

# Examining the role of orthographic neighbourhood effects during lateralised lexical decision paradigms: A meta-analysis

Adam Parker

03/06/2020

This markdown file covers the proposed analysis for a meta-analysis of orthographic Nsize effects in the right and left visual field. As we are primarily interested in the effect size in each hemifield, rather than the reliability of interaction between visual field and Nsize, we choose to conduct two meta-analyses testing the main effect in each.

This Rmarkdown accompanies a pre-registration on the OSF: XXX.

## Random effects models

We chose to use a random effects model as the studies will differ with several respects: language of participants, Nsizes, etc.

This is done seperately for the LVF and RVF.

The below output includes the estimate from the random-effects models, measures of heterogeneity, and precision which are all covered in the registration document. The anlayses are run using the `rma()` function from the **metafor** package. The effect sizes and variance are calculated using a function (ES) that uses the equations presented in Borenstein et al. (2009) for calculating variances for within subjects designs.

We also conduct sensitvty analysis where the estimated correlation is replaced.

## Right visual field

First, read the data.

```
data <- read.csv("extracted_dat.csv", na.strings = " ")
# now get the correlations for the RVF
corrR <- mean(data$r.1, na.rm= TRUE)
# create RVF data
RVF_dat <- dplyr::select(data, Source.name, Source.year, RVF.High.N, RVF.High.N.sd, RVF.Low.N, RVF.Low.N.sd)
# rename for package
RVF_dat <- RVF_dat %>%
  rename(
    highN      = RVF.High.N,
    highN_sd   = RVF.High.N.sd,
    lowN       = RVF.Low.N,
    lowN_sd    = RVF.Low.N.sd)
# add columns
RVF_dat$study <- paste0(RVF_dat$Source.name, " (", RVF_dat$Source.year, ")")
```

```
RVF_dat$corr <- corrR
RVF_dat$VF <- "RVF"
# ensure formatted correctly
RVF_dat$sample_size <- as.numeric(RVF_dat$Sample.size)
RVF_dat$highN <- as.numeric(as.character(RVF_dat$highN))
RVF_dat$highN_sd <- as.numeric(as.character(RVF_dat$highN_sd))
RVF_dat$lowN <- as.numeric(as.character(RVF_dat$lowN))
RVF_dat$lowN_sd <- as.numeric(as.character(RVF_dat$lowN_sd))
RVF_dat$corr <- as.numeric(as.character(RVF_dat$corr))
```

Now that we have the simulated data, we need to calculate the mean difference for a within subjects design. We do this using a function based on the formula presented in Borenstein, Hedges, Higgins, and Rothstein (2009).

```
ES<- function(df, M1, M2, S1, S2, N, r){
  # mean difference
  df$MD<- M1-M2 # equation 4.12
  # standard deviation of the difference
  df$sdif<- sqrt(S1^2 + S2^2 - 2*r*S1*S2) # equation 4.15
  # variance
  df$var<- (df$sdif^2)/N # equation 4.13
  # standard error
  df$se<- sqrt(df$var) # equation 4.14
  return(df)
}
```

Now we have data simulated, start proposed analysis.

```
# calculate the ES
#W do this for RVF
RVF.ES <- ES(RVF_dat, RVF_dat$highN, RVF_dat$lowN, RVF_dat$highN_sd, RVF_dat$lowN_sd, RVF_dat$sample_size)
# now run the meta analysis
RVF.model <- rma(yi= MD, vi= var, sei= se, data=RVF.ES, method="SJ", test="knha" , digits=3)
```

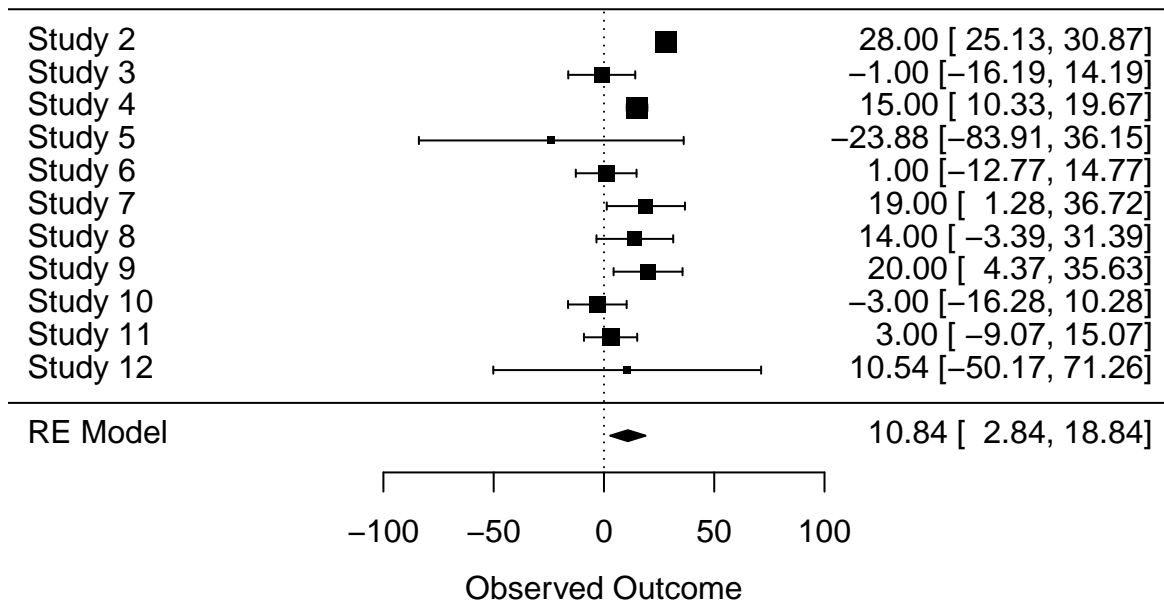
```
## Warning in rma(yi = MD, vi = var, sei = se, data = RVF.ES, method = "SJ", :
## Studies with NAs omitted from model fitting.
```

```
# print
RVF.model
```

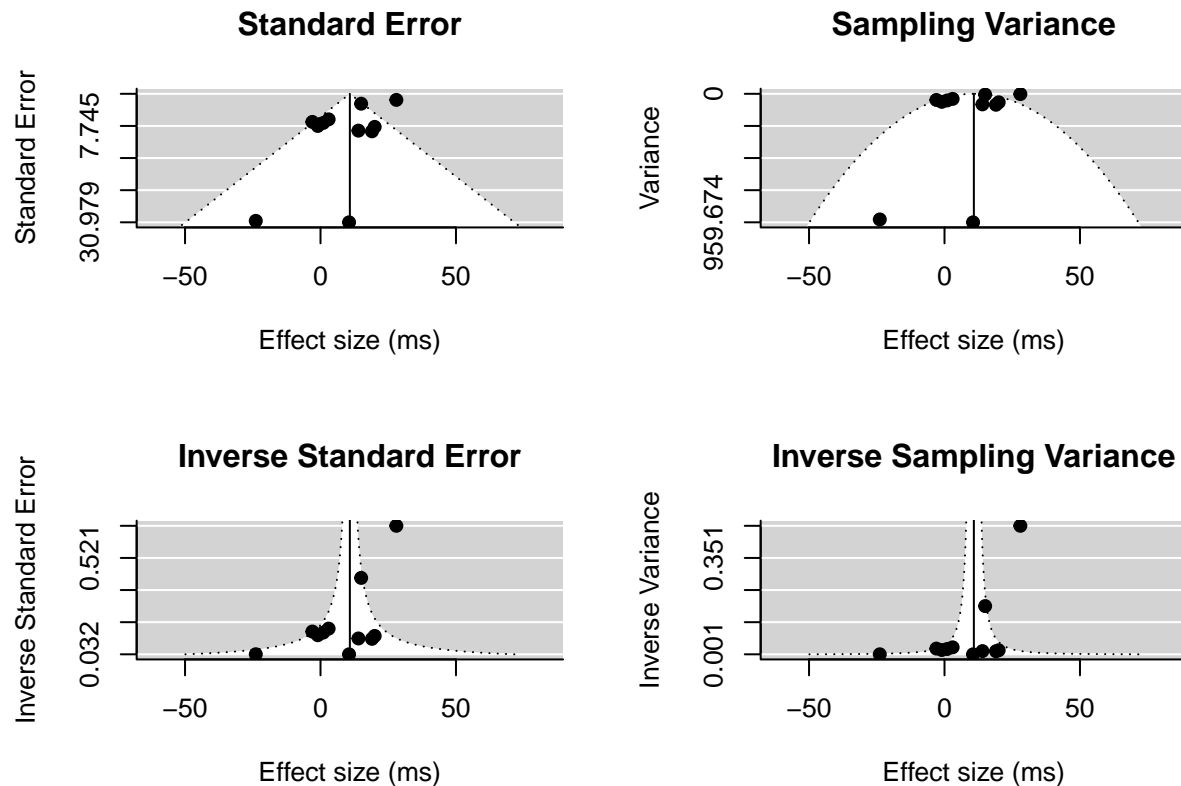
```
##
## Random-Effects Model (k = 11; tau^2 estimator: SJ)
##
## tau^2 (estimated amount of total heterogeneity): 100.782 (SE = 58.098)
## tau (square root of estimated tau^2 value): 10.039
## I^2 (total heterogeneity / total variability): 81.75%
## H^2 (total variability / sampling variability): 5.48
##
## Test for Heterogeneity:
## Q(df = 10) = 69.345, p-val < .001
##
## Model Results:
```

```
##
## estimate      se    tval    pval   ci.lb   ci.ub
##   10.842   3.589   3.021   0.013   2.844   18.839 *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# plot
forest(RVF.model)
```



```
# FUNNEL plots
# set up 2x2 array for plotting
par(mfrow=c(2,2))
# draw funnel plots
funnel(RVF.model, main="Standard Error", xlab="Effect size (ms)")
funnel(RVF.model, yaxis="vi", main="Sampling Variance", xlab="Effect size (ms)")
funnel(RVF.model, yaxis="seinv", main="Inverse Standard Error", xlab="Effect size (ms)")
funnel(RVF.model, yaxis="vinv", main="Inverse Sampling Variance", xlab="Effect size (ms)")
```



The estimate is significantly different from zero, so we can conclude that the effect is truly inhibitory in the RVF/left hemisphere. Now let's run some sensitivity analyses using predetermined levels of correlation.

```
# specify correlations
corlist <- c(.2, .5, .8, 1)
# create temporary data to write to before placing within forCurve
senseDat <- data.frame(matrix(ncol = 5, nrow = length(corlist)))
colnames(senseDat) <- c("cor", "beta", "CI-", "CI+", "p")
# set row counter
myrow <- 0
# start sensitivity analysis
for (i in corlist) {

  RVF.ES <- ES(RVF_dat, RVF_dat$highN, RVF_dat$lowN, RVF_dat$highN_sd, RVF_dat$lowN_sd, RVF_dat$sample_)
  # now run the meta analysis
  RVF.model <- rma(yi= MD, vi= var, sei= se, data=RVF.ES, method="SJ", test="knha" , digits=3)

  # flip counter
  myrow <- myrow+1
  # take the useful estimates
  senseDat$cor[myrow] <- as.numeric(i) # cor
  senseDat$beta[myrow] <- as.numeric(RVF.model[2]) # beta
  senseDat$p[myrow] <- as.numeric(RVF.model[5]) # p
  senseDat$CI-[myrow] <- as.numeric(RVF.model[6]) # lower CI
  senseDat$CI+[myrow] <- as.numeric(RVF.model[7]) # lower CI
}
```

```
}
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = RVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = RVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = RVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

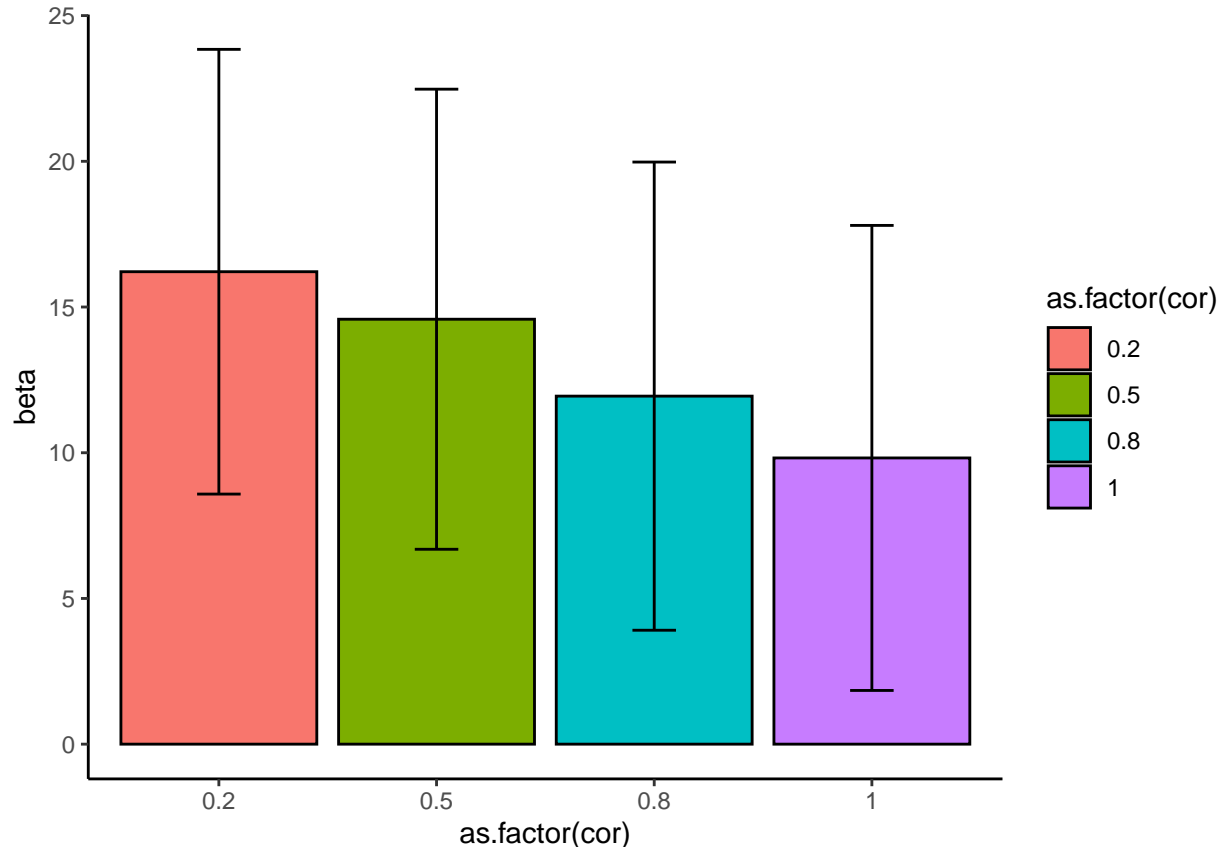
```
## Warning in rma(yi = MD, vi = var, sei = se, data = RVF.ES, method = "SJ", :  
## There are outcomes with non-positive sampling variances.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = RVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = RVF.ES, method = "SJ", :  
## Cannot compute Q-test, I2, or H2 when there are non-positive sampling  
## variances in the data.
```

```
# now plot the outcome
```

```
ggplot(senseDat, aes(x=as.factor(cor), y=beta, fill=as.factor(cor))) +  
  geom_bar(stat="identity", color="black", position=position_dodge()) +  
  geom_errorbar(aes(ymin='CI-', ymax='CI+'), width=.2, position=position_dodge(.9)) +  
  theme_classic()
```



Regardless of the correlation, it appears that the effect is inhibitory.

## Left visual field

First, we run again for LVF data.

```
# now get the correlations for the RVF
corrL <- mean(data$r, na.rm= TRUE)
# create RVF data
LVF_dat <- dplyr::select(data, Source.name, Source.year, LVF.High.N, LVF.High.N.sd, LVF.Low.N, LVF.Low.N.sd)
# rename for package
LVF_dat <- LVF_dat %>%
  rename(
    highN      = LVF.High.N,
    highN_sd   = LVF.High.N.sd,
    lowN       = LVF.Low.N,
    lowN_sd    = LVF.Low.N.sd)
# add columns
LVF_dat$study <- paste0(LVF_dat$Source.name, " (", LVF_dat$Source.year, ")")
LVF_dat$corr <- corrL
LVF_dat$VF <- "LVF"
# ensure formatted correctly
LVF_dat$sample_size <- as.numeric(LVF_dat$Sample.size)
LVF_dat$highN <- as.numeric(as.character(LVF_dat$highN))
LVF_dat$highN_sd <- as.numeric(as.character(LVF_dat$highN_sd))
LVF_dat$lowN <- as.numeric(as.character(LVF_dat$lowN))
LVF_dat$lowN_sd <- as.numeric(as.character(LVF_dat$lowN_sd))
LVF_dat$corr <- as.numeric(as.character(LVF_dat$corr))
```

Now run the analysis for the LVF.

```
# calculate the ES
#W do this for LVF
LVF.ES <- ES(LVF_dat, LVF_dat$highN, LVF_dat$lowN, LVF_dat$highN_sd, LVF_dat$lowN_sd, LVF_dat$sample_size)
# now run the meta analysis
LVF.model <- rma(yi= MD, vi= var, sei= se, data=LVF.ES, method="SJ", test="knha" , digits=3)
```

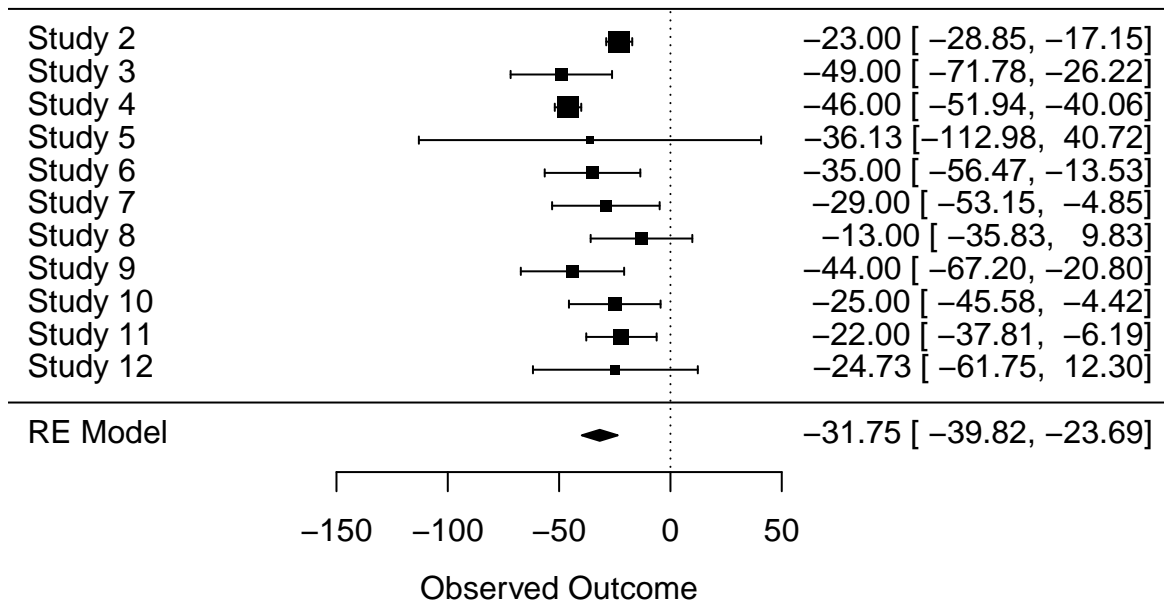
```
## Warning in rma(yi = MD, vi = var, sei = se, data = LVF.ES, method = "SJ", :
## Studies with NAs omitted from model fitting.
```

```
# print
LVF.model
```

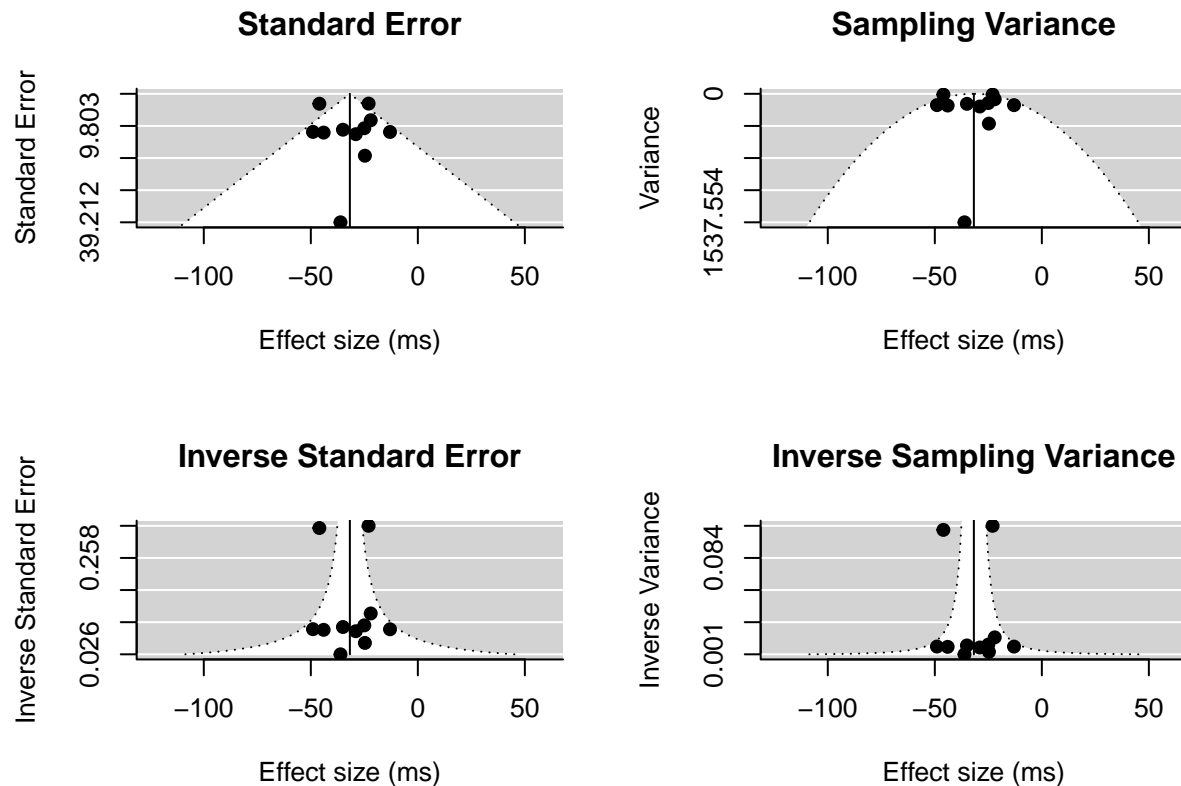
```
##
## Random-Effects Model (k = 11; tau^2 estimator: SJ)
##
## tau^2 (estimated amount of total heterogeneity): 73.892 (SE = 43.217)
## tau (square root of estimated tau^2 value):      8.596
## I^2 (total heterogeneity / total variability):   59.43%
## H^2 (total variability / sampling variability):   2.47
##
## Test for Heterogeneity:
```

```
## Q(df = 10) = 38.107, p-val < .001
##
## Model Results:
##
## estimate      se      tval    pval    ci.lb    ci.ub
## -31.755  3.619  -8.774  <.001  -39.818  -23.691  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# plot
forest(LVF.model)
```



```
# FUNNEL plots
# set up 2x2 array for plotting
par(mfrow=c(2,2))
# draw funnel plots
funnel(LVF.model, main="Standard Error", xlab="Effect size (ms)")
funnel(LVF.model, yaxis="vi", main="Sampling Variance", xlab="Effect size (ms)")
funnel(LVF.model, yaxis="seinv", main="Inverse Standard Error", xlab="Effect size (ms)")
funnel(LVF.model, yaxis="vinv", main="Inverse Sampling Variance", xlab="Effect size (ms)")
```



The estimate is significantly different from zero, so we can conclude that the effect is truly facilitatory in the LVF/right hemisphere hemisphere. Now let's run some sensitivity analyses using predetermined levels of correlation.

```
# specify correlations
corlist <- c(.2, .5, .8, 1)
# create temporary data to write to before placing within forCurve
senseDat <- data.frame(matrix(ncol = 5, nrow = length(corlist)))
colnames(senseDat) <- c("cor", "beta", "CI-", "CI+", "p")
# set row counter
myrow <- 0
# start sensitivity analysis
for (i in corlist) {

  LVF.ES <- ES(LVF_dat, LVF_dat$highN, LVF_dat$lowN, LVF_dat$highN_sd, LVF_dat$lowN_sd, LVF_dat$sample_
  # now run the meta analysis
  LVF.model <- rma(yi= MD, vi= var, sei= se, data=LVF.ES, method="SJ", test="knha" , digits=3)

  # flip counter
  myrow <- myrow+1
  # take the useful estimates
  senseDat$cor[myrow] <- as.numeric(i) # cor
  senseDat$beta[myrow] <- as.numeric(LVF.model[2]) # beta
  senseDat$p[myrow] <- as.numeric(LVF.model[5]) # p
  senseDat$CI-[myrow] <- as.numeric(LVF.model[6]) # lower CI
  senseDat$CI+[myrow] <- as.numeric(LVF.model[7]) # lower CI
}
```



```
}
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = LVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = LVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

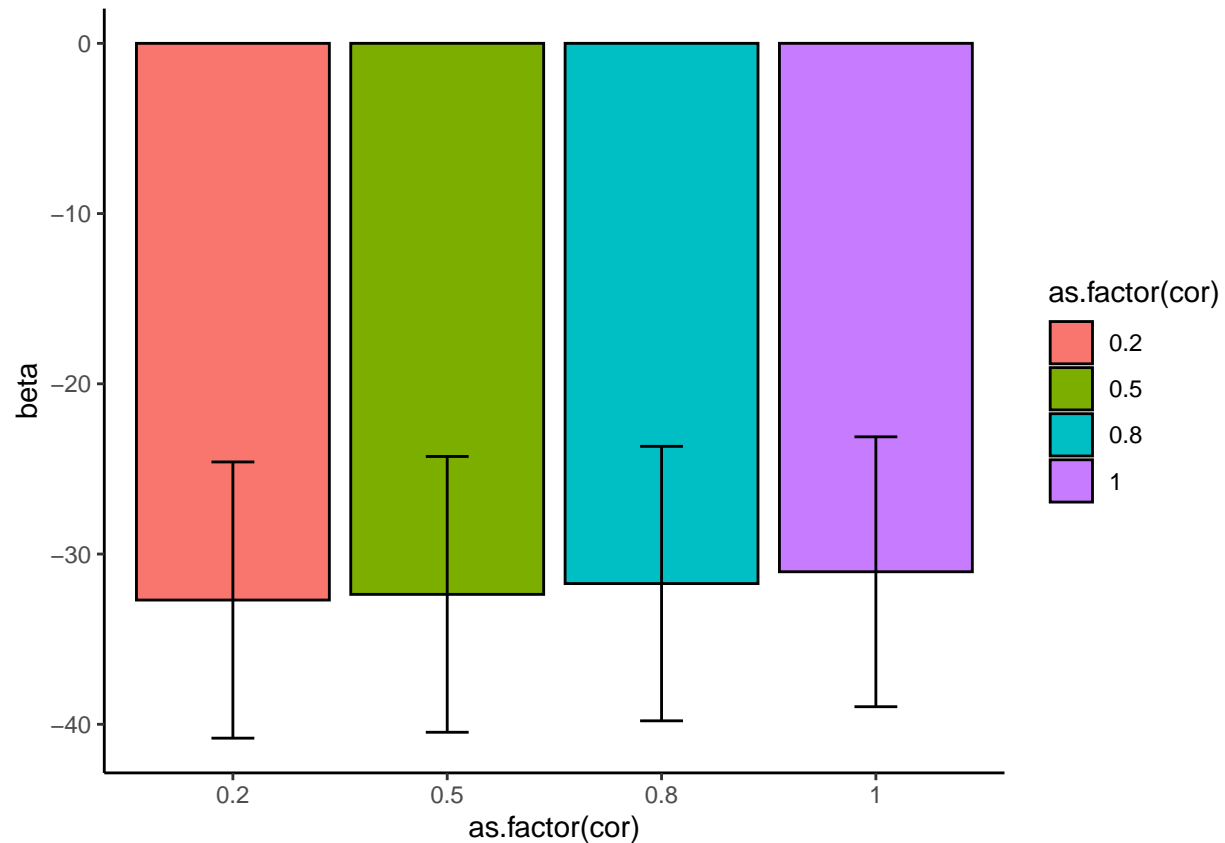
```
## Warning in rma(yi = MD, vi = var, sei = se, data = LVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = LVF.ES, method = "SJ", :  
## There are outcomes with non-positive sampling variances.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = LVF.ES, method = "SJ", :  
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, sei = se, data = LVF.ES, method = "SJ", :  
## Cannot compute Q-test,  $I^2$ , or  $H^2$  when there are non-positive sampling  
## variances in the data.
```

```
# now plot the outcome  
ggplot(senseDat, aes(x=as.factor(cor), y=beta, fill=as.factor(cor))) +  
  geom_bar(stat="identity", color="black", position=position_dodge()) +  
  geom_errorbar(aes(ymin='CI-', ymax='CI+', width=.2, position=position_dodge(.9)) +  
    theme_classic()
```



Regardless of the correlation used, the effect is always inhibitory.

## Exploratory analysis

This indicates that

### Difference between the visual fields.

This exploratory analysis lets us check that there is a reliable difference between the two across visual fields.

```
# combined visual fields
two.VF <- rbind(RVF.ES, LVF.ES)
# set contrasts
two.VF$VF <- as.factor(two.VF$VF)
contrasts(two.VF$VF) <- contr.sum
# now run lm
rma(yi= MD, vi= var, mods = ~ factor(VF), data=two.VF)
```

```
## Warning in rma(yi = MD, vi = var, mods = ~factor(VF), data = two.VF): There are
## outcomes with non-positive sampling variances.
```

```
## Warning in rma(yi = MD, vi = var, mods = ~factor(VF), data = two.VF): Studies
## with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, mods = ~factor(VF), data = two.VF): Cannot
## compute QE-test, I^2, or H^2 when there are non-positive sampling variances in
## the data.

##
## Mixed-Effects Model (k = 22; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      129.5184 (SE = 45.2225)
## tau (square root of estimated tau^2 value):             11.3806
## R^2 (amount of heterogeneity accounted for):             76.40%
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 61.3627, p-val < .0001
##
## Model Results:
##
##              estimate      se      zval      pval      ci.lb      ci.ub
## intrcpt          -31.0458  3.6004  -8.6228 <.0001   -38.1025  -23.9890 ***
## factor(VF)RVF     40.8511  5.2150   7.8334 <.0001    30.6300   51.0723 ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This indicates that there is a real strong statistical difference between the two fields. We knew this, but it is reassuring to know.

## Number of trials

With few trials, the estimates means may not be so reliable. Let's have a look at how the number of word trials per condition influence the effects.

```
data$study <- paste0(data$Source.name, " (", data$Source.year, ")")
# get number of trials
trials.dat <- dplyr::select(data, study, Number.of.word.trials)
# merge them
two.VF<- cbind(two.VF, trials.dat)
# now run lm
rma(yi= MD, vi= var, mods = ~ Number.of.word.trials, data=two.VF)
```

```
## Warning in rma(yi = MD, vi = var, mods = ~Number.of.word.trials, data = two.VF):
## There are outcomes with non-positive sampling variances.
```

```
## Warning in rma(yi = MD, vi = var, mods = ~Number.of.word.trials, data = two.VF):
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, mods = ~Number.of.word.trials, data = two.VF):
## Cannot compute QE-test, I^2, or H^2 when there are non-positive sampling
## variances in the data.
```

```
##
## Mixed-Effects Model (k = 22; tau^2 estimator: REML)
```

```
##
## tau^2 (estimated amount of residual heterogeneity):      576.1317 (SE = 194.6378)
## tau (square root of estimated tau^2 value):             24.0027
## R^2 (amount of heterogeneity accounted for):             0.00%
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.0426, p-val = 0.8365
##
## Model Results:
##
##               estimate      se      zval      pval      ci.lb      ci.ub
## intrcpt          -9.4084  12.8639  -0.7314  0.4645  -34.6212  15.8044
## Number.of.word.trials -0.0207   0.1004  -0.2063  0.8365   -0.2176   0.1761
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This didn't seem to have a very large influence.

## Size of the N manipulation

```
data$ndif <- as.numeric(data$High.N) - as.numeric(data$Low.N)
# get number ndif
ndif.dat <- dplyr::select(data, study, ndif)
# merge them
two.VF<- cbind(two.VF, ndif.dat)
# now run lm
rma(yi= MD, vi= var, mods = ~ factor(VF) * ndif, data=two.VF)
```

```
## Warning in rma(yi = MD, vi = var, mods = ~factor(VF) * ndif, data = two.VF):
## There are outcomes with non-positive sampling variances.
```

```
## Warning in rma(yi = MD, vi = var, mods = ~factor(VF) * ndif, data = two.VF):
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(yi = MD, vi = var, mods = ~factor(VF) * ndif, data = two.VF):
## Cannot compute QE-test, I^2, or H^2 when there are non-positive sampling
## variances in the data.
```

```
##
## Mixed-Effects Model (k = 22; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):      114.5327 (SE = 42.6749)
## tau (square root of estimated tau^2 value):             10.7020
## R^2 (amount of heterogeneity accounted for):             79.13%
##
## Test of Moderators (coefficients 2:4):
## QM(df = 3) = 73.4534, p-val < .0001
##
## Model Results:
```

```
##
##           estimate      se      zval      pval      ci.lb      ci.ub
## intrcpt      -34.0992  12.9481  -2.6335  0.0085  -59.4770   -8.7214   **
## factor(VF)RVF    70.1328  18.4619   3.7988  0.0001   33.9482  106.3174   ***
## ndif           0.3825   1.5534   0.2462  0.8055   -2.6622    3.4271
## factor(VF)RVF:ndif -3.6304   2.2099  -1.6427  0.1004   -7.9618    0.7010
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

There isn't much support for the strength of the manipulation influencing this.