

005__analyses__supplemental.R

kleinri

2019-06-19

```
# Many Labs 4 Analysis Script - Supplement
# Coder: Rick Klein raklein22@gmail.com
# OSF: https://osf.io/8ccnw/
```

```
#ANALYSIS SCRIPT - Supplement
# This script contains _most_ supplemental analyses. At the moment, some
# remain in the main script, and this only includes analyses performed after
# the dissertation was accepted.
```

```
# Open the .rproj file in R Studio to avoid setting the working directory.
# Otherwise, call setwd() with a path to the /ml4/ folder
# All file paths are relative from the working directory.
```

```
library(metafor)
```

```
## Loading required package: Matrix
```

```
## Loading 'metafor' package (version 2.1-0). For an overview
## and introduction to the package please type: help(metafor).
```

```
library(metaSEM)
```

```
## Loading required package: OpenMx
```

```
## To take full advantage of multiple cores, use:
```

```
##   mxOption(key='Number of Threads', value=parallel::detectCores()) #now
```

```
##   Sys.setenv(OMP_NUM_THREADS=parallel::detectCores()) #before library(OpenMx)
```

```
##
```

```
## Attaching package: 'OpenMx'
```

```
## The following objects are masked from 'package:Matrix':
```

```
##
```

```
##   %&%, expm
```

```
## "SLSQP" is set as the default optimizer in OpenMx.
```

```
## mxOption(NULL, "Gradient algorithm") is set at "central".
```

```
## mxOption(NULL, "Optimality tolerance") is set at "6.3e-14".
```

```
## mxOption(NULL, "Gradient iterations") is set at "2".
```

```
library(haven)
```

```
library(psych)
```

```
##
```

```
## Attaching package: 'psych'
```

```
## The following object is masked from 'package:OpenMx':
```

```
##
```

```
##   tr
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(effsize)
```

```
##
## Attaching package: 'effsize'
## The following object is masked from 'package:psych':
##
##   cohen.d
```

```
library(GPArotation)
```

```
library(tidyverse)
```

```
## Warning: replacing previous import 'dplyr::vars' by 'rlang::vars' when
## loading 'dbplyr'
```

```
## Registered S3 methods overwritten by 'ggplot2':
```

```
##   method      from
##   [.quosures   rlang
##   c.quosures   rlang
##   print.quosures rlang
```

```
## -- Attaching packages ----- tid
```

```
## v ggplot2 3.1.1      v purrr  0.3.2
## v tibble  2.1.3      v stringr 1.4.0
## v tidyr   0.8.3.9000 v forcats 0.4.0
## v readr   1.3.1
```

```
## -- Conflicts ----- tidy
```

```
## x ggplot2::%+%( ) masks psych::%+%( )
## x ggplot2::alpha() masks psych::alpha()
## x tidyr::expand() masks Matrix::expand()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()
## x tidyr::pack() masks Matrix::pack()
## x purrr::rerun() masks metaSEM::rerun()
## x tidyr::unpack() masks Matrix::unpack()
```

```
#####
#
# First suppl analysis: repeat the primary meta-analyses for pro- and anti-
# ratings separately (as opposed to creating a difference score)
#
#####
# I'm re-using the primary analysis script and just editing it. Admittedly, this
```

```

# is not the most efficient code, but it should work.

# Implementing pro-only first:

#read in deidentified aggregate dataset
merged <- readRDS("./data/public/merged_deidentified.rds")

###ANALYSIS 0: no exclusions###

#Function to generate required stats for meta-analysis. No exclusions.
analysis_pro0 <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition == 'tv'])
  n_ms <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition == 'ms'])
  sd_tv <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv'], na.rm = TRUE) #sd
  sd_ms <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms'], na.rm = TRUE) #sd
  mean_tv <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv'], na.rm = TRUE)
  mean_ms <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms'], na.rm = TRUE)
  expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is expert or not
  d_diff <- (mean_ms - mean_tv)/ sqrt((sd_ms^2+sd_tv^2)/2) #computes Cohen's D effect size
  nhst <- t.test(data$proauth_avg~data$ms_condition, subset = data$source==sitesource)
  t <- nhst$statistic
  df <- nhst$parameter
  p.value <- nhst$p.value
  result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff)
  return(result)
}

#above function is run for each site identifier
riverside_results <- analysis_pro0(merged, "riverside")
azusa_results <- analysis_pro0(merged, "azusa")
cnj_results <- analysis_pro0(merged, "cnj")
illinois_results <- analysis_pro0(merged, "illinois")
ithaca_results <- analysis_pro0(merged, "ithaca")
kansas_inhouse_results <- analysis_pro0(merged, "kansas_inhouse")
occid_results <- analysis_pro0(merged, "occid")
pace_expert_results <- analysis_pro0(merged, "pace_expert")
sou_inhouse_results <- analysis_pro0(merged, "sou_inhouse")
ufl_results <- analysis_pro0(merged, "ufl")
upenn_results <- analysis_pro0(merged, "upenn")
uwmadison_expert_results <- analysis_pro0(merged, "uwmadison_expert")
uwmadison_inhouse_results <- analysis_pro0(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_pro0(merged, "wesleyan_inhouse")
wpi_results <- analysis_pro0(merged, "wpi")
kansas_expert_results <- analysis_pro0(merged, "kansas_expert")
plu_results <- analysis_pro0(merged, "plu")
ashland_results <- analysis_pro0(merged, "ashland")
vcu_results <- analysis_pro0(merged, "vcu")
byui_results <- analysis_pro0(merged, "byui")
pace_inhouse_results <- analysis_pro0(merged, "pace_inhouse")

#merges results from above into a single data frame

```

```

combinedresults_pro0 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,
  upenn_results,
  uwmadison_expert_results,
  uwmadison_inhouse_results,
  vcu_results,
  wesleyan_inhouse_results,
  wpi_results,
  byui_results,
  pace_inhouse_results
)

#Computing SE and sampling variance with metafor package.
# yi (the standardized mean difference effect size) and vi (the sampling variance) to be used in meta-a
# n1i numeric number of participants in the intervention group
# m1i numeric mean number of days off work/school in the intervention group
# sd1i numeric standard deviation of the number of days off work/school in the intervention group
# n2i numeric number of participants in the control/comparison group
# m2i numeric mean number of days off work/school in the control/comparison group
# sd2i numeric standard deviation of the number of days off work/school in the control/comparison group
#Appends yi and vi to the data object.
combinedresults_pro0 <- escalc(n1i = n_ms, n2i = n_tv, m1i = mean_ms, m2i = mean_tv,
                             sd1i = sd_ms, sd2i = sd_tv, data = combinedresults_pro0, measure = "SMD",
                             append = TRUE)

#saves .csv file
write.csv(combinedresults_pro0, "../data/public/combinedresults_pro0.csv", row.names = FALSE)

###ANALYSIS 1: Exclusion set 1###
#1. Wrote something for both writing prompts
#2. Completed all six items evaluating the essay authors)

#Function to generate required stats for meta-analysis.
analysis_pro1 <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition
n_ms <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition
sd_tv <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete
sd_ms <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete
mean_tv <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete

```

```

mean_ms <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete == 0)])
expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is expert or not
d_diff <- (mean_ms - mean_tv) / sqrt((sd_ms^2+sd_tv^2)/2) #computes Cohen's D effect size
nhst <- t.test(data$proauth_avg~data$ms_condition, subset = data$source==sitesource & (data$msincomplete == 0))
t <- nhst$statistic
df <- nhst$parameter
p.value <- nhst$p.value
result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff)
return(result)
}

```

```

#above function is run for each site identifier
riverside_results <- analysis_pro1(merged, "riverside")
azusa_results <- analysis_pro1(merged, "azusa")
cnj_results <- analysis_pro1(merged, "cnj")
illinois_results <- analysis_pro1(merged, "illinois")
ithaca_results <- analysis_pro1(merged, "ithaca")
kansas_inhouse_results <- analysis_pro1(merged, "kansas_inhouse")
occid_results <- analysis_pro1(merged, "occid")
pace_expert_results <- analysis_pro1(merged, "pace_expert")
sou_inhouse_results <- analysis_pro1(merged, "sou_inhouse")
ufl_results <- analysis_pro1(merged, "ufl")
upenn_results <- analysis_pro1(merged, "upenn")
uwmadison_expert_results <- analysis_pro1(merged, "uwmadison_expert")
uwmadison_inhouse_results <- analysis_pro1(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_pro1(merged, "wesleyan_inhouse")
wpi_results <- analysis_pro1(merged, "wpi")
kansas_expert_results <- analysis_pro1(merged, "kansas_expert")
plu_results <- analysis_pro1(merged, "plu")
ashland_results <- analysis_pro1(merged, "ashland")
vcu_results <- analysis_pro1(merged, "vcu")
byui_results <- analysis_pro1(merged, "byui")
pace_inhouse_results <- analysis_pro1(merged, "pace_inhouse")

```

```

#merges results from above into a single data frame

```

```

combinedresults_pro1 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,
  upenn_results,
  uwmadison_expert_results,
  uwmadison_inhouse_results,
  vcu_results,

```

```

wesleyan_inhouse_results,
wpi_results,
byui_results,
pace_inhouse_results
)

# This uses the metafor package to compute yi (the standardized mean difference effect size) and vi (th
# Appends this to the data object.
combinedresults_pro1 <- escalc(n1i = n_ms, n2i = n_tv, m1i = mean_ms, m2i = mean_tv,
                             sd1i = sd_ms, sd2i = sd_tv, data = combinedresults_pro1, measure = "SMD",
                             append = TRUE)

#saves .csv file
write.csv(combinedresults_pro1, "./data/public/combinedresults_pro1.csv", row.names = FALSE)

###ANALYSIS 2: Exclusion set 2###
#1. Wrote something for both writing prompts
#2. Completed all six items evaluating the essay authors
#3. Identify as White (race == 1)
#4. Born in USA (countryofbirth == 1)

#Function to generate required stats for meta-analysis.
analysis_pro2 <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition
  n_ms <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition
  sd_tv <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete
  sd_ms <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete
  mean_tv <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomp
  mean_ms <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomp
  expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is expert o
  d_diff <- (mean_ms - mean_tv)/ sqrt((sd_ms^2+sd_tv^2)/2) #computes Cohen's D effect size
  nhst <- t.test(data$proauth_avg~data$ms_condition, subset = data$source==sitesource & (data$msincompl
  t <- nhst$statistic
  df <- nhst$parameter
  p.value <- nhst$p.value
  result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff
  return(result)
}

#in-house sites don't necessarily have the data necessary to implement these exclusions
#Below, analysis1 (basic exclusions) is run for in-house, while analysis 2 is run for expert versions

#expert sites
riverside_results <- analysis_pro2(merged, "riverside")
cnj_results <- analysis_pro2(merged, "cnj")
occid_results <- analysis_pro2(merged, "occid")
pace_expert_results <- analysis_pro2(merged, "pace_expert")
uwmadison_expert_results <- analysis_pro2(merged, "uwmadison_expert")
kansas_expert_results <- analysis_pro2(merged, "kansas_expert")
ashland_results <- analysis_pro2(merged, "ashland")
vcu_results <- analysis_pro2(merged, "vcu")

```

```

byui_results <- analysis_pro2(merged, "byui")

#inhouse sites
azusa_results <- analysis_pro1(merged, "azusa")
illinois_results <- analysis_pro1(merged, "illinois")
ithaca_results <- analysis_pro1(merged, "ithaca")
kansas_inhouse_results <- analysis_pro1(merged, "kansas_inhouse")
sou_inhouse_results <- analysis_pro1(merged, "sou_inhouse")
ufl_results <- analysis_pro1(merged, "ufl")
upenn_results <- analysis_pro1(merged, "upenn")
uwmadison_inhouse_results <- analysis_pro1(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_pro1(merged, "wesleyan_inhouse")
wpi_results <- analysis_pro1(merged, "wpi")
plu_results <- analysis_pro1(merged, "plu")
pace_inhouse_results <- analysis_pro1(merged, "pace_inhouse")

#merges results from above into a single data frame
combinedresults_pro2 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,
  upenn_results,
  uwmadison_expert_results,
  uwmadison_inhouse_results,
  vcu_results,
  wesleyan_inhouse_results,
  wpi_results,
  byui_results,
  pace_inhouse_results
)

# This uses the metafor package to compute yi (the standardized mean difference effect size) and vi (th
# Appends this to the data object.
combinedresults_pro2 <- escalc(n1i = n_ms, n2i = n_tv, m1i = mean_ms, m2i = mean_tv,
                             sd1i = sd_ms, sd2i = sd_tv, data = combinedresults_pro2, measure = "SMD",
                             append = TRUE)

# saves .csv file
write.csv(combinedresults_pro2, "../data/public/combinedresults_pro2.csv", row.names = FALSE)

###ANALYSIS 3: Exclusion set 3###
# 1. Wrote something for both writing prompts
# 2. Completed all six items evaluating the essay authors

```



```

# 3. Identify as White
# 4. Born in USA
# 5. Score a 7 or higher on the American Identity item

# Function to generate required stats for meta-analysis.
analysis_pro3 <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition == 'tv'])
  n_ms <- length(data$proauth_avg[!is.na(data$proauth_avg) & data$source==sitesource & data$ms_condition == 'ms'])
  sd_tv <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete == 0)])
  sd_ms <- sd(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete == 0)])
  mean_tv <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete == 0)])
  mean_ms <- mean(data$proauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete == 0)])
  expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is expert or not
  d_diff <- (mean_ms - mean_tv) / sqrt((sd_ms^2 + sd_tv^2) / 2) #computes Cohen's D effect size
  nhst <- t.test(data$proauth_avg ~ data$ms_condition, subset = data$source==sitesource & (data$msincomplete == 0))
  t <- nhst$statistic
  df <- nhst$parameter
  p.value <- nhst$p.value
  result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff)
  return(result)
}

# in-house sites don't necessarily have the data necessary to implement these exclusions
# Below, analysis1 (basic exclusions) is run for in-house, while analysis 3 is run for expert versions

# expert sites
riverside_results <- analysis_pro3(merged, "riverside")
cnj_results <- analysis_pro3(merged, "cnj")
occid_results <- analysis_pro3(merged, "occid")
pace_expert_results <- analysis_pro3(merged, "pace_expert")
uwmadison_expert_results <- analysis_pro3(merged, "uwmadison_expert")
kansas_expert_results <- analysis_pro3(merged, "kansas_expert")
ashland_results <- analysis_pro3(merged, "ashland")
vcu_results <- analysis_pro3(merged, "vcu")
byui_results <- analysis_pro3(merged, "byui")

# inhouse sites
azusa_results <- analysis_pro1(merged, "azusa")
illinois_results <- analysis_pro1(merged, "illinois")
ithaca_results <- analysis_pro1(merged, "ithaca")
kansas_inhouse_results <- analysis_pro1(merged, "kansas_inhouse")
sou_inhouse_results <- analysis_pro1(merged, "sou_inhouse")
ufl_results <- analysis_pro1(merged, "ufl")
upenn_results <- analysis_pro1(merged, "upenn")
uwmadison_inhouse_results <- analysis_pro1(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_pro1(merged, "wesleyan_inhouse")
wpi_results <- analysis_pro1(merged, "wpi")
plu_results <- analysis_pro1(merged, "plu")
pace_inhouse_results <- analysis_pro1(merged, "pace_inhouse")

# merges results from above into a single data frame

```



```

combinedresults_pro3 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,
  upenn_results,
  uwmadison_expert_results,
  uwmadison_inhouse_results,
  vcu_results,
  wesleyan_inhouse_results,
  wpi_results,
  byui_results,
  pace_inhouse_results
)

# This uses the metafor package to compute yi (the standardized mean difference effect size) and vi (th
# Appends this to the data object.
combinedresults_pro3 <- escalc(n1i = n_ms, n2i = n_tv, m1i = mean_ms, m2i = mean_tv,
                             sd1i = sd_ms, sd2i = sd_tv, data = combinedresults_pro3, measure = "SMD",
                             append = TRUE)

# saves .csv file
write.csv(combinedresults_pro3, "./data/public/combinedresults_pro3.csv", row.names = FALSE)

# reads in csv files from above, just to confirm we can start with those files
combinedresults_pro0 <- read.csv("./data/public/combinedresults_pro0.csv")
combinedresults_pro1 <- read.csv("./data/public/combinedresults_pro1.csv")
combinedresults_pro2 <- read.csv("./data/public/combinedresults_pro2.csv")
combinedresults_pro3 <- read.csv("./data/public/combinedresults_pro3.csv")

# analyses repeated for each set of exclusion criteria
# three-level random-effects meta-analysis in MetaSEM
summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_pro0))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_pro0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error   lbound   ubound z value Pr(>|z|)
## Intercept -2.0707e-02  4.3825e-02 -1.0660e-01  6.5189e-02 -0.4725   0.6366
## Tau2_2     1.0000e-10  6.3654e-02 -1.2476e-01  1.2476e-01  0.0000   1.0000
## Tau2_3     1.0000e-10  6.3325e-02 -1.2412e-01  1.2412e-01  0.0000   1.0000
##

```

```
## Q statistic on the homogeneity of effect sizes: 20.57539
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4224922
##
## Heterogeneity indices (based on the estimated Tau2):
##               Estimate
## I2_2 (Typical v: Q statistic)      0
## I2_3 (Typical v: Q statistic)      0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: -6.073144
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_pro1))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_pro1)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -2.3136e-02  4.4176e-02 -1.0972e-01  6.3447e-02 -0.5237   0.6005
## Tau2_2     1.0000e-10  6.0256e-02 -1.1810e-01  1.1810e-01  0.0000   1.0000
## Tau2_3     1.0000e-10  6.0002e-02 -1.1760e-01  1.1760e-01  0.0000   1.0000
##
## Q statistic on the homogeneity of effect sizes: 19.65102
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4799414
##
## Heterogeneity indices (based on the estimated Tau2):
##               Estimate
## I2_2 (Typical v: Q statistic)      0
## I2_3 (Typical v: Q statistic)      0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: -6.847261
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_pro2))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_pro2)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
```

```
##           Estimate   Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept  1.1563e-02  4.8703e-02 -8.3893e-02  1.0702e-01  0.2374  0.8123
## Tau2_2     1.0001e-10  4.1460e-02 -8.1259e-02  8.1259e-02  0.0000  1.0000
## Tau2_3     1.0000e-10  4.1069e-02 -8.0494e-02  8.0494e-02  0.0000  1.0000
##
## Q statistic on the homogeneity of effect sizes: 19.873
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4658991
##
## Heterogeneity indices (based on the estimated Tau2):
##           Estimate
## I2_2 (Typical v: Q statistic)      0
## I2_3 (Typical v: Q statistic)      0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: -1.748979
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)
```

```
summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_pro3))
```

```
##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_pro3)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept  9.7805e-03  5.3349e-02 -9.4783e-02  1.1434e-01  0.1833  0.8545
## Tau2_2     1.0019e-10  3.5920e-02 -7.0402e-02  7.0402e-02  0.0000  1.0000
## Tau2_3     1.0000e-10  3.4538e-02 -6.7694e-02  6.7694e-02  0.0000  1.0000
##
## Q statistic on the homogeneity of effect sizes: 21.76236
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.3535435
##
## Heterogeneity indices (based on the estimated Tau2):
##           Estimate
## I2_2 (Typical v: Q statistic)      0
## I2_3 (Typical v: Q statistic)      0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: 5.006565
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)
```

#Notes: I² for level 2 indicates the percent of total variance explained by effects within sites, and I

forest plots for each

```

# ### All forest plots now outdated in favor of metaviz.R
# ### All forest plots now outdated in favor of metaviz.R
# ### All forest plots now outdated in favor of metaviz.R
data <- combinedresults1
# same forst plot, but using rma so it plots the aggregate
dev.off()
png("./output/comb1.randomeffects.png", type='cairo')
par(mar=c(4,4,1,4)) #decreasing margins
forest(rma(yi= data$yi, vi=data$vi, slab=data$sitesource))
par(cex=1, font=2) #bold font
text(-3.3, 20.5, "Location", pos=4) #adds location label using x, y coord
text(3.8, 20.5, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
dev.off()
#
data <- combinedresults2
# same forst plot, but using rma so it plots the aggregate
dev.off()
png("./output/comb2.randomeffects.png", type='cairo')
par(mar=c(4,4,1,4)) #decreasing margins
forest(rma(yi= data$yi, vi=data$vi, slab=data$sitesource))
par(cex=1, font=2) #bold font
text(-5.1, 20.5, "Location", pos=4) #adds location label using x, y coord
text(6.6, 20.5, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
dev.off()
#
data <- combinedresults3
# same forst plot, but using rma so it plots the aggregate
dev.off()
png("./output/comb3.randomeffects.png", type='cairo')
par(mar=c(4,4,1,4)) #decreasing margins
forest(rma(yi= data$yi, vi=data$vi, slab=data$sitesource))
par(cex=1, font=2) #bold font
text(-6, 20.5, "Location", pos=4) #adds location label using x, y coord
text(6.5, 20.5, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
dev.off()

# a covariate of study version (in-house or expert-designed) is added to create a three-level mixed-eff
# note the openMX status, sometimes indicates a potential problem
summary(mixed_pro0 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro0))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error   lbound   ubound z value Pr(>|z|)
## Intercept -1.9979e-02  5.4704e-02 -1.2720e-01  8.7238e-02 -0.3652   0.7149
## Slope_1    -2.0379e-03  9.9973e-02 -1.9798e-01  1.9390e-01 -0.0204   0.9837
## Tau2_2      1.0000e-10  7.0050e-02 -1.3730e-01  1.3730e-01  0.0000   1.0000
## Tau2_3      1.0002e-10  6.8883e-02 -1.3501e-01  1.3501e-01  0.0000   1.0000
##
## Q statistic on the homogeneity of effect sizes: 20.57539
## Degrees of freedom of the Q statistic: 20

```

```
## P value of the Q statistic: 0.4224922
##
## Explained variances (R2):
##           Level 2 Level 3
## Tau2 (no predictor)      1e-10      0
## Tau2 (with predictors)   1e-10      0
## R2                      0e+00      0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4
## Degrees of freedom: 17
## -2 log likelihood: -6.07367
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( mixed_pro1 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro1))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro1)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std.Error      lbound      ubound z value Pr(>|z|)
## Intercept -2.4601e-02 5.5309e-02 -1.3300e-01 8.3803e-02 -0.4448 0.6565
## Slope_1    4.0784e-03 1.0087e-01 -1.9363e-01 2.0178e-01 0.0404 0.9677
## Tau2_2     1.0000e-10 6.9559e-02 -1.3633e-01 1.3633e-01 0.0000 1.0000
## Tau2_3     1.0012e-10 6.8101e-02 -1.3348e-01 1.3348e-01 0.0000 1.0000
##
## Q statistic on the homogeneity of effect sizes: 19.65102
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4799414
##
## Explained variances (R2):
##           Level 2 Level 3
## Tau2 (no predictor)      1e-10      0
## Tau2 (with predictors)   1e-10      0
## R2                      0e+00      0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4
## Degrees of freedom: 17
## -2 log likelihood: -6.849359
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( mixed_pro2 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro2))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro2)
##
## 95% confidence intervals: z statistic approximation
```

```

## Coefficients:
##           Estimate   Std.Error    lbound      ubound z value Pr(>|z|)
## Intercept -2.4601e-02  5.5875e-02 -1.3411e-01  8.4913e-02 -0.4403  0.6597
## Slope_1    1.4282e-01  1.1665e-01 -8.5818e-02  3.7146e-01  1.2243  0.2208
## Tau2_2     1.0000e-10  6.6387e-02 -1.3012e-01  1.3012e-01  0.0000  1.0000
## Tau2_3     1.0000e-10  6.2130e-02 -1.2177e-01  1.2177e-01  0.0000  1.0000
##
## Q statistic on the homogeneity of effect sizes: 19.873
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4658991
##
## Explained variances (R2):
##                Level 2 Level 3
## Tau2 (no predictor)  1.0001e-10    0
## Tau2 (with predictors) 1.0000e-10    0
## R2                  1.0853e-04    0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4
## Degrees of freedom: 17
## -2 log likelihood: -3.56347
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( mixed_pro3 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro3))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro3)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error    lbound      ubound z value Pr(>|z|)
## Intercept -2.4601e-02  5.6416e-02 -1.3517e-01  8.5973e-02 -0.4361  0.6628
## Slope_1    2.0375e-01  1.3635e-01 -6.3483e-02  4.7098e-01  1.4944  0.1351
## Tau2_2     1.0000e-10  5.2374e-02 -1.0265e-01  1.0265e-01  0.0000  1.0000
## Tau2_3     1.0000e-10  4.9167e-02 -9.6366e-02  9.6366e-02  0.0000  1.0000
##
## Q statistic on the homogeneity of effect sizes: 21.76236
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.3535435
##
## Explained variances (R2):
##                Level 2 Level 3
## Tau2 (no predictor)  1.0019e-10    0
## Tau2 (with predictors) 1.0000e-10    0
## R2                  1.9361e-03    0
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4
## Degrees of freedom: 17
## -2 log likelihood: 2.545569
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.

```

```
## Other values may indicate problems.)
# Notes: The R2 for the version predictor will be reported for both level 2 and level 3, although in the
# constraining the variance to test if it significantly worsens the model
summary( fixed_pro0 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro0, RE2.constraints = 0, RE3.constraints = 0)

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro0,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std. Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.0199794  0.0530369 -0.1239299  0.0839710 -0.3767   0.7064
## Slope_1    -0.0020379  0.0887914 -0.1760660  0.1719901 -0.0230   0.9817
##
## Q statistic on the homogeneity of effect sizes: 20.57539
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4224922
##
## Explained variances (R2):
##                Level 2 Level 3
## Tau2 (no predictor)    1e-10     0
## Tau2 (with predictors)    NA      NA
## R2                     NA      NA
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: -6.07367
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( fixed_pro1 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro1, RE2.constraints = 0, RE3.constraints = 0)

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro1,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std. Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.0246006  0.0533522 -0.1291690  0.0799678 -0.4611   0.6447
## Slope_1    0.0040784  0.0890345 -0.1704260  0.1785827  0.0458   0.9635
##
## Q statistic on the homogeneity of effect sizes: 19.65102
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4799414
##
## Explained variances (R2):
##                Level 2 Level 3
```



```

## Tau2 (no predictor)      1e-10      0
## Tau2 (with predictors)   NA        NA
## R2                      NA        NA
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: -6.849359
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( fixed_pro2 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro2, RE2.constraints = 0, RE3.constraints = 0)

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro2,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.024601  0.053352 -0.129169  0.079968 -0.4611  0.6447
## Slope_1    0.142821  0.106026 -0.064987  0.350628  1.3470  0.1780
##
## Q statistic on the homogeneity of effect sizes: 19.873
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.4658991
##
## Explained variances (R2):
##                Level 2 Level 3
## Tau2 (no predictor)  1.0001e-10      0
## Tau2 (with predictors)    NA      NA
## R2                    NA      NA
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: -3.56347
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( fixed_pro3 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_pro3, RE2.constraints = 0, RE3.constraints = 0)

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_pro3,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.024601  0.053352 -0.129169  0.079968 -0.4611  0.6447
## Slope_1    0.203749  0.129880 -0.050810  0.458309  1.5688  0.1167

```

```
##
## Q statistic on the homogeneity of effect sizes: 21.76236
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.3535435
##
## Explained variances (R2):
##               Level 2 Level 3
## Tau2 (no predictor) 1.0019e-10 0
## Tau2 (with predictors) NA NA
## R2 NA NA
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: 2.545569
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)
# compare if there is a significant difference in model fit, chi square difference test
anova(mixed_pro0, fixed_pro0)

##               base               comparison ep minus2LL df AIC
## 1 Meta analysis with ML <NA> 4 -6.07367 17 -40.07367
## 2 Meta analysis with ML Meta analysis with ML 2 -6.07367 19 -44.07367
##               diffLL diffdf p
## 1 NA NA NA
## 2 -2.085844e-09 2 1
anova(mixed_pro1, fixed_pro1)

##               base               comparison ep minus2LL df AIC
## 1 Meta analysis with ML <NA> 4 -6.849359 17 -40.84936
## 2 Meta analysis with ML Meta analysis with ML 2 -6.849359 19 -44.84936
##               diffLL diffdf p
## 1 NA NA NA
## 2 -1.118946e-08 2 1
anova(mixed_pro2, fixed_pro2)

##               base               comparison ep minus2LL df AIC
## 1 Meta analysis with ML <NA> 4 -3.56347 17 -37.56347
## 2 Meta analysis with ML Meta analysis with ML 2 -3.56347 19 -41.56347
##               diffLL diffdf p
## 1 NA NA NA
## 2 -1.665738e-08 2 1
anova(mixed_pro3, fixed_pro3)

##               base               comparison ep minus2LL df AIC
## 1 Meta analysis with ML <NA> 4 2.545569 17 -31.45443
## 2 Meta analysis with ML Meta analysis with ML 2 2.545569 19 -35.45443
##               diffLL diffdf p
## 1 NA NA NA
## 2 -1.453634e-08 2 1
```

```

# Repeating analyses of "expert" sites in the aggregate, ignoring site dependence.
# This is a simple alternative and useful for most stringent exclusion criteria which drastically reduce
# read in .rds data
data <- readRDS("./data/public/merged_deidentified.rds")
# selecting only expert labs
data <- subset(data, expert==1)

###ANALYSIS 0: no exclusions###
# t.test and descriptive statistics per condition from psych package
t.test(data$proauth_avg~data$ms_condition)

##
## Welch Two Sample t-test
##
## data: data$proauth_avg by data$ms_condition
## t = -0.4294, df = 798.85, p-value = 0.6677
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.2074399 0.1329736
## sample estimates:
## mean in group ms mean in group tv
## 6.701878 6.739111

describeBy(data$proauth_avg, group = data$ms_condition)

##
## Descriptive statistics by group
## group: ms
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 426 6.7 1.3 7 6.74 1.48 2.67 9 6.33 -0.35 -0.27
## se
## X1 0.06
## -----
## group: tv
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 375 6.74 1.16 7 6.78 0.99 2.67 9 6.33 -0.45 0.2
## se
## X1 0.06

effsize::cohen.d(data$proauth_avg~data$ms_condition,pooled=TRUE,paired=FALSE,
na.rm=TRUE, hedges.correction=TRUE,
conf.level=0.95)

##
## Hedges's g
##
## g estimate: -0.03015799 (negligible)
## 95 percent confidence interval:
## lower upper
## -0.1691619 0.1088459

###ANALYSIS 1: Exclusion set 1###
# 1. Wrote something for both writing prompts
data <- subset(data, (data$msincomplete == 0 | is.na(data$msincomplete)))
# 2. Completed all six items evaluating the essay authors)
data <- subset(data, (!is.na(data$prou3) & !is.na(data$prou4) & !is.na(data$prou5) & !is.na(data$ant

```

```

# t.test and descriptive statistics per condition from psych package
t.test(data$proauth_avg~data$ms_condition)

##
## Welch Two Sample t-test
##
## data: data$proauth_avg by data$ms_condition
## t = -0.40404, df = 796.95, p-value = 0.6863
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.2056349 0.1354316
## sample estimates:
## mean in group ms mean in group tv
## 6.704009 6.739111

describeBy(data$proauth_avg, group = data$ms_condition)

##
## Descriptive statistics by group
## group: ms
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 424 6.7 1.3 7 6.75 1.48 2.67 9 6.33 -0.36 -0.27
## se
## X1 0.06
## -----
## group: tv
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 375 6.74 1.16 7 6.78 0.99 2.67 9 6.33 -0.45 0.2
## se
## X1 0.06

effsize::cohen.d(data$proauth_avg~data$ms_condition,pooled=TRUE,paired=FALSE,
na.rm=TRUE, hedges.correction=TRUE,
conf.level=0.95)

##
## Hedges's g
##
## g estimate: -0.02841381 (negligible)
## 95 percent confidence interval:
## lower upper
## -0.1675707 0.1107431

###ANALYSIS 2: Exclusion set 2###
# 1. Wrote something for both writing prompts
# 2. Completed all six items evaluating the essay authors
# 3. Identify as White (race == 1)
data <- subset(data, data$race == 1)
# 4. Born in USA (countryofbirth == 1)
data <- subset(data, data$countryofbirth == 1)
# t.test and descriptive statistics per condition from psych package
t.test(data$proauth_avg~data$ms_condition)

##
## Welch Two Sample t-test
##

```

```
## data: data$proauth_avg by data$ms_condition
## t = 1.2548, df = 440.54, p-value = 0.2102
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.07916149 0.35874450
## sample estimates:
## mean in group ms mean in group tv
## 6.872312 6.732520
```

```
describeBy(data$proauth_avg, group = data$ms_condition)
```

```
##
## Descriptive statistics by group
## group: ms
## vars n mean sd median trimmed mad min max range skew kurtosis se
## X1 1 248 6.87 1.2 7 6.9 1.48 3 9 6 -0.3 -0.21 0.08
## -----
## group: tv
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 205 6.73 1.16 7 6.79 0.99 2.67 9 6.33 -0.57 0.59
## se
## X1 0.08
```

```
effsize::cohen.d(data$proauth_avg~data$ms_condition,pooled=TRUE,paired=FALSE,
na.rm=TRUE, hedges.correction=TRUE,
conf.level=0.95) #this is incorrectly indicating a negative value, I'm not sure why bu
```

```
##
## Hedges's g
##
## g estimate: 0.1178321 (negligible)
## 95 percent confidence interval:
## lower upper
## -0.06783481 0.30349910
```

```
###ANALYSIS 3: Exclusion set 3###
# 1. Wrote something for both writing prompts
# 2. Completed all six items evaluating the essay authors
# 3. Identify as White
# 4. Born in USA
# 5. Score a 7 or higher on the American Identity item
data <- subset(data, data$americanid >= 7)
# t.test and descriptive statistics per condition from psych package
t.test(data$proauth_avg~data$ms_condition)
```

```
##
## Welch Two Sample t-test
##
## data: data$proauth_avg by data$ms_condition
## t = 1.5325, df = 259.3, p-value = 0.1266
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.0610248 0.4893910
## sample estimates:
## mean in group ms mean in group tv
## 7.038031 6.823848
```

```
describeBy(data$proauth_avg, group = data$ms_condition)
```

```
##
## Descriptive statistics by group
## group: ms
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 149 7.04 1.14 7 7.06 1.48 4.33 9 4.67 -0.18 -0.74
## se
## X1 0.09
## -----
## group: tv
## vars n mean sd median trimmed mad min max range skew kurtosis se
## X1 1 123 6.82 1.15 7 6.85 0.99 3 9 6 -0.39 0.14 0.1
```

```
effsize::cohen.d(data$proauth_avg-data$ms_condition,pooled=TRUE,paired=FALSE,
na.rm=TRUE, hedges.correction=TRUE,
conf.level=0.95) #this is incorrectly indicating a positive value, reversing sign in t
```

```
##
## Hedges's g
##
## g estimate: 0.1863663 (negligible)
## 95 percent confidence interval:
## lower upper
## -0.05399807 0.42673065
```

```
###Conducting a small meta-analysis of only the in-house data to provide a summary of those results in
# Read in summary.csv which used basic exclusion rules, Exclusion Set 1
data <- read.csv("./data/public/combinedresults_pro1.csv")
# subset to in-house rows only
data <- subset(data, expert==0)
# conduct random effects meta-analysis
summary( meta(y = yi, v = vi, data = data))
```

```
##
## Call:
## meta(y = yi, v = vi, data = data)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
## Estimate Std.Error lbound ubound z value
## Intercept1 -2.4601e-02 5.6147e-02 -1.3465e-01 8.5446e-02 -0.4381
## Tau2_1_1 1.0000e-10 1.6110e-02 -3.1574e-02 3.1574e-02 0.0000
## Pr(>|z|)
## Intercept1 0.6613
## Tau2_1_1 1.0000
##
## Q statistic on the homogeneity of effect sizes: 9.890359
## Degrees of freedom of the Q statistic: 11
## P value of the Q statistic: 0.5402757
##
## Heterogeneity indices (based on the estimated Tau2):
## Estimate
## Intercept1: I2 (Q statistic) 0
##
```

```

## Number of studies (or clusters): 12
## Number of observed statistics: 12
## Number of estimated parameters: 2
## Degrees of freedom: 10
## -2 log likelihood: -5.85792
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

# # forest plot
# dev.off()
# par(mar=c(4,4,1,4)) #decreasing margins
# forest(x= data$yi, vi=data$vi, slab=data$location)
# par(cex=1, font=2)#bold font
# text(-3.3, 13, "Location", pos=4) #adds location label using x, y coord
# text(3.8, 13, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
#
# # same forest plot, but using rma so it plots the aggregate
# dev.off()
# png("../output/inhousemeta.png", type='cairo')
# par(mar=c(4,4,1,4)) #decreasing margins
# forest(rma(yi= data$yi, vi=data$vi, slab=data$location))
# par(cex=1, font=2) #bold font
# text(-3.3, 13, "Location", pos=4) #adds location label using x, y coord
# text(3.8, 13, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
# dev.off()

# sample funnel plot
# funnel(rma(yi= data$yi, vi=data$vi, slab=data$location))

# Focused analysis of sites with "expert" or "a lot of knowledge about TMT" leads
# Still using exclusion set 1
# Read data
data <- merged
# Applying exclusion criteria 1
# 1. Wrote something for both writing prompts
data <- subset(data, (data$msincomplete == 0 | is.na(data$msincomplete)))
# 2. Completed all six items evaluating the essay authors)
data <- subset(data, (!is.na(data$prou3) & !is.na(data$prou4) & !is.na(data$prou5) & !is.na(data$ant
# Selecting only the below sites:
#University of Wisconsin, Madison, WI (in-house)
#The College of New Jersey
#University of Kansas (Expert)
#University of Kansas (in-house)
#Pace University (expert)
#Virginia Commonwealth University, Richmond, VA
data <- subset(data, data$source=="uwmadison_inhouse" | data$source=="cnj" | data$source=="kansas_expert
# Applying the same levels fix as earlier, only because it caused problems in
# cohen.d() below. May not be necessary anymore.
data$ms_condition <- factor(data$ms_condition, levels = c("ms", "tv"))
# Analyses using that subset
t.test(data$proauth_avg~data$ms_condition)

##
## Welch Two Sample t-test
##

```



```

## data: data$proauth_avg by data$ms_condition
## t = 0.49098, df = 524.4, p-value = 0.6236
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.2152873 0.3587548
## sample estimates:
## mean in group ms mean in group tv
## 6.327423 6.255689

describeBy(data$proauth_avg, group = data$ms_condition)

##
## Descriptive statistics by group
## group: ms
## vars n mean sd median trimmed mad min max range skew kurtosis se
## X1 1 282 6.33 1.71 6.67 6.47 1.48 1 9 8 -0.79 0.26 0.1
## -----
## group: tv
## vars n mean sd median trimmed mad min max range skew kurtosis se
## X1 1 249 6.26 1.66 6.33 6.41 1.48 1 9 8 -0.81 0.32 0.1

effsize::cohen.d(data$proauth_avg~data$ms_condition,pooled=TRUE,paired=FALSE,
na.rm=TRUE, hedges.correction=TRUE,
conf.level=0.95) #this was previously incorrectly indicating a positive value? Had to

##
## Hedges's g
##
## g estimate: 0.04255275 (negligible)
## 95 percent confidence interval:
## lower upper
## -0.1282972 0.2134027

#####
#####
#####
# Analyzing anti- ratings now
#####
#####
#####

#read in deidentified aggregate dataset
merged <- readRDS("../data/public/merged_deidentified.rds")

###ANALYSIS 0: no exclusions###

#Function to generate required stats for meta-analysis. No exclusions.
analysis_anti0 <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condit

```

```

n_ms <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condit
sd_tv <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv'], na.rm = TRUE) #sd
sd_ms <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms'], na.rm = TRUE) #sd
mean_tv <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv'], na.rm = TRUE)
mean_ms <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms'], na.rm = TRUE)
expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is experto
d_diff <- (mean_ms - mean_tv) / sqrt((sd_ms^2+sd_tv^2)/2) #computes Cohen's D effect size
nhst <- t.test(data$antiauth_avg~data$ms_condition, subset = data$source==sitesource)
t <- nhst$statistic
df <- nhst$parameter
p.value <- nhst$p.value
result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff
return(result)
}

```

```

#above function is run for each site identifier
riverside_results <- analysis_anti0(merged, "riverside")
azusa_results <- analysis_anti0(merged, "azusa")
cnj_results <- analysis_anti0(merged, "cnj")
illinois_results <- analysis_anti0(merged, "illinois")
ithaca_results <- analysis_anti0(merged, "ithaca")
kansas_inhouse_results <- analysis_anti0(merged, "kansas_inhouse")
occid_results <- analysis_anti0(merged, "occid")
pace_expert_results <- analysis_anti0(merged, "pace_expert")
sou_inhouse_results <- analysis_anti0(merged, "sou_inhouse")
ufl_results <- analysis_anti0(merged, "ufl")
upenn_results <- analysis_anti0(merged, "upenn")
uwmadison_expert_results <- analysis_anti0(merged, "uwmadison_expert")
uwmadison_inhouse_results <- analysis_anti0(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_anti0(merged, "wesleyan_inhouse")
wpi_results <- analysis_anti0(merged, "wpi")
kansas_expert_results <- analysis_anti0(merged, "kansas_expert")
plu_results <- analysis_anti0(merged, "plu")
ashland_results <- analysis_anti0(merged, "ashland")
vcu_results <- analysis_anti0(merged, "vcu")
byui_results <- analysis_anti0(merged, "byui")
pace_inhouse_results <- analysis_anti0(merged, "pace_inhouse")

```

```

#merges results from above into a single data frame
combinedresults_anti0 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,

```

```

upenn_results,
uwmadison_expert_results,
uwmadison_inhouse_results,
vcu_results,
wesleyan_inhouse_results,
wpi_results,
byui_results,
pace_inhouse_results
)

#Computing SE and sampling variance with metafor package.
# yi (the standardized mean difference effect size) and vi (the sampling variance) to be used in meta-a
# n1i numeric number of participants in the intervention group
# m1i numeric mean number of days off work/school in the intervention group
# sd1i numeric standard deviation of the number of days off work/school in the intervention group
# n2i numeric number of participants in the control/comparison group
# m2i numeric mean number of days off work/school in the control/comparison group
# sd2i numeric standard deviation of the number of days off work/school in the control/comparison group
#Appends yi and vi to the data object.
combinedresults_anti0 <- escalc(n1i = n_ms, n2i = n_tv, m1i = mean_ms, m2i = mean_tv,
                               sd1i = sd_ms, sd2i = sd_tv, data = combinedresults_anti0, measure = "SMD",
                               append = TRUE)

#saves .csv file
write.csv(combinedresults_anti0, "../data/public/combinedresults_anti0.csv", row.names = FALSE)

###ANALYSIS 1: Exclusion set 1###
#1. Wrote something for both writing prompts
#2. Completed all six items evaluating the essay authors)

#Function to generate required stats for meta-analysis.
analysis_antil <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condit
  n_ms <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condit
  sd_tv <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomple
  sd_ms <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomple
  mean_tv <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomple
  mean_ms <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomple
  expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is expert o
  d_diff <- (mean_ms - mean_tv)/ sqrt((sd_ms^2+sd_tv^2)/2) #computes Cohen's D effect size
  nhst <- t.test(data$antiauth_avg~data$ms_condition, subset = data$source==sitesource & (data$msincomple
  t <- nhst$statistic
  df <- nhst$parameter
  p.value <- nhst$p.value
  result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff
  return(result)
}

#above function is run for each site identifier
riverside_results <- analysis_antil(merged, "riverside")
azusa_results <- analysis_antil(merged, "azusa")

```

```

cnj_results <- analysis_anti1(merged, "cnj")
illinois_results <- analysis_anti1(merged, "illinois")
ithaca_results <- analysis_anti1(merged, "ithaca")
kansas_inhouse_results <- analysis_anti1(merged, "kansas_inhouse")
occid_results <- analysis_anti1(merged, "occid")
pace_expert_results <- analysis_anti1(merged, "pace_expert")
sou_inhouse_results <- analysis_anti1(merged, "sou_inhouse")
ufl_results <- analysis_anti1(merged, "ufl")
upenn_results <- analysis_anti1(merged, "upenn")
uwmadison_expert_results <- analysis_anti1(merged, "uwmadison_expert")
uwmadison_inhouse_results <- analysis_anti1(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_anti1(merged, "wesleyan_inhouse")
wpi_results <- analysis_anti1(merged, "wpi")
kansas_expert_results <- analysis_anti1(merged, "kansas_expert")
plu_results <- analysis_anti1(merged, "plu")
ashland_results <- analysis_anti1(merged, "ashland")
vcu_results <- analysis_anti1(merged, "vcu")
byui_results <- analysis_anti1(merged, "byui")
pace_inhouse_results <- analysis_anti1(merged, "pace_inhouse")

```

#merges results from above into a single data frame

```

combinedresults_anti1 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,
  upenn_results,
  uwmadison_expert_results,
  uwmadison_inhouse_results,
  vcu_results,
  wesleyan_inhouse_results,
  wpi_results,
  byui_results,
  pace_inhouse_results
)

```

This uses the metafor package to compute yi (the standardized mean difference effect size) and vi (the variance of yi)
Appends this to the data object.

```

combinedresults_anti1 <- escalc(n1i = n_ms, n2i = n_tv, m1i = mean_ms, m2i = mean_tv,
                               sd1i = sd_ms, sd2i = sd_tv, data = combinedresults_anti1, measure = "SMD",
                               append = TRUE)

```

#saves .csv file

```

write.csv(combinedresults_anti1, "../data/public/combinedresults_anti1.csv", row.names = FALSE)

```

```

####ANALYSIS 2: Exclusion set 2####
#1. Wrote something for both writing prompts
#2. Completed all six items evaluating the essay authors
#3. Identify as White (race == 1)
#4. Born in USA (countryofbirth == 1)

#Function to generate required stats for meta-analysis.
analysis_anti2 <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condit
  n_ms <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condit
  sd_tv <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete
  sd_ms <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete
  mean_tv <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete
  mean_ms <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete
  expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is expert or not
  d_diff <- (mean_ms - mean_tv)/ sqrt((sd_ms^2+sd_tv^2)/2) #computes Cohen's D effect size
  nhst <- t.test(data$antiauth_avg~data$ms_condition, subset = data$source==sitesource & (data$msincomplete
  t <- nhst$statistic
  df <- nhst$parameter
  p.value <- nhst$p.value
  result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff)
  return(result)
}

#in-house sites don't necessarily have the data necessary to implement these exclusions
#Below, analysis1 (basic exclusions) is run for in-house, while analysis 2 is run for expert versions

#expert sites
riverside_results <- analysis_anti2(merged, "riverside")
cnj_results <- analysis_anti2(merged, "cnj")
occid_results <- analysis_anti2(merged, "occid")
pace_expert_results <- analysis_anti2(merged, "pace_expert")
uwmadison_expert_results <- analysis_anti2(merged, "uwmadison_expert")
kansas_expert_results <- analysis_anti2(merged, "kansas_expert")
ashland_results <- analysis_anti2(merged, "ashland")
vcu_results <- analysis_anti2(merged, "vcu")
byui_results <- analysis_anti2(merged, "byui")

#inhouse sites
azusa_results <- analysis_anti1(merged, "azusa")
illinois_results <- analysis_anti1(merged, "illinois")
ithaca_results <- analysis_anti1(merged, "ithaca")
kansas_inhouse_results <- analysis_anti1(merged, "kansas_inhouse")
sou_inhouse_results <- analysis_anti1(merged, "sou_inhouse")
ufl_results <- analysis_anti1(merged, "ufl")
upenn_results <- analysis_anti1(merged, "upenn")
uwmadison_inhouse_results <- analysis_anti1(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_anti1(merged, "wesleyan_inhouse")
wpi_results <- analysis_anti1(merged, "wpi")
plu_results <- analysis_anti1(merged, "plu")
pace_inhouse_results <- analysis_anti1(merged, "pace_inhouse")

```

```

#merges results from above into a single data frame
combinedresults_anti2 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,
  upenn_results,
  uwmadison_expert_results,
  uwmadison_inhouse_results,
  vcu_results,
  wesleyan_inhouse_results,
  wpi_results,
  byui_results,
  pace_inhouse_results
)

# This uses the metafor package to compute yi (the standardized mean difference effect size) and vi (th
# Appends this to the data object.
combinedresults_anti2 <- escalc(nli = n_ms, n2i = n_tv, m1i = mean_ms, m2i = mean_tv,
                               sd1i = sd_ms, sd2i = sd_tv, data = combinedresults_anti2, measure = "SMD",
                               append = TRUE)

# saves .csv file
write.csv(combinedresults_anti2, "../data/public/combinedresults_anti2.csv", row.names = FALSE)

####ANALYSIS 3: Exclusion set 3###
# 1. Wrote something for both writing prompts
# 2. Completed all six items evaluating the essay authors
# 3. Identify as White
# 4. Born in USA
# 5. Score a 7 or higher on the American Identity item

# Function to generate required stats for meta-analysis.
analysis_anti3 <- function(data, sitesource)
{
  location <- merged$location[data$source==sitesource][1] #saves first row from location variable
  n_tv <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condition == 'tv'])
  n_ms <- length(data$antiauth_avg[!is.na(data$antiauth_avg) & data$source==sitesource & data$ms_condition == 'ms'])
  sd_tv <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete == 0)])
  sd_ms <- sd(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete == 0)])
  mean_tv <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'tv' & (data$msincomplete == 0)])
  mean_ms <- mean(data$antiauth_avg[data$source==sitesource & data$ms_condition == 'ms' & (data$msincomplete == 0)])
  expert <- mean(merged$expert[data$source==sitesource]) #shortcut to indicate whether site is expert or not
  d_diff <- (mean_ms - mean_tv) / sqrt((sd_ms^2 + sd_tv^2) / 2) #computes Cohen's D effect size
}

```

```

nhst <- t.test(data$antiauth_avg~data$ms_condition, subset = data$source==sitesource & (data$msincomp
t <- nhst$statistic
df <- nhst$parameter
p.value <- nhst$p.value
result <- data.frame(location, sitesource, expert, n_tv, mean_tv, sd_tv, n_ms, mean_ms, sd_ms, d_diff
return(result)
}

# in-house sites don't necessarily have the data necessary to implement these exclusions
# Below, analysis1 (basic exclusions) is run for in-house, while analysis 3 is run for expert versions

# expert sites
riverside_results <- analysis_anti3(merged, "riverside")
cnj_results <- analysis_anti3(merged, "cnj")
occid_results <- analysis_anti3(merged, "occid")
pace_expert_results <- analysis_anti3(merged, "pace_expert")
uwmadison_expert_results <- analysis_anti3(merged, "uwmadison_expert")
kansas_expert_results <- analysis_anti3(merged, "kansas_expert")
ashland_results <- analysis_anti3(merged, "ashland")
vcu_results <- analysis_anti3(merged, "vcu")
byui_results <- analysis_anti3(merged, "byui")

# inhouse sites
azusa_results <- analysis_anti1(merged, "azusa")
illinois_results <- analysis_anti1(merged, "illinois")
ithaca_results <- analysis_anti1(merged, "ithaca")
kansas_inhouse_results <- analysis_anti1(merged, "kansas_inhouse")
sou_inhouse_results <- analysis_anti1(merged, "sou_inhouse")
ufl_results <- analysis_anti1(merged, "ufl")
upenn_results <- analysis_anti1(merged, "upenn")
uwmadison_inhouse_results <- analysis_anti1(merged, "uwmadison_inhouse")
wesleyan_inhouse_results <- analysis_anti1(merged, "wesleyan_inhouse")
wpi_results <- analysis_anti1(merged, "wpi")
plu_results <- analysis_anti1(merged, "plu")
pace_inhouse_results <- analysis_anti1(merged, "pace_inhouse")

# merges results from above into a single data frame
combinedresults_anti3 <- rbind(
  ashland_results,
  azusa_results,
  cnj_results,
  illinois_results,
  ithaca_results,
  kansas_expert_results,
  kansas_inhouse_results,
  occid_results,
  pace_expert_results,
  plu_results,
  riverside_results,
  sou_inhouse_results,
  ufl_results,
  upenn_results,
  uwmadison_expert_results,

```



```

uwmadison_inhouse_results,
vcu_results,
wesleyan_inhouse_results,
wpi_results,
byui_results,
pace_inhouse_results
)

# This uses the metafor package to compute yi (the standardized mean difference effect size) and vi (th
# Appends this to the data object.
combinedresults_anti3 <- escalc(nli = n_ms, n2i = n_tv, mli = mean_ms, m2i = mean_tv,
                               sdli = sd_ms, sd2i = sd_tv, data = combinedresults_anti3, measure = "SMD",
                               append = TRUE)

# saves .csv file
write.csv(combinedresults_anti3, "../data/public/combinedresults_anti3.csv", row.names = FALSE)

# reads in csv files from above, just to confirm we can start with those files
combinedresults_anti0 <- read.csv("../data/public/combinedresults_anti0.csv")
combinedresults_anti1 <- read.csv("../data/public/combinedresults_anti1.csv")
combinedresults_anti2 <- read.csv("../data/public/combinedresults_anti2.csv")
combinedresults_anti3 <- read.csv("../data/public/combinedresults_anti3.csv")

# analyses repeated for each set of exclusion criteria
# three-level random-effects meta-analysis in MetaSEM
summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_anti0))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_anti0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error   lbound   ubound z value Pr(>|z|)
## Intercept -5.1829e-02  4.9917e-02 -1.4966e-01  4.6006e-02 -1.0383   0.2991
## Tau2_2     1.0000e-10         NA         NA         NA      NA      NA
## Tau2_3     1.2170e-02         NA         NA         NA      NA      NA
##
## Q statistic on the homogeneity of effect sizes: 25.57261
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.1803955
##
## Heterogeneity indices (based on the estimated Tau2):
##           Estimate
## I2_2 (Typical v: Q statistic)   0.000
## I2_3 (Typical v: Q statistic)   0.239
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: -1.892966
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

```

```
## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_anti1))
```

```
##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_anti1)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -5.0982e-02  4.8536e-02 -1.4611e-01  4.4146e-02 -1.0504   0.2935
## Tau2_2     1.0000e-10           NA           NA           NA      NA      NA
## Tau2_3     9.7033e-03           NA           NA           NA      NA      NA
##
## Q statistic on the homogeneity of effect sizes: 24.25077
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.2316096
##
## Heterogeneity indices (based on the estimated Tau2):
##           Estimate
## I2_2 (Typical v: Q statistic)  0.0000
## I2_3 (Typical v: Q statistic)  0.1992
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: -2.818028
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)
```

```
## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_anti2))
```

```
##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_anti2)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -5.6280e-02  5.7622e-02 -1.6922e-01  5.6657e-02 -0.9767   0.3287
## Tau2_2     1.0000e-10           NA           NA           NA      NA      NA
## Tau2_3     1.8814e-02           NA           NA           NA      NA      NA
##
## Q statistic on the homogeneity of effect sizes: 31.12158
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.05360574
##
## Heterogeneity indices (based on the estimated Tau2):
##           Estimate
## I2_2 (Typical v: Q statistic)  0.0000
## I2_3 (Typical v: Q statistic)  0.2901
##
```

```

## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: 8.525611
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
summary( meta3(y=yi, v=vi, cluster=location, data=combinedresults_anti3))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, data = combinedresults_anti3)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##              Estimate Std.Error lbound ubound z value Pr(>|z|)
## Intercept -3.2027e-02      NA      NA      NA      NA      NA
## Tau2_2     1.0000e-10      NA      NA      NA      NA      NA
## Tau2_3     1.2887e-02      NA      NA      NA      NA      NA
##
## Q statistic on the homogeneity of effect sizes: 26.74513
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.1425881
##
## Heterogeneity indices (based on the estimated Tau2):
##              Estimate
## I2_2 (Typical v: Q statistic)  0.0000
## I2_3 (Typical v: Q statistic)  0.1993
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 3
## Degrees of freedom: 18
## -2 log likelihood: 9.490174
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
#Notes: I? for level 2 indicates the percent of total variance explained by effects within sites, and I

# # forest plots for each
# ### All forest plots now outdated in favor of metaviz.R
# ### All forest plots now outdated in favor of metaviz.R
# ### All forest plots now outdated in favor of metaviz.R
# data <- combinedresults1
# # same forst plot, but using rma so it plots the aggregate
# dev.off()
# png("./output/comb1.randomeffects.png", type='cairo')
# par(mar=c(4,4,1,4)) #decreasing margins
# forest(rma(yi= data$yi, vi=data$vi, slab=data$sitesource))
# par(cex=1, font=2) #bold font
# text(-3.3, 20.5, "Location", pos=4) #adds location label using x, y coord

```

```

# text(3.8, 20.5, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
# dev.off()
#
# data <- combinedresults2
# # same forst plot, but using rma so it plots the aggregate
# dev.off()
# png("./output/comb2.randomeffects.png", type='cairo')
# par(mar=c(4,4,1,4)) #decreasing margins
# forest(rma(yi= data$yi, vi=data$vi, slab=data$sitesource))
# par(cex=1, font=2) #bold font
# text(-5.1, 20.5, "Location", pos=4) #adds location label using x, y coord
# text(6.6, 20.5, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
# dev.off()
#
# data <- combinedresults3
# # same forst plot, but using rma so it plots the aggregate
# dev.off()
# png("./output/comb3.randomeffects.png", type='cairo')
# par(mar=c(4,4,1,4)) #decreasing margins
# forest(rma(yi= data$yi, vi=data$vi, slab=data$sitesource))
# par(cex=1, font=2) #bold font
# text(-6, 20.5, "Location", pos=4) #adds location label using x, y coord
# text(6.5, 20.5, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
# dev.off()

# a covariate of study version (in-house or expert-designed) is added to create a three-level mixed-eff
# note the openMX status, sometimes indicates a potential problem
summary( mixed_anti0 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti0))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error   lbound   ubound z value Pr(>|z|)
## Intercept -3.4640e-02  6.3819e-02 -1.5972e-01  9.0443e-02 -0.5428   0.5873
## Slope_1    -4.3571e-02  9.9250e-02 -2.3810e-01  1.5096e-01 -0.4390   0.6607
## Tau2_2      1.0000e-10         NA         NA         NA      NA      NA
## Tau2_3      1.1361e-02         NA         NA         NA      NA      NA
##
## Q statistic on the homogeneity of effect sizes: 25.57261
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.1803955
##
## Explained variances (R2):
##               Level 2 Level 3
## Tau2 (no predictor)    1e-10  0.0122
## Tau2 (with predictors)  1e-10  0.0114
## R2                     0e+00  0.0665
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4

```

```

## Degrees of freedom: 17
## -2 log likelihood: -2.079378
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
summary( mixed_anti1 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti1))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti1)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -3.4519e-02  6.2360e-02 -1.5674e-01  8.7704e-02 -0.5535   0.5799
## Slope_1    -4.1639e-02  9.6960e-02 -2.3168e-01  1.4840e-01 -0.4294   0.6676
## Tau2_2      1.0000e-10         NA         NA         NA         NA         NA
## Tau2_3      8.8283e-03         NA         NA         NA         NA         NA
##
## Q statistic on the homogeneity of effect sizes: 24.25077
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.2316096
##
## Explained variances (R2):
##           Level 2 Level 3
## Tau2 (no predictor)      1e-10  0.0097
## Tau2 (with predictors)    1e-10  0.0088
## R2                        0e+00  0.0902
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4
## Degrees of freedom: 17
## -2 log likelihood: -2.99444
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
summary( mixed_anti2 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti2))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti2)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate   Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -4.4108e-02  6.4127e-02 -1.6979e-01  8.1578e-02 -0.6878   0.4916
## Slope_1    -3.9412e-02  1.1996e-01 -2.7454e-01  1.9571e-01 -0.3285   0.7425
## Tau2_2      1.0000e-10         NA         NA         NA         NA         NA
## Tau2_3      1.8346e-02         NA         NA         NA         NA         NA
##
## Q statistic on the homogeneity of effect sizes: 31.12158

```

```

## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.05360574
##
## Explained variances (R2):
##               Level 2 Level 3
## Tau2 (no predictor)      1e-10 0.0188
## Tau2 (with predictors)   1e-10 0.0183
## R2                      0e+00 0.0248
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4
## Degrees of freedom: 17
## -2 log likelihood: 8.420149
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
summary( mixed_anti3 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti3))

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti3)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##               Estimate Std.Error      lbound      ubound z value Pr(>|z|)
## Intercept -3.8248e-02 2.9146e-02 -9.5374e-02 1.8878e-02 -1.3123 0.1894
## Slope_1    2.9817e-02 1.3808e-01 -2.4081e-01 3.0045e-01 0.2159 0.8290
## Tau2_2     1.0000e-10      NA      NA      NA      NA      NA
## Tau2_3     1.2982e-02      NA      NA      NA      NA      NA
##
## Q statistic on the homogeneity of effect sizes: 26.74513
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.1425881
##
## Explained variances (R2):
##               Level 2 Level 3
## Tau2 (no predictor)      1e-10 0.0129
## Tau2 (with predictors)   1e-10 0.0130
## R2                      0e+00 0.0000
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 4
## Degrees of freedom: 17
## -2 log likelihood: 9.444407
## OpenMx status1: 5 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

## Warning in print.summary.meta(x): OpenMx status1 is neither 0 or 1. You are advised to 'rerun' it ag
# Notes: The R² for the version predictor will be reported for both level 2 and level 3, although in th
# constraining the variance to test if it significantly worsens the model

```

```
summary( fixed_anti0 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti0, RE2.c

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti0,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.024552  0.053179 -0.128781  0.079676 -0.4617   0.6443
## Slope_1    -0.049334  0.088837 -0.223451  0.124784 -0.5553   0.5787
##
## Q statistic on the homogeneity of effect sizes: 25.57261
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.1803955
##
## Explained variances (R2):
##                Level 2 Level 3
## Tau2 (no predictor)    1e-10 0.0122
## Tau2 (with predictors)      NA      NA
## R2                      NA      NA
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: -1.351333
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( fixed_anti1 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti1, RE2.c

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti1,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.025226  0.053424 -0.129934  0.079482 -0.4722   0.6368
## Slope_1    -0.047740  0.089038 -0.222251  0.126771 -0.5362   0.5918
##
## Q statistic on the homogeneity of effect sizes: 24.25077
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.2316096
##
## Explained variances (R2):
##                Level 2 Level 3
## Tau2 (no predictor)    1e-10 0.0097
## Tau2 (with predictors)      NA      NA
## R2                      NA      NA
##
```



```

## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: -2.523955
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( fixed_anti2 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti2, RE2.c

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti2,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.025226  0.053424 -0.129934  0.079482 -0.4722   0.6368
## Slope_1   -0.042649  0.106241 -0.250878  0.165580 -0.4014   0.6881
##
## Q statistic on the homogeneity of effect sizes: 31.12158
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.05360574
##
## Explained variances (R2):
##                Level 2 Level 3
## Tau2 (no predictor)    1e-10 0.0188
## Tau2 (with predictors)    NA    NA
## R2                     NA    NA
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: 9.52034
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

summary( fixed_anti3 <- meta3(y=yi, v=vi, cluster=location, x=expert, data=combinedresults_anti3, RE2.c

##
## Call:
## meta3(y = yi, v = vi, cluster = location, x = expert, data = combinedresults_anti3,
##       RE2.constraints = 0, RE3.constraints = 0)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##           Estimate Std.Error    lbound    ubound z value Pr(>|z|)
## Intercept -0.025226  0.053424 -0.129934  0.079482 -0.4722   0.6368
## Slope_1    0.024459  0.129666 -0.229682  0.278601  0.1886   0.8504
##
## Q statistic on the homogeneity of effect sizes: 26.74513
## Degrees of freedom of the Q statistic: 20
## P value of the Q statistic: 0.1425881

```

```
##
## Explained variances (R2):
##               Level 2 Level 3
## Tau2 (no predictor)      1e-10 0.0129
## Tau2 (with predictors)    NA      NA
## R2                      NA      NA
##
## Number of studies (or clusters): 18
## Number of observed statistics: 21
## Number of estimated parameters: 2
## Degrees of freedom: 19
## -2 log likelihood: 9.998565
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

# compare if there is a significant difference in model fit, chi square difference test
anova(mixed_anti0, fixed_anti0)
```

```
##               base               comparison ep  minus2LL df      AIC
## 1 Meta analysis with ML               <NA>  4 -2.079378 17 -36.07938
## 2 Meta analysis with ML Meta analysis with ML  2 -1.351333 19 -39.35133
##      diffLL diffdf              p
## 1          NA      NA          NA
## 2 0.7280452      2 0.6948755
```

```
anova(mixed_anti1, fixed_anti1)
```

```
##               base               comparison ep  minus2LL df      AIC
## 1 Meta analysis with ML               <NA>  4 -2.994440 17 -36.99444
## 2 Meta analysis with ML Meta analysis with ML  2 -2.523955 19 -40.52396
##      diffLL diffdf              p
## 1          NA      NA          NA
## 2 0.4704844      2 0.7903794
```

```
anova(mixed_anti2, fixed_anti2)
```

```
##               base               comparison ep  minus2LL df      AIC
## 1 Meta analysis with ML               <NA>  4 8.420149 17 -25.57985
## 2 Meta analysis with ML Meta analysis with ML  2 9.520340 19 -28.47966
##      diffLL diffdf              p
## 1          NA      NA          NA
## 2 1.100191      2 0.5768947
```

```
anova(mixed_anti3, fixed_anti3)
```

```
##               base               comparison ep  minus2LL df      AIC
## 1 Meta analysis with ML               <NA>  4 9.444407 17 -24.55559
## 2 Meta analysis with ML Meta analysis with ML  2 9.998565 19 -28.00143
##      diffLL diffdf              p
## 1          NA      NA          NA
## 2 0.5541577      2 0.7579947
```

```
# Repeating analyses of "expert" sites in the aggregate, ignoring site dependence.
# This is a simple alternative and useful for most stringent exclusion criteria which drastically reduce
# read in .rds data
data <- readRDS("./data/public/merged_deidentified.rds")
# selecting only expert labs
data <- subset(data, expert==1)
```

```
###ANALYSIS 0: no exclusions###
# t.test and descriptive statistics per condition from psych package
t.test(data$antiauth_avg~data$ms_condition)
```

```
##
## Welch Two Sample t-test
##
## data: data$antiauth_avg by data$ms_condition
## t = -0.91454, df = 783.21, p-value = 0.3607
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3632229 0.1323434
## sample estimates:
## mean in group ms mean in group tv
## 5.473005 5.588444
```

```
describeBy(data$antiauth_avg, group = data$ms_condition)
```

```
##
## Descriptive statistics by group
## group: ms
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 426 5.47 1.77 5.67 5.52 1.98 1 9 8 -0.28 -0.39
## se
## X1 0.09
## -----
## group: tv
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 375 5.59 1.79 5.67 5.66 1.98 1 9 8 -0.37 -0.36
## se
## X1 0.09
```

```
effsize::cohen.d(data$antiauth_avg~data$ms_condition,pooled=TRUE,paired=FALSE,
na.rm=TRUE, hedges.correction=TRUE,
conf.level=0.95)
```

```
##
## Hedges's g
##
## g estimate: -0.06475689 (negligible)
## 95 percent confidence interval:
## lower upper
## -0.20378917 0.07427538
```

```
###ANALYSIS 1: Exclusion set 1###
```

```
# 1. Wrote something for both writing prompts
```

```
data <- subset(data, (data$msincomplete == 0 | is.na(data$msincomplete)))
```

```
# 2. Completed all six items evaluating the essay authors)
```

```
data <- subset(data, (!is.na(data$prou3) & !is.na(data$prou4) & !is.na(data$prou5) & !is.na(data$antiauth_avg)))
```

```
# t.test and descriptive statistics per condition from psych package
```

```
t.test(data$antiauth_avg~data$ms_condition)
```

```
##
## Welch Two Sample t-test
##
## data: data$antiauth_avg by data$ms_condition
```

```
## t = -0.91364, df = 782.71, p-value = 0.3612
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3638690 0.1327349
## sample estimates:
## mean in group ms mean in group tv
## 5.472877 5.588444

describeBy(data$antiauth_avg, group = data$ms_condition)

##
## Descriptive statistics by group
## group: ms
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 424 5.47 1.77 5.67 5.52 1.98 1 9 8 -0.28 -0.4
## se
## X1 0.09
## -----
## group: tv
## vars n mean sd median trimmed mad min max range skew kurtosis
## X1 1 375 5.59 1.79 5.67 5.66 1.98 1 9 8 -0.37 -0.36
## se
## X1 0.09

effsize::cohen.d(data$antiauth_avg-data$ms_condition,pooled=TRUE,paired=FALSE,
na.rm=TRUE, hedges.correction=TRUE,
conf.level=0.95)

##
## Hedges's g
##
## g estimate: -0.06475339 (negligible)
## 95 percent confidence interval:
## lower upper
## -0.20393963 0.07443285

####ANALYSIS 2: Exclusion set 2###
# 1. Wrote something for both writing prompts
# 2. Completed all six items evaluating the essay authors
# 3. Identify as White (race == 1)
data <- subset(data, data$race == 1)
# 4. Born in USA (countryofbirth == 1)
data <- subset(data, data$countryofbirth == 1)
# t.test and descriptive statistics per condition from psych package
t.test(data$antiauth_avg~data$ms_condition)

##
## Welch Two Sample t-test
##
## data: data$antiauth_avg by data$ms_condition
## t = -0.07626, df = 434.65, p-value = 0.9392
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3433484 0.3176993
## sample estimates:
## mean in group ms mean in group tv
```

```
##          5.344086          5.356911
describeBy(data$antiauth_avg, group = data$ms_condition)

##
## Descriptive statistics by group
## group: ms
##   vars   n mean   sd median trimmed  mad min max range  skew kurtosis
## X1      1 248 5.34 1.78   5.33    5.39 1.98   1  9   8 -0.25   -0.45
##      se
## X1 0.11
## -----
## group: tv
##   vars   n mean   sd median trimmed  mad min max range  skew kurtosis   se
## X1      1 205 5.36 1.78   5.67    5.43 1.98   1  9   8 -0.4   -0.39 0.12

effsize::cohen.d(data$antiauth_avg-data$ms_condition,pooled=TRUE,paired=FALSE,
                 na.rm=TRUE, hedges.correction=TRUE,
                 conf.level=0.95) #this is incorrectly indicating a negative value, I'm not sure why bu

##
## Hedges's g
##
## g estimate: -0.007188503 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1926966  0.1783196

###ANALYSIS 3: Exclusion set 3###
# 1. Wrote something for both writing prompts
# 2. Completed all six items evaluating the essay authors
# 3. Identify as White
# 4. Born in USA
# 5. Score a 7 or higher on the American Identity item
data <- subset(data, data$americanid >= 7)
# t.test and descriptive statistics per condition from psych package
t.test(data$antiauth_avg-data$ms_condition)

##
## Welch Two Sample t-test
##
## data: data$antiauth_avg by data$ms_condition
## t = 0.40559, df = 263.1, p-value = 0.6854
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3399641  0.5163523
## sample estimates:
## mean in group ms mean in group tv
##      5.082774      4.994580

describeBy(data$antiauth_avg, group = data$ms_condition)

##
## Descriptive statistics by group
## group: ms
##   vars   n mean   sd median trimmed  mad min max range  skew kurtosis
## X1      1 149 5.08 1.81     5    5.08 1.98   1  9   8 -0.03   -0.6
```

```

##      se
## X1 0.15
## -----
## group: tv
##      vars   n mean   sd median trimmed  mad min  max range  skew kurtosis
## X1      1 123 4.99 1.76      5    5.04 1.98   1 8.67  7.67 -0.22   -0.53
##      se
## X1 0.16

effsize::cohen.d(data$antiauth_avg~data$ms_condition,pooled=TRUE,paired=FALSE,
                  na.rm=TRUE, hedges.correction=TRUE,
                  conf.level=0.95) #this is incorrectly indicating a positive value, reversing sign in t

##
## Hedges's g
##
## g estimate: 0.04912773 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1907571  0.2890126

###Conducting a small meta-analysis of only the in-house data to provide a summary of those results in
# Read in summary .csv which used basic exclusion rules, Exclusion Set 1
data <- read.csv("./data/public/combinedresults_antil.csv")
# subset to in-house rows only
data <- subset(data, expert==0)
# conduct random effects meta-analysis
summary( meta(y = yi, v = vi, data = data))

##
## Call:
## meta(y = yi, v = vi, data = data)
##
## 95% confidence intervals: z statistic approximation
## Coefficients:
##      Estimate Std. Error   lbound   ubound z value Pr(>|z|)
## Intercept1 -0.034056  0.066351 -0.164102  0.095990 -0.5133  0.6078
## Tau2_1_1    0.013576  0.023351 -0.032192  0.059343  0.5814  0.5610
##
## Q statistic on the homogeneity of effect sizes: 17.39158
## Degrees of freedom of the Q statistic: 11
## P value of the Q statistic: 0.09681464
##
## Heterogeneity indices (based on the estimated Tau2):
##      Estimate
## Intercept1: I2 (Q statistic)  0.2764
##
## Number of studies (or clusters): 12
## Number of observed statistics: 12
## Number of estimated parameters: 2
## Degrees of freedom: 10
## -2 log likelihood: 1.143547
## OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
## Other values may indicate problems.)

```

```

# # forest plot
# dev.off()
# par(mar=c(4,4,1,4)) #decreasing margins
# forest(x= data$yi, vi=data$vi, slab=data$location)
# par(cex=1, font=2)#bold font
# text(-3.3, 13, "Location", pos=4) #adds location label using x, y coord
# text(3.8, 13, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
#
# # same forest plot, but using rma so it plots the aggregate
# dev.off()
# png("./output/inhousemeta.png", type='cairo')
# par(mar=c(4,4,1,4)) #decreasing margins
# forest(rma(yi= data$yi, vi=data$vi, slab=data$location))
# par(cex=1, font=2) #bold font
# text(-3.3, 13, "Location", pos=4) #adds location label using x, y coord
# text(3.8, 13, "SMD [95% CI]", pos=2) #adds standardized mean diff label using x y coord
# dev.off()

# sample funnel plot
# funnel(rma(yi= data$yi, vi=data$vi, slab=data$location))

# Focused analysis of sites with "expert" or "a lot of knowledge about TMT" leads
# Still using exclusion set 1
# Read data
data <- merged
# Applying exclusion criteria 1
# 1. Wrote something for both writing prompts
data <- subset(data, (data$msincomplete == 0 | is.na(data$msincomplete)))
# 2. Completed all six items evaluating the essay authors)
data <- subset(data, (!is.na(data$prou3) & !is.na(data$prou4) & !is.na(data$prou5) & !is.na(data$antiauth_avg)))
# Selecting only the below sites:
#University of Wisconsin, Madison, WI (in-house)
#The College of New Jersey
#University of Kansas (Expert)
#University of Kansas (in-house)
#Pace University (expert)
#Virginia Commonwealth University, Richmond, VA
data <- subset(data, data$source=="uwmadison_inhouse" | data$source=="cnj" | data$source=="kansas_expert")
# Applying the same levels fix as earlier, only because it caused problems in
# cohen.d() below. May not be necessary anymore.
data$ms_condition <- factor(data$ms_condition, levels = c("ms", "tv"))
# Analyses using that subset
t.test(data$antiauth_avg~data$ms_condition)

##
## Welch Two Sample t-test
##
## data: data$antiauth_avg by data$ms_condition
## t = -0.08774, df = 523.16, p-value = 0.9301
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3251113 0.2973122
## sample estimates:
## mean in group ms mean in group tv

```

```
##          5.276596          5.290495
```

```
describeBy(data$antiauth_avg, group = data$ms_condition)
```

```
##
```

```
## Descriptive statistics by group
```

```
## group: ms
```

```
##      vars   n mean    sd median trimmed  mad min max range  skew kurtosis
## X1      1 282 5.28 1.84   5.33   5.34 1.98   1  9   8 -0.28   -0.45
```

```
##      se
```

```
## X1 0.11
```

```
## -----
```

```
## group: tv
```

```
##      vars   n mean    sd median trimmed  mad min max range  skew kurtosis
## X1      1 249 5.29 1.81   5.33   5.34 1.98   1  9   8 -0.24   -0.51
```

```
##      se
```

```
## X1 0.11
```

```
effsize::cohen.d(data$antiauth_avg-data$ms_condition,pooled=TRUE,paired=FALSE,
```

```
na.rm=TRUE, hedges.correction=TRUE,
```

```
conf.level=0.95) #this was previously incorrectly indicating a positive value? Had to
```

```
##
```

```
## Hedges's g
```

```
##
```

```
## g estimate: -0.007610053 (negligible)
```

```
## 95 percent confidence interval:
```

```
##      lower      upper
```

```
## -0.1784414  0.1632213
```