

# Data Sets for Many-Analyses

Julia Haaf

2/28/2020

This document takes the original data and subsets it using different exclusion criteria. It is therefore a documentation of all data sets used in the many-analyses.

```
sitesource_label = case_when(source == "uff" ~ "University of Florida", source == "occid" ~ "Occidental College", source == "ashland" ~ "Ashland University", source == "ithaca" ~ "Ithaca College", source == "riverside" ~ "University of California, Riverside", source == "wesleyan_inhouse" ~ "Wesleyan University", source == "uwmadison_expert" ~ "University of Wisconsin", source == "uwmadison_inhouse" ~ "University of Wisconsin", source == "vcu" ~ "Virginia Commonwealth University", source == "sou_inhouse" ~ "Southern Oregon University", source == "plu" ~ "Pacific Lutheran University", source == "byui" ~ "Brigham Young University - Idaho", source == "azusa" ~ "Azusa Pacific University", source == "cnj" ~ "The College of New Jersey", source == "wpi" ~ "Worcester Polytechnic Institute", source == "illinois" ~ "University of Illinois", source == "kansas_expert" ~ "University of Kansas", source == "kansas_inhouse" ~ "University of Kansas", source == "upenn" ~ "University of Pennsylvania", source == "pace_inhouse" ~ "Pace University", source == "pace_expert" ~ "Pace University") )
```

```
merged <- readRDS("data/merged_deidentified_full.rds")
```

```
# head(merged)
nrow(merged)
```

```
## [1] 2281
```

```
# table(merged$race)
# sum(is.na(merged$race))
# sum(is.na(merged$countryofbirth))
```

```
levels(merged$countryofbirth)
```

```
## [1] "" "1"
## [3] "2" "4"
## [5] "6" "America"
## [7] "Bulgaria" "china"
## [9] "China" "India"
## [11] "India " "Mexico"
## [13] "Philippines" "South Korea"
## [15] "South Korea " "The United States of America."
## [17] "U.S" "U.S.A"
## [19] "United State" "united states"
## [21] "United States" "united states "
## [23] "United States of America" "us"
## [25] "US" "usa"
## [27] "USA"
```

```
merged$usborn <- ifelse(merged$countryofbirth %in% c(1, "united states", "United State", "United States", "us", "US", "usa", "USA", "U.S.A", "U.S.", "The United States of America.", "America"), 1, 0)
```

```
levels(merged$race)
```

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

```

## [2] "0"
## [3] "1"
## [4] "1 AND 2"
## [5] "1 AND 4 AND 5"
## [6] "1 AND 6"
## [7] "1,2"
## [8] "1,2,4"
## [9] "1,3"
## [10] "1,3,4"
## [11] "1,4"
## [12] "1,4,5"
## [13] "1,4,6"
## [14] "1,5"
## [15] "1,6"
## [16] "2"
## [17] "2,3,4"
## [18] "2,6"
## [19] "3"
## [20] "4"
## [21] "4,5"
## [22] "4,6"
## [23] "5"
## [24] "6"
## [25] "African American and White "
## [26] "Caucasion & Hispanic/Latina"
## [27] "hispanic"
## [28] "Hispanic"
## [29] "hispanic "
## [30] "Irish American "
## [31] "Latina"
## [32] "Latino"
## [33] "Muslim/American"
## [34] "N"
## [35] "Thai and Caucasian"
## [36] "white "
## [37] "white/ native american"
## [38] "White/Caucasian "
## [39] "White/Caucasian/Welsh,German,Italian"

merged$iswhite <- ifelse(merged$race %in% c(1, "White/Caucasian ", "White/Caucasian/Welsh,German,Italian",
, "1 AND 2", "1 AND 4 AND 5", "1 AND 6", "1,2", "1,2,4", "1,3", "1,3,4", "1,4", "1,4,5", "1,4,6", "1,5", "1,6", "white/ native american", "African American and White ", "Caucasion & Hispanic/Latina", "Thai and Caucasian"), 1, 0)

# Create variables, indexes, and exclusion rules ----
# compute exclusion rules
merged <- mutate(merged,
  # Exclusion rule 1:
  #1. Wrote something for both writing prompts
  #2. Completed all six items evaluating the essay authors)
  pass_ER1 = !is.na(pro_minus_anti) & (msincomplete == 0 | is.na(msincomplete)) & # completed all 6 items
  !is.na(prou3) & !is.na(prou4) & !is.na(prou5) & # P provided all 3 ratings of pr
  !is.na(anti3) & !is.na(anti4) & !is.na(anti5), # P provided all 3 ratings of an
  # Exclusion rule 2:

```

```

# as above, plus
#3. Identify as White (race == 1)
#4. Born in USA (countryofbirth == 1)
pass_ER2 = pass_ER1 &
  (iswhite == 1) & # white ps, NA race discarded
  (usborn == 1),
# Exclusion rule 3:
# as above, plus
# 5. Score a 7 or higher on the American Identity item
pass_ER3 = pass_ER2 &
  (americanid >= 7 & !is.na(americanid)) # strongly ID as american, NAs discarded
)

```

## Exclusion Criteria

Person-level exclusion (Taken from Klein et al., 2019):

1. All participants who did not complete the materials.
2. 1 + All participants who do not identify as white and who were born outside of the United States.
3. 2 + All participants who responded lower than 7 on an American Identity item.

Study-level exclusion based on  $N$  (from the preregistration):

1. No study-level exclusion.
2. All studies that have less than 60 participants collected.
3. All studies that have less than 40 participants per cell.

Study-level exclusion based on expert advice (from the comment):

1. No exclusion based on expert advice.
2. In-house studies are excluded.

All exclusion criteria are crossed in the following way:

```

person.ex <- 1:3
n.ex <- 1:3
expert.ex <- 1:2

crit <- expand.grid(person.ex, n.ex, expert.ex)
nrow(crit)

## [1] 18

```

## Original Analysis

We aim at reanalyzing the key findings of Klein et al (2019) using the exclusion settings (1,1,1), (2,1,1), and (3,1,1).

```

merged.111 <- subset(merged, pass_ER1 == TRUE)
nrow(merged.111)

## [1] 2211

#write data
write.csv2(merged.111, "data/reanalysis_111.csv", row.names = FALSE)
metaset.111 <- run.analysis(merged.111)
write.csv2(metaset.111, "data/metaset_111.csv", row.names = FALSE)

```

```
merged.211 <- subset(merged, pass_ER2 == TRUE)
nrow(merged.211)

## [1] 637

#write data
write.csv2(merged.211, "data/reanalysis_211.csv", row.names = FALSE)
metaset.211 <- run.analysis(merged.211)
write.csv2(metaset.211, "data/metaset_211.csv", row.names = FALSE)

merged.311 <- subset(merged, pass_ER3 == TRUE)
nrow(merged.311)

## [1] 277

#write data
write.csv2(merged.311, "data/reanalysis_311.csv", row.names = FALSE)
metaset.311 <- run.analysis(merged.311)
write.csv2(metaset.311, "data/metaset_311.csv", row.names = FALSE)
```

## Analysis for the Main Claim of the Comment

We aim at reanalyzing the main claim of Chatard et al (2020) using the exclusion settings (1,3,2), (2,3,2), and (3,3,2). Note that the study-level exclusions are based on the number of participants for exclusion criterion 1.

```
nbysourcebycond <- with(merged, tapply(dv_order, list(source, ms_condition), length))
include3 <- names(which(rowSums(nbysourcebycond >= 40) == 2))
```

```
merged.132 <- subset(merged.111, source %in% include3 & expert == 1)
nrow(merged.132)
```

```
## [1] 700
```

```
write.csv2(merged.132, "data/reanalysis_132.csv", row.names = FALSE)
metaset.132 <- run.analysis(merged.132)
write.csv2(metaset.132, "data/metaset_132.csv", row.names = FALSE)
```

```
merged.232 <- subset(merged.211, source %in% include3 & expert == 1)
nrow(merged.232)
```

```
## [1] 386
```

```
write.csv2(merged.232, "data/reanalysis_232.csv", row.names = FALSE)
metaset.232 <- run.analysis(merged.232)
write.csv2(metaset.232, "data/metaset_232.csv", row.names = FALSE)
```

```
merged.332 <- subset(merged.311, source %in% include3 & expert == 1)
nrow(merged.332)
```

```
## [1] 229
```

```
write.csv2(merged.332, "data/reanalysis_332.csv", row.names = FALSE)
metaset.332 <- run.analysis(merged.332)
write.csv2(metaset.332, "data/metaset_332.csv", row.names = FALSE)
```

```
metafor::rma(metaset.332$yi, sei = metaset.332$sei, method="FE")
```

```
##
```

```
## Fixed-Effects Model (k = 7)
```

```
##
## Test for Heterogeneity:
## Q(df = 6) = 2.8145, p-val = 0.8317
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
##    0.1930    0.1351    1.4291    0.1530   -0.0717    0.4577
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## All Other Variants

### Exclusion .1.2

```
merged.112 <- subset(merged.111, expert == 1)
nrow(merged.112)

## [1] 799

write.csv2(merged.112, "data/reanalysis_112.csv", row.names = FALSE)
metaset.112 <- run.analysis(merged.112)
write.csv2(metaset.112, "data/metaset_112.csv", row.names = FALSE)

merged.212 <- subset(merged.211, expert == 1)
nrow(merged.212)

## [1] 463

write.csv2(merged.212, "data/reanalysis_212.csv", row.names = FALSE)
metaset.212 <- run.analysis(merged.212)
write.csv2(metaset.212, "data/metaset_212.csv", row.names = FALSE)

merged.312 <- subset(merged.311, expert == 1)
nrow(merged.312)

## [1] 277

write.csv2(merged.312, "data/reanalysis_312.csv", row.names = FALSE)
metaset.312 <- run.analysis(merged.312)
write.csv2(metaset.312, "data/metaset_312.csv", row.names = FALSE)
```

### Exclusion .2.1

```
nbysource <- with(merged, tapply(dv_order, source, length))
include2 <- names(which(nbysource > 60))

merged.121 <- subset(merged.111, source %in% include2)
nrow(merged.121)

## [1] 2053

write.csv2(merged.121, "data/reanalysis_121.csv", row.names = FALSE)
metaset.121 <- run.analysis(merged.121)
write.csv2(merged.121, "data/metaset_121.csv", row.names = FALSE)
```

```
merged.221 <- subset(merged.211, source %in% include2)
nrow(merged.212)
```

```
## [1] 463
```

```
write.csv2(merged.221, "data/reanalysis_221.csv", row.names = FALSE)
metaset.221 <- run.analysis(merged.221)
write.csv2(merged.221, "data/metaset_221.csv", row.names = FALSE)
```

```
merged.321 <- subset(merged.311, source %in% include2)
nrow(merged.321)
```

```
## [1] 229
```

```
write.csv2(merged.321, "data/reanalysis_321.csv", row.names = FALSE)
metaset.321 <- run.analysis(merged.321)
write.csv2(metaset.321, "data/metaset_321.csv", row.names = FALSE)
```

## Exclusion .2.2

```
merged.122 <- subset(merged.111, source %in% include2 & expert == 1)
nrow(merged.122)
```

```
## [1] 700
```

```
write.csv2(merged.122, "data/reanalysis_122.csv", row.names = FALSE)
metaset.122 <- run.analysis(merged.122)
write.csv2(merged.122, "data/metaset_122.csv", row.names = FALSE)
```

```
merged.222 <- subset(merged.211, source %in% include2 & expert == 1)
nrow(merged.222)
```

```
## [1] 386
```

```
write.csv2(merged.222, "data/reanalysis_222.csv", row.names = FALSE)
metaset.222 <- run.analysis(merged.222)
write.csv2(merged.222, "data/metaset_222.csv", row.names = FALSE)
```

```
merged.322 <- subset(merged.311, source %in% include2 & expert == 1)
nrow(merged.322)
```

```
## [1] 229
```

```
write.csv2(merged.322, "data/reanalysis_322.csv", row.names = FALSE)
metaset.322 <- run.analysis(merged.322)
write.csv2(metaset.322, "data/metaset_322.csv", row.names = FALSE)
```

## Exclusion .3.1

```
merged.131 <- subset(merged.111, source %in% include3)
nrow(merged.131)
```

```
## [1] 1852
```

```
write.csv2(merged.131, "data/reanalysis_131.csv", row.names = FALSE)
metaset.131 <- run.analysis(merged.131)
write.csv2(metaset.131, "data/metaset_131.csv", row.names = FALSE)
```

```
merged.231 <- subset(merged.211, source %in% include3)
nrow(merged.231)
```

```
## [1] 549
```

```
write.csv2(merged.231, "data/reanalysis_231.csv", row.names = FALSE)
metaset.231 <- run.analysis(merged.231)
write.csv2(metaset.231, "data/metaset_231.csv", row.names = FALSE)
```

```
merged.331 <- subset(merged.311, source %in% include3 & expert == 1)
nrow(merged.331)
```

```
## [1] 229
```

```
write.csv2(merged.331, "data/reanalysis_331.csv", row.names = FALSE)
metaset.331 <- run.analysis(merged.331)
write.csv2(metaset.331, "data/metaset_331.csv", row.names = FALSE)
```