent of correct responses per Condition
with(raw_data, table(Condition, correct)/c(table(Condition)))
```

```
##           correct
## Condition      0      1
## Atypical 0.11111111 0.88888889
## Typical 0.04615385 0.95384615
# Exclude incorrect trials
raw_data <- subset(raw_data, correct==1)

# Check that each subject has at least one typical and atypical trial left
table(with(raw_data, table(subject_nr, Condition))>0)

##
## TRUE
## 120
```

## Import mouse-tracking data

```
mt_data <- mt_import_mousetrap(raw_data)
```

## Trajectory preprocessing

```
mt_data <- mt_remap_symmetric(mt_data)
mt_data <- mt_align_start(mt_data)
mt_data <- mt_measures(mt_data)
mt_data <- mt_time_normalize(mt_data)
```

## Analyses

### Logging resolution

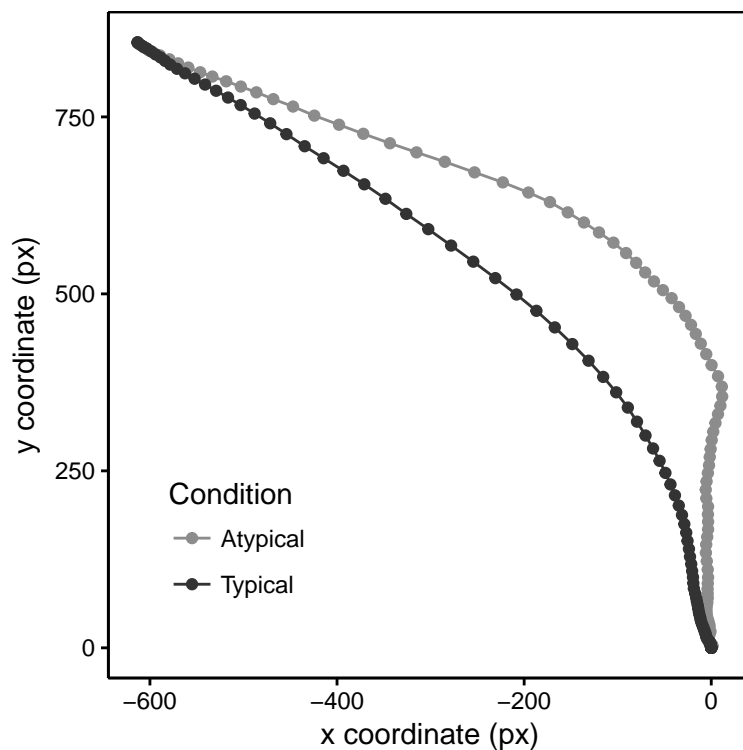
```
mt_check_resolution(mt_data, desired = 10)
```

```
## $summary
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  1.000  10.000  10.000   9.981  10.000  40.000
##
## $sd
## [1] 0.430673
##
## $frequencies
## log_diffs
##      1      2      3      4      5      6      7      8      9     10     11     12
##    93    133     85    128     92    146     63    124     85 212191    182    10
##    13     14     15     16     17     20     23     25     27     29     32     33
##     9     10      4      3      3      1      1      1      2      2      3      3
##    34     36     40
##     2      2      1
##
## $relative_frequencies
## log_diffs
```

```
##      1      2      3      4      5      6      7      8      9     10     11     12
## 0.0004 0.0006 0.0004 0.0006 0.0004 0.0007 0.0003 0.0006 0.0004 0.9944 0.0009 0.0000
##     13     14     15     16     17     20     23     25     27     29     32     33
## 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
##     34     36     40
## 0.0000 0.0000 0.0000
##
## $frequencies_desired
## log_diffs_class
## smaller desired greater
##      949 212191      239
##
## $relative_frequencies_desired
## log_diffs_class
## smaller desired greater
## 0.0044 0.9944 0.0011
```

## Plot average time-normalized trajectories

```
mt_plot_aggregate(mt_data, use = "tn_trajectories", points=TRUE,
  x = "xpos", y = "ypos", color = "Condition", subject_id = "subject_nr")+
  theme(legend.position=c(.2,.2))+
  xlab("x coordinate (px)") + ylab("y coordinate (px)")+
  scale_color_grey(start = .55, end = .20)
```



```
# ggsave("figures/Figure5.png",width=8.5, height=10,unit="cm",dpi=1200)
# ggsave("figures/Figure5.pdf",width=8.5, height=10,unit="cm")
# ggsave("figures/Figure5.eps",width=8.5, height=10,unit="cm")
```

## Comparison of measures based on time-normalized vs. raw trajectories

```
mt_data <- mt_measures(mt_data, use="tn_trajectories", save_as="tn_measures")
diag(cor(mt_data$measures[, -1], mt_data$tn_measures[, -1]))
```

##	xpos_max	xpos_min	ypos_max	ypos_min	MAD
##	0.9999950	0.9999948	0.9999992	0.9985444	0.9998875
##	MAD_time	MD_above	MD_above_time	MD_below	MD_below_time
##	0.9990594	0.9999922	0.9978766	0.9998770	0.9718700
##	AD	AUC	xpos_flips	ypos_flips	xpos_reversals
##	0.9999912	0.9999688	0.9977138	0.9897775	1.0000000
##	ypos_reversals	RT	initiation_time	idle_time	
##	1.0000000	1.0000000	0.9997353	0.9929374	

## Comparison of MAD aggregated per participant

```
agg_mad <- mt_aggregate_per_subject(mt_data, subject_id = "subject_nr",
  use_variables = "MAD", use2_variables = "Condition")

agg_mad %>%
  group_by(Condition) %>%
  select(MAD, Condition) %>%
  summarize_all(.funs = c("mean", "sd"))
```

##	#	A tibble:	2	x	3
##	Condition	mean	sd		
##	<fctr>	<dbl>	<dbl>		
##	1	Atypical	343.7954	218.6368	
##	2	Typical	172.2093	110.8397	

```
t.test(MAD~Condition, data=agg_mad, paired=TRUE)
```

```
##
## Paired t-test
##
## data: MAD by Condition
## t = 6.731, df = 59, p-value = 7.706e-09
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 120.5772 222.5949
## sample estimates:
## mean of the differences
## 171.5861
```

```
# Cohen's dz
mad_diff <- with(agg_mad, MAD[Condition=="Atypical"] - MAD[Condition=="Typical"])
mean(mad_diff)/sd(mad_diff)
```

```
## [1] 0.8689738
```

## Bimodality analyses

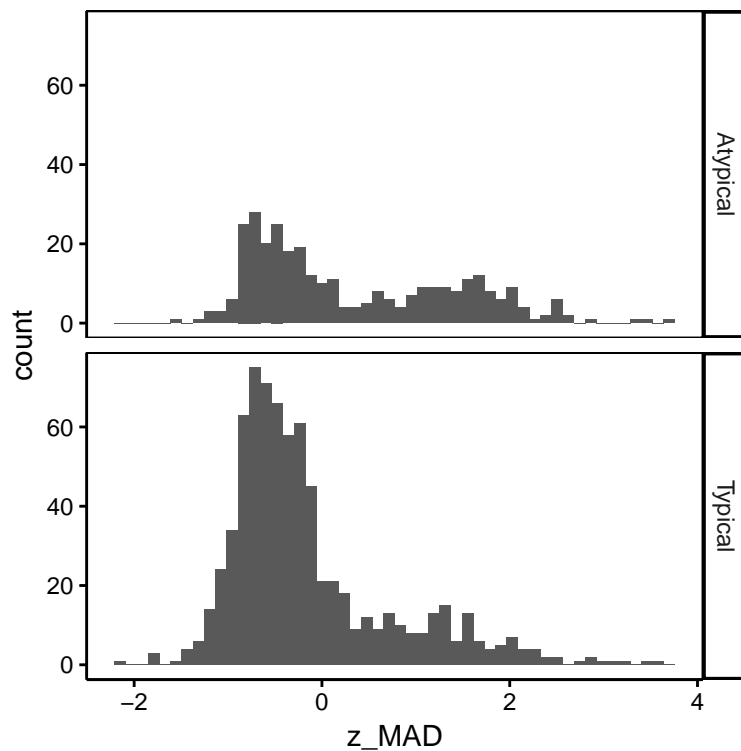
```
# Standardize MAD per participant
mt_data <- mt_standardize(mt_data,
  use_variables = "MAD", within = "subject_nr", prefix = "z_")

# Calculate bimodality coefficient
mt_check_bimodality(mt_data, use_variables = "z_MAD",
  grouping_variables = "Condition", methods = "BC")

## $BC
## Condition z_MAD
## 1 Atypical 0.5934582
## 2 Typical 0.6078283

# Merge trial level data (needed for distribution qplot with facets)
results <- merge(mt_data$data, mt_data$measures, by="mt_id")

# Distribution
qplot(x=z_MAD, data=results, bins=50)+facet_grid(Condition~.)
```



## Repeat analyses without large MAD values

```
# Exclude trials with abs(z_MAD)>1.5
table(abs(mt_data$measures$z_MAD)<=1.5)/nrow(mt_data$measures)

##
## FALSE TRUE
```

```

## 0.1165414 0.8834586
mt_data_sub <- mt_subset(mt_data,abs(z_MAD)<=1.5,check="measures")

# Check that each subject has at least on typical and atypical trial left
table(with(mt_data_sub$data,table(subject_nr,Condition))>0)

##
## TRUE
## 120

# Comparison of MAD aggregated per participant
agg_mad_sub <- mt_aggregate_per_subject(mt_data_sub, subject_id = "subject_nr",
  use_variables = "MAD", use2_variables = "Condition")

agg_mad_sub %>%
  group_by(Condition) %>%
  select(MAD,Condition) %>%
  summarize_all(.funs = c("mean","sd"))

## # A tibble: 2 x 3
##   Condition    mean      sd
##   <fctr>      <dbl>   <dbl>
## 1 Atypical  200.7050 209.2332
## 2 Typical   121.7434 117.4904

t.test(MAD~Condition, data=agg_mad_sub, paired=TRUE)

##
## Paired t-test
##
## data: MAD by Condition
## t = 4.1599, df = 59, p-value = 0.0001047
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  40.97963 116.94341
## sample estimates:
## mean of the differences
##              78.96152

```