Report III.

František Kalvas

2022-03-18

## Introduction

Main focus of the third report is to present reduced model. We agreed during meeting at 2022-03-18 that we:

1. will use only ESBG measure (not to uncertain reader whether the results are case of model or polarization measure)
2. will omit the ‘vaguely-speak’ mode (since it’s not intuitive and we hardly finding natural examples)
3. rename variables (see details bellow)

## Loading data

Data are at <http://github.com/frantisek901/Spirals/Experiment>. Experiment is still running and I, FranČesko, from time to time actualize the \*.csv files at GitHub, then I run script experiment.R which loads the data. Later version probably finds better names for variables, but now, I use default names from NetLogo experiment.

Who is not interested in working with megabytes of \*.csv files, might use compiled \*.RData, there are three files: shortData1D.RData, shortData.RData and shortData4D.RData, which are main data files from experiments running only 365 steps in 1D, 2D and 4D opinion spaces, these data are extended by extra simulations with low size of small-world network neighborhood, very narrow boundary and high probability of speaking.

For avoiding statistical artifact we sampling data – for each combination of important variables same number of observations/simulations. Here we must note, that per 1 simulation not using identity we have 3 simulations using identity, since it makes no sense to vary Narrownes of identity group in case we are not using identity in the model.

Now we load and join these data and factorize and rename selected variables:

## Loading stored data  
load("shortData.RData")  
load("shortData1D.RData")  
load("shortData4D.RData")  
  
  
## Filtering data objects  
res = filter(res, mode == "openly-listen")  
res1D = filter(res1D, mode == "openly-listen")  
res4D = filter(res4D, mode == "openly-listen")  
  
  
## We control the ratios of boundary and identity use  
num = 300  
df = sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.1 & res4D$id\_threshold == 0.39,], num) %>%  
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.1 & res4D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.1 & res4D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res4D[!res4D$`use\_identity?` & res4D$boundary == 0.1,], num)) %>%   
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.22 & res4D$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.22 & res4D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.22 & res4D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res4D[!res4D$`use\_identity?` & res4D$boundary == 0.22,], num)) %>%   
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.28 & res4D$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.28 & res4D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.28 & res4D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res4D[!res4D$`use\_identity?` & res4D$boundary == 0.28,], num)) %>%   
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.34 & res4D$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.34 & res4D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res4D[res4D$`use\_identity?` & res4D$boundary == 0.34 & res4D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res4D[!res4D$`use\_identity?` & res4D$boundary == 0.34,], num)) %>%   
  
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.1 & res1D$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.1 & res1D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.1 & res1D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res1D[!res1D$`use\_identity?` & res1D$boundary == 0.1,], num)) %>%   
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.22 & res1D$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.22 & res1D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.22 & res1D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res1D[!res1D$`use\_identity?` & res1D$boundary == 0.22,], num)) %>%   
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.28 & res1D$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.28 & res1D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.28 & res1D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res1D[!res1D$`use\_identity?` & res1D$boundary == 0.28,], num)) %>%   
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.34 & res1D$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.34 & res1D$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res1D[res1D$`use\_identity?` & res1D$boundary == 0.34 & res1D$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res1D[!res1D$`use\_identity?` & res1D$boundary == 0.34,], num)) %>%   
   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.1 & res$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.1 & res$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.1 & res$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res[!res$`use\_identity?` & res$boundary == 0.1,], num)) %>%   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.22 & res$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.22 & res$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.22 & res$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res[!res$`use\_identity?` & res$boundary == 0.22,], num)) %>%   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.28 & res$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.28 & res$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.28 & res$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res[!res$`use\_identity?` & res$boundary == 0.28,], num)) %>%   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.34 & res$id\_threshold == 0.39,], num)) %>%  
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.34 & res$id\_threshold == 0.49,], num)) %>%   
 add\_row(sample\_n(res[res$`use\_identity?` & res$boundary == 0.34 & res$id\_threshold == 0.59,], num)) %>%   
 add\_row(sample\_n(res[!res$`use\_identity?` & res$boundary == 0.34,], num)) %>%   
  
 ## Selecting variables:  
 select(opinions, boundary, `use\_identity?`, id\_threshold, `conformity-level`, `p-speaking-level`,   
 `tolerance-level`, `p-random`, `n-neis`, ESBG\_365) %>%   
   
 ## Changing some variables to factors:  
 mutate(#id\_threshold = if\_else(!`use\_identity?`, NA\_real\_, id\_threshold),  
 opinions = factor(opinions), boundary = factor(boundary), id\_threshold = factor(id\_threshold),  
 identity = if\_else(`use\_identity?`, "Identity used", "Identity not used") %>% factor()) %>%   
 relocate(identity, .after = `use\_identity?`) %>% select(-`use\_identity?`) %>%   
  
 ## Renaming variables according 2022-03-18 meeting:  
 prejmenuj(1:10, c("Opinion dimensions:", "Acceptability of different opinion:", "Identity:",   
 "Narrowness of identity group:", "conformity", "speaking", "tolerance", "random links",  
 "close links", "ESBG"))

## Regressions

With new package jtools I succeeded in formatting results into nice table.

# Control variables  
ec = (lm(ESBG~conformity+speaking+tolerance+`random links`+`close links`, df))  
  
# boundary + ec  
eb = (lm(ESBG~`Acceptability of different opinion:`+conformity+speaking+tolerance+`random links`+`close links`, df))  
  
# identity + eb  
ei = (lm(ESBG~`Narrowness of identity group:`+`Identity:`+`Acceptability of different opinion:`+conformity+speaking+tolerance+`random links`+`close links`, df))  
  
# Full model: opinions + ei  
ef = (lm(ESBG~`Opinion dimensions:`+`Narrowness of identity group:`+`Identity:`+`Acceptability of different opinion:`+conformity+speaking+tolerance+`random links`+`close links`, df))  
  
# Only used variables  
eu = (lm(ESBG~`Opinion dimensions:`+`Narrowness of identity group:`+`Identity:`+`Acceptability of different opinion:`, df))  
  
# Only used variables with interactions  
eui = (lm(ESBG~`Identity:`+`Narrowness of identity group:`\*`Acceptability of different opinion:`\*`Opinion dimensions:`, df))  
  
export\_summs(ec, eb, ei, ef, eu)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
| (Intercept) | 0.10 \*\*\* | 0.11 \*\*\* | -0.09 \*\*\* | 0.18 \*\*\* | 0.15 \*\*\* |
|  | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) |
| conformity | -0.02 | -0.02 | -0.02 | -0.02 |  |
|  | (0.02) | (0.02) | (0.02) | (0.02) |  |
| speaking | -0.06 \*\*\* | -0.06 \*\*\* | -0.06 \*\*\* | -0.01 |  |
|  | (0.01) | (0.01) | (0.01) | (0.01) |  |
| tolerance | 0.03 \*\*\* | 0.02 \*\*\* | 0.02 \*\*\* | -0.01 \* |  |
|  | (0.01) | (0.01) | (0.01) | (0.00) |  |
| `random links` | 0.53 \*\*\* | 0.54 \*\*\* | 0.53 \*\*\* | -0.00 |  |
|  | (0.01) | (0.01) | (0.01) | (0.01) |  |
| `close links` | -0.00 | -0.00 \* | -0.00 \* | -0.00 \*\* |  |
|  | (0.00) | (0.00) | (0.00) | (0.00) |  |
| `Acceptability of different opinion:`0.22 |  | -0.01 \* | -0.01 \*\* | -0.01 \*\*\* | -0.01 \*\*\* |
|  |  | (0.00) | (0.00) | (0.00) | (0.00) |
| `Acceptability of different opinion:`0.28 |  | -0.02 \*\*\* | -0.02 \*\*\* | -0.02 \*\*\* | -0.02 \*\*\* |
|  |  | (0.00) | (0.00) | (0.00) | (0.00) |
| `Acceptability of different opinion:`0.34 |  | -0.03 \*\*\* | -0.03 \*\*\* | -0.03 \*\*\* | -0.03 \*\*\* |
|  |  | (0.00) | (0.00) | (0.00) | (0.00) |
| `Narrowness of identity group:`0.49 |  |  | 0.09 \*\*\* | 0.09 \*\*\* | 0.09 \*\*\* |
|  |  |  | (0.00) | (0.00) | (0.00) |
| `Narrowness of identity group:`0.59 |  |  | 0.09 \*\*\* | 0.09 \*\*\* | 0.09 \*\*\* |
|  |  |  | (0.00) | (0.00) | (0.00) |
| `Identity:`Identity used |  |  | 0.19 \*\*\* | 0.19 \*\*\* | 0.19 \*\*\* |
|  |  |  | (0.00) | (0.00) | (0.00) |
| `Opinion dimensions:`2 |  |  |  | -0.16 \*\*\* | -0.16 \*\*\* |
|  |  |  |  | (0.00) | (0.00) |
| `Opinion dimensions:`4 |  |  |  | -0.24 \*\*\* | -0.24 \*\*\* |
|  |  |  |  | (0.00) | (0.00) |
| N | 14400 | 14400 | 14400 | 14400 | 14400 |
| R2 | 0.10 | 0.10 | 0.31 | 0.54 | 0.54 |
| \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05. | | | | | |

export\_summs(eu, eui)

|  |  |  |
| --- | --- | --- |
|  | Model 1 | Model 2 |
| (Intercept) | 0.15 \*\*\* | 0.21 \*\*\* |
|  | (0.00) | (0.01) |
| `Opinion dimensions:`2 | -0.16 \*\*\* | -0.21 \*\*\* |
|  | (0.00) | (0.01) |
| `Opinion dimensions:`4 | -0.24 \*\*\* | -0.31 \*\*\* |
|  | (0.00) | (0.01) |
| `Narrowness of identity group:`0.49 | 0.09 \*\*\* | 0.05 \*\*\* |
|  | (0.00) | (0.01) |
| `Narrowness of identity group:`0.59 | 0.09 \*\*\* | 0.01 |
|  | (0.00) | (0.01) |
| `Identity:`Identity used | 0.19 \*\*\* | 0.19 \*\*\* |
|  | (0.00) | (0.00) |
| `Acceptability of different opinion:`0.22 | -0.01 \*\*\* | -0.01 |
|  | (0.00) | (0.01) |
| `Acceptability of different opinion:`0.28 | -0.02 \*\*\* | -0.04 \*\*\* |
|  | (0.00) | (0.01) |
| `Acceptability of different opinion:`0.34 | -0.03 \*\*\* | -0.07 \*\*\* |
|  | (0.00) | (0.01) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.22 |  | -0.10 \*\*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.22 |  | 0.04 \*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.28 |  | -0.08 \*\*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.28 |  | 0.08 \*\*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.34 |  | -0.05 \*\*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.34 |  | 0.13 \*\*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.49:`Opinion dimensions:`2 |  | 0.04 \*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.59:`Opinion dimensions:`2 |  | -0.01 |
|  |  | (0.01) |
| `Narrowness of identity group:`0.49:`Opinion dimensions:`4 |  | 0.04 \*\*\* |
|  |  | (0.01) |
| `Narrowness of identity group:`0.59:`Opinion dimensions:`4 |  | -0.00 |
|  |  | (0.01) |
| `Acceptability of different opinion:`0.22:`Opinion dimensions:`2 |  | -0.05 \*\*\* |
|  |  | (0.01) |
| `Acceptability of different opinion:`0.28:`Opinion dimensions:`2 |  | -0.06 \*\*\* |
|  |  | (0.01) |
| `Acceptability of different opinion:`0.34:`Opinion dimensions:`2 |  | -0.04 \*\*\* |
|  |  | (0.01) |
| `Acceptability of different opinion:`0.22:`Opinion dimensions:`4 |  | -0.03 \*\* |
|  |  | (0.01) |
| `Acceptability of different opinion:`0.28:`Opinion dimensions:`4 |  | -0.01 |
|  |  | (0.01) |
| `Acceptability of different opinion:`0.34:`Opinion dimensions:`4 |  | 0.02 |
|  |  | (0.01) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.22:`Opinion dimensions:`2 |  | 0.15 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.22:`Opinion dimensions:`2 |  | 0.09 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.28:`Opinion dimensions:`2 |  | 0.17 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.28:`Opinion dimensions:`2 |  | 0.10 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.34:`Opinion dimensions:`2 |  | 0.15 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.34:`Opinion dimensions:`2 |  | 0.06 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.22:`Opinion dimensions:`4 |  | 0.15 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.22:`Opinion dimensions:`4 |  | 0.06 \*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.28:`Opinion dimensions:`4 |  | 0.12 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.28:`Opinion dimensions:`4 |  | 0.03 |
|  |  | (0.02) |
| `Narrowness of identity group:`0.49:`Acceptability of different opinion:`0.34:`Opinion dimensions:`4 |  | 0.09 \*\*\* |
|  |  | (0.02) |
| `Narrowness of identity group:`0.59:`Acceptability of different opinion:`0.34:`Opinion dimensions:`4 |  | -0.01 |
|  |  | (0.02) |
| N | 14400 | 14400 |
| R2 | 0.54 | 0.60 |
| \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05. | | |

The improvement by key variables is evident, BIC clearly show this, BIC difference of full model and model only with selected variables is 25.0493931. Also BIC shows that 3-way interaction improves the model significantly: -1681.4443037.

### NOTE:

We renamed variables and I also factorized in principle numerical variables that are using only 3 or 4 values. Here is table with old names, new names and minimum value (contrast):

| Old name | New name | Minimum or contrast (old value/ new value) |
| --- | --- | --- |
| opinions | Opinion dimensions | 1 |
| boundary | Acceptability of different opinion | 0.1 |
| use\_identity? | Identity | FALSE/ Identity not used |
| id\_threshold | Narrowness of identity group | 0.39 |

I hope it help in interpreting results.

Also note that we stopped using old variable mode, since we can naturally interprept only mode openly-listen, but we still have problem to naturally explain and work with mode vaguely-speak, so we stick with the only one mode we know how to work with.

Now, again, just for sure, we compare BIC of medels:

paste("BIC comparison of full vs. selected variables model:", round(BIC(ef) - BIC(eu), 1))

## [1] "BIC comparison of full vs. selected variables model: 25"

paste("BIC comparison of main effects vs. interactions model:", round(BIC(eui) - BIC(eu), 1))

## [1] "BIC comparison of main effects vs. interactions model: -1681.4"

## Graph

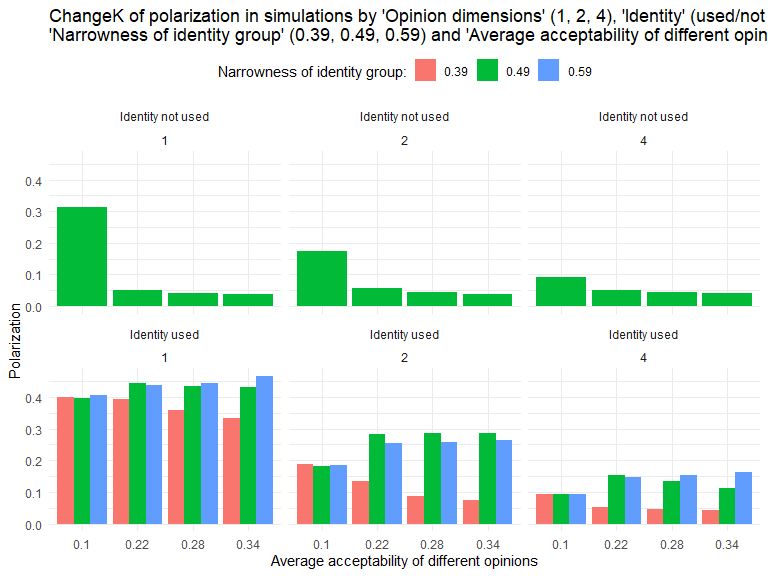
Now, let’s show our results graphically!

### Drawing graphs

dfg = df %>%   
 group\_by(`Opinion dimensions:`, `Identity:`, `Narrowness of identity group:`, `Acceptability of different opinion:`) %>%  
 summarise(ESBG = mean(ESBG)) %>% ungroup() %>% group\_by(`Opinion dimensions:`, `Identity:`) %>%   
 arrange(`Opinion dimensions:`, `Identity:`, `Narrowness of identity group:`, `Acceptability of different opinion:`)

## `summarise()` has grouped output by 'Opinion dimensions:', 'Identity:', 'Narrowness of identity group:'. You can override using the `.groups` argument.

ggplot(dfg, aes(x = `Acceptability of different opinion:`, y = ESBG,   
 fill = `Narrowness of identity group:`)) +  
 facet\_wrap(vars(`Identity:`, `Opinion dimensions:`), ncol=3) +  
 geom\_col(position = position\_dodge()) +  
 labs(title = "ChangeK of polarization in simulations by 'Opinion dimensions' (1, 2, 4), 'Identity' (used/not used), \n'Narrowness of identity group' (0.39, 0.49, 0.59) and 'Average acceptability of different opinions' (0.1, 0.22, 0.28, 0.34)",  
 x = "Average acceptability of different opinions", y = "Polarization") +  
 theme\_minimal() +  
 theme(legend.position = "top")



ggplot(dfg, aes(x = `Acceptability of different opinion:`, y = ESBG,   
 col = `Narrowness of identity group:`, group = `Narrowness of identity group:`)) +  
 facet\_wrap(vars(`Identity:`, `Opinion dimensions:`), ncol=3) +  
 geom\_point() + geom\_line() +  
 labs(title = "Change of polarization in simulations by 'Opinion dimensions' (1, 2, 4), 'Identity' (used/not used), \n'Narrowness of identity group' (0.39, 0.49, 0.59) and 'Average acceptability of different opinions' (0.1, 0.22, 0.28, 0.34)",  
 x = "Average acceptability of different opinions", y = "Polarization") +  
 theme\_minimal() +  
 theme(legend.position = "top")

