# Ideology Correlations for Moral/Emotional Contagion

*Julian Wills and Billy Brady*

*November 5th, 2015*

## Correlations, Linear Models, and Unpacking Interactions

### Variable Guide:

**tID = ideology of original tweet author (negative = liberal; 0 = centrist; postive = conservative)**

**rtID = ideology of retweeter (in most cases, averaged per tweet)**

**ex.tID = ideological extremity (i.e. absolute value) of author (0 = centrist)**

**ex.rtID = ideological extremity (i.e. absolute value) of retweeter**

**diffID = ideological difference (absolute value) between author and (avg.) retweeter (0 = no difference)**

**M = moral (1); nonmoral (-1)**

**E = emotional (1); unemotional (-1)**

**cond = Four Conditions: M_E, M_NE, NM_E, NM_NE**

---

### Output of Models

```
# after adjusting for ideology of author, retweeters, and ideological extremity
lm(diffID ~ tID+rtID*M+rtID*E+M*E,data=dCjoin2) %>% summary() #main model
```

```
##
## Call:
## lm(formula = diffID ~ tID + rtID * M + rtID * E + M * E, data = dCjoin2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2404 -0.3409 -0.1208  0.2204  3.0307
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)  0.692340   0.005675 121.991     < 2e-16 ***
## tID          0.131735   0.006440  20.455     < 2e-16 ***
## rtID         0.104043   0.007365  14.127     < 2e-16 ***
## M           -0.056783   0.005579 -10.178     < 2e-16 ***
```

```
## E            0.011849   0.005566   2.129      0.0333 *
## rtID:M      -0.032949   0.006611  -4.984 0.000000631 ***
## rtID:E       0.025531   0.006180   4.131 0.000036338 ***
## M:E         -0.007284   0.005171  -1.409      0.1590
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5188 on 13216 degrees of freedom
## Multiple R-squared:  0.09611,    Adjusted R-squared:  0.09564
## F-statistic: 200.8 on 7 and 13216 DF,  p-value: < 2.2e-16
```

```r
# morality plays a bigger role than emotion when it comes to ideological diversity
lm(diffID ~ tID+rtID+ex.tID*M+ex.tID*E,data=dCjoin2) %>% summary() #emo int w/ ex.tID
```

```
##
## Call:
## lm(formula = diffID ~ tID + rtID + ex.tID * M + ex.tID * E, data = dCjoin2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5486 -0.3405 -0.1207  0.2304  2.9905
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.529326   0.009561  55.363  < 2e-16 ***
## tID          0.151042   0.006458  23.387  < 2e-16 ***
## rtID         0.091655   0.006630  13.824  < 2e-16 ***
## ex.tID       0.228819   0.011088  20.637  < 2e-16 ***
## M            0.029591   0.009420   3.141  0.00169 **
## E            0.014241   0.008978   1.586  0.11271
## ex.tID:M    -0.111851   0.010833 -10.325  < 2e-16 ***
## ex.tID:E    -0.023883   0.010206  -2.340  0.01929 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5113 on 13216 degrees of freedom
## Multiple R-squared:  0.1222, Adjusted R-squared:  0.1217
## F-statistic: 262.8 on 7 and 13216 DF,  p-value: < 2.2e-16
```

```r
lm(diffID ~ tID+rtID+rtID*M+rtID*E,data=dCjoin2) %>% summary() #emo int w/ rtID
```

```
##
## Call:
## lm(formula = diffID ~ tID + rtID + rtID * M + rtID * E, data = dCjoin2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2311 -0.3410 -0.1213  0.2206  3.0262
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)  0.691621   0.005653 122.356     < 2e-16 ***
## tID          0.131771   0.006440  20.460     < 2e-16 ***
```

```
## rtID         0.103877   0.007364  14.106     < 2e-16 ***
## M           -0.057722   0.005539 -10.420     < 2e-16 ***
## E            0.009333   0.005272   1.770      0.0767 .
## rtID:M      -0.032671   0.006609  -4.944 0.000000776 ***
## rtID:E       0.026110   0.006167   4.234 0.000023113 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5188 on 13217 degrees of freedom
## Multiple R-squared:  0.09598,    Adjusted R-squared:  0.09557
## F-statistic: 233.9 on 6 and 13217 DF,  p-value: < 2.2e-16
```

```r
lm(diffID ~ tID+rtID+tID*M+tID*E,data=dCjoin2) %>% summary()
```

```
##
## Call:
## lm(formula = diffID ~ tID + rtID + tID * M + tID * E, data = dCjoin2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1279 -0.3407 -0.1206  0.2203  3.0207
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.689315   0.005652 121.966  < 2e-16 ***
## tID          0.127240   0.006973  18.247  < 2e-16 ***
## rtID         0.098109   0.006763  14.507  < 2e-16 ***
## M           -0.044816   0.005470  -8.194 2.76e-16 ***
## E           -0.001973   0.005166  -0.382    0.703
## tID:M        0.008567   0.006333   1.353    0.176
## tID:E       -0.006908   0.005871  -1.177    0.239
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5195 on 13217 degrees of freedom
## Multiple R-squared:  0.09354,    Adjusted R-squared:  0.09313
## F-statistic: 227.3 on 6 and 13217 DF,  p-value: < 2.2e-16
```

```r
lm(diffID ~ ex.tID+ex.rtID*M+ex.rtID*E+M*E,data=dCjoin2) %>% summary() #huge morality interaction
```

```
##
## Call:
## lm(formula = diffID ~ ex.tID + ex.rtID * M + ex.rtID * E + M *
##     E, data = dCjoin2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9705 -0.3885 -0.1452  0.2472  2.9761
##
## Coefficients:
##              Estimate Std. Error t value         Pr(>|t|)
## (Intercept)  0.407822   0.013431  30.364          < 2e-16 ***
## ex.tID       0.094516   0.010414   9.076          < 2e-16 ***
```

```
## ex.rtID      0.171401   0.013705  12.507          < 2e-16 ***
## M            0.009813   0.012009   0.817            0.414
## E           -0.002869   0.011720  -0.245            0.807
## ex.rtID:M   -0.095029   0.013332  -7.128 0.00000000000107 ***
## ex.rtID:E   -0.015480   0.012877  -1.202            0.229
## M:E         -0.006337   0.005360  -1.182            0.237
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5366 on 13216 degrees of freedom
## Multiple R-squared:  0.03291,    Adjusted R-squared:  0.0324
## F-statistic: 64.26 on 7 and 13216 DF,  p-value: < 2.2e-16
```

```
lm(diffID ~ ex.tID*ex.rtID*M+ex.rtID*E+M*E,data=dCjoin2) %>% summary() #morality 3-way interaction
```

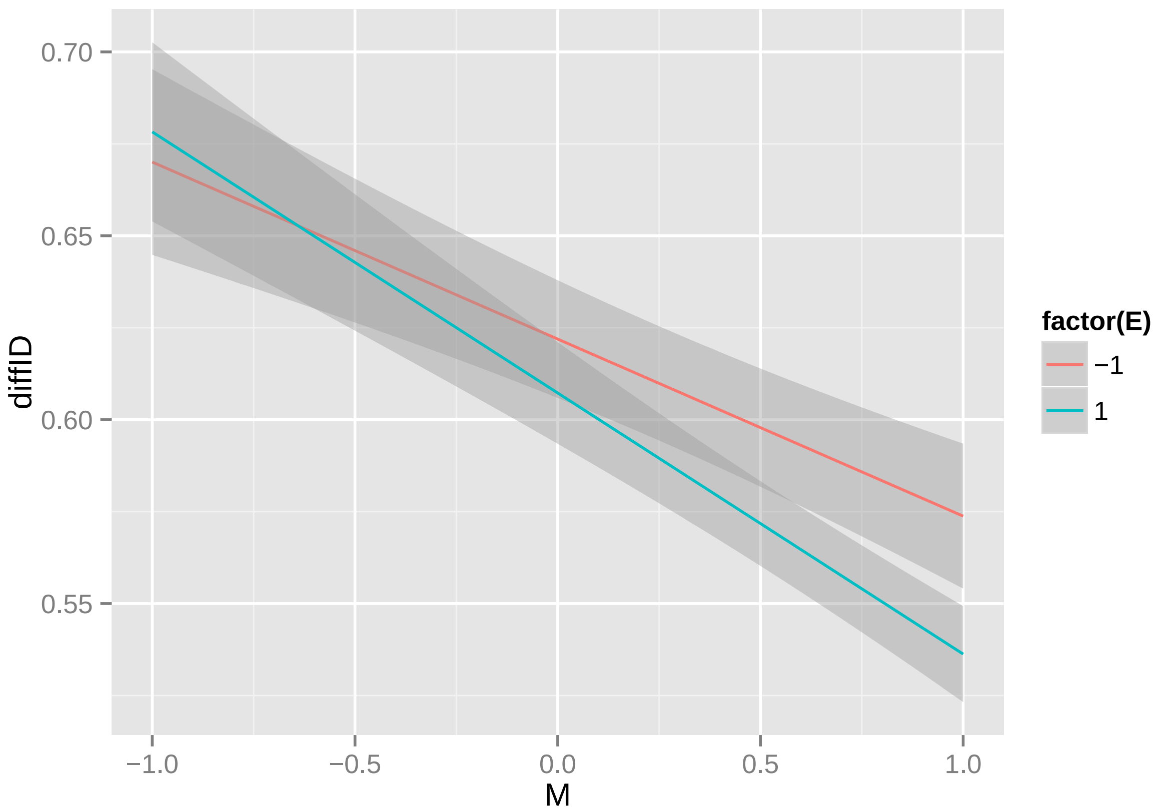
```
##
## Call:
## lm(formula = diffID ~ ex.tID * ex.rtID * M + ex.rtID * E + M *
##     E, data = dCjoin2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9175 -0.3457 -0.0928  0.2280  4.0485
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -0.233553   0.019646 -11.888   <2e-16 ***
## ex.tID             1.004703   0.023686  42.418   <2e-16 ***
## ex.rtID            0.950614   0.023169  41.029   <2e-16 ***
## M                  0.007826   0.019666   0.398   0.6907
## E                 -0.027112   0.010922  -2.482   0.0131 *
## ex.tID:ex.rtID    -1.045816   0.025842 -40.470   <2e-16 ***
## ex.tID:M          -0.037193   0.023663  -1.572   0.1160
## ex.rtID:M         -0.016212   0.023174  -0.700   0.4842
## ex.rtID:E          0.016855   0.012009   1.404   0.1605
## M:E               -0.006993   0.004995  -1.400   0.1615
## ex.tID:ex.rtID:M -0.058078   0.025799  -2.251   0.0244 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4994 on 13213 degrees of freedom
## Multiple R-squared:  0.1625, Adjusted R-squared:  0.1619
## F-statistic: 256.4 on 10 and 13213 DF,  p-value: < 2.2e-16
```

---

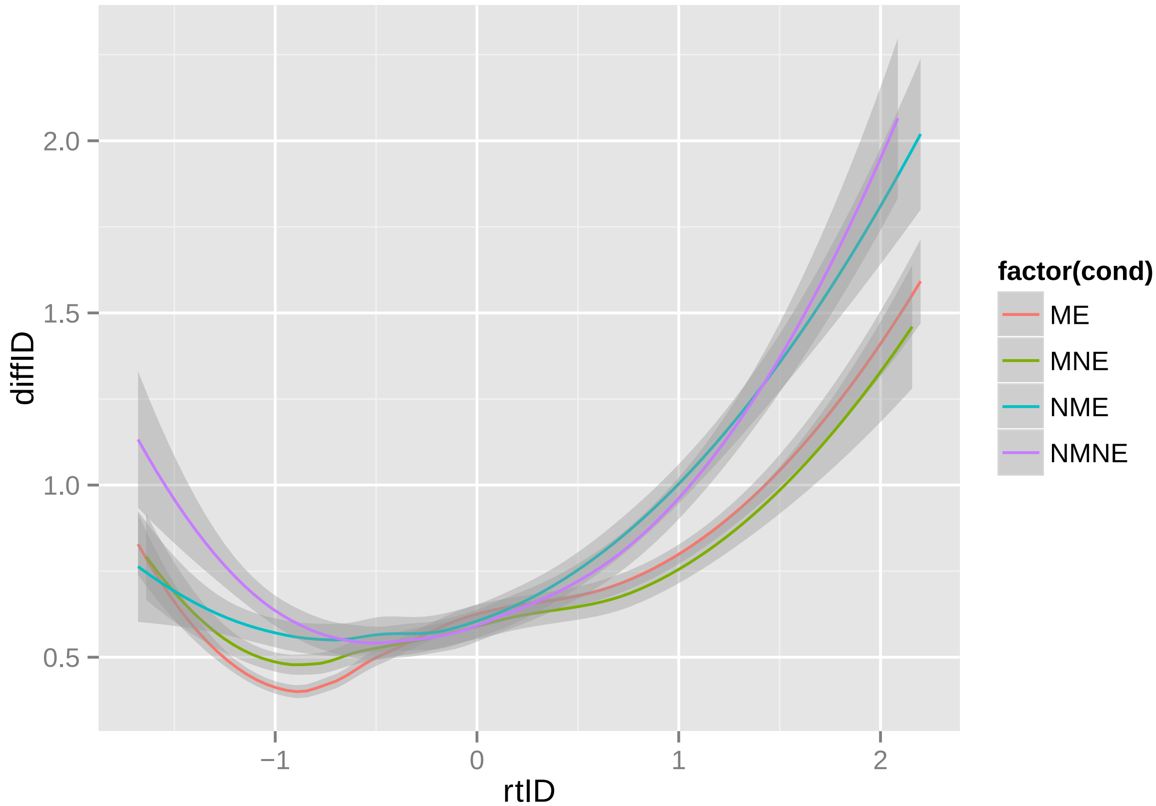
# Plots:

### In a moral context, emotional tweets accrue even narrower support

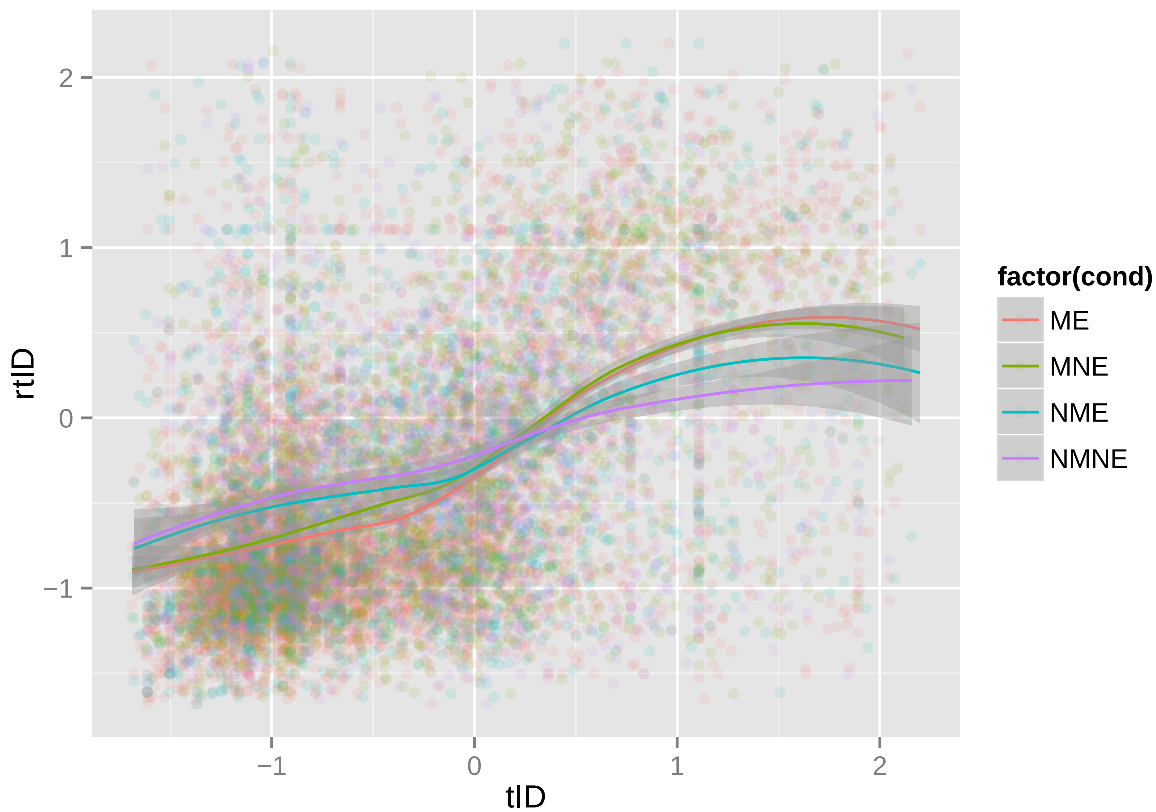
```r
ggplot(dCjoin2,aes(x=M,y=diffID,color=factor(E))) +
  geom_smooth(method="lm")
```




```r
ggplot(dCjoin2,aes(x=rtID,y=diffID,color=factor(cond))) +
  geom_smooth(method="loess")
```



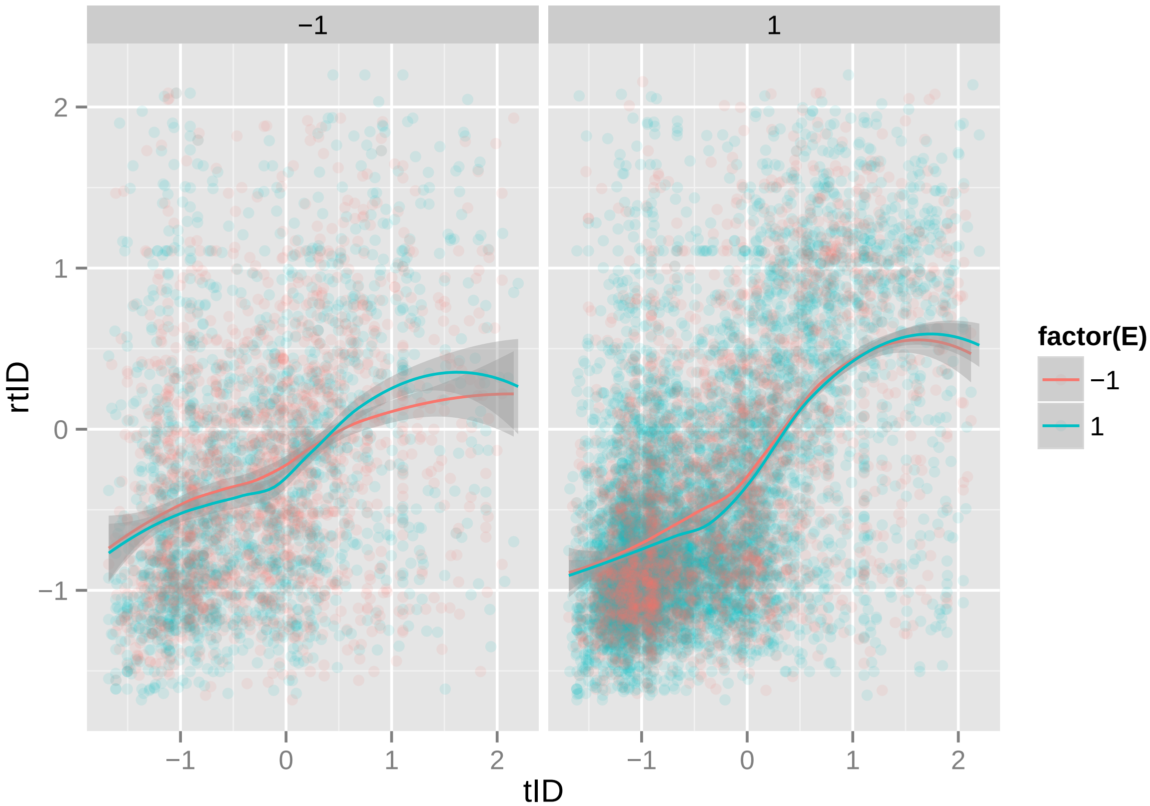

```
ggplot(dCjoin2,aes(x=tID,y=rtID,color=factor(cond))) +
  geom_point(alpha=.1) +
  geom_smooth(method="loess")
```



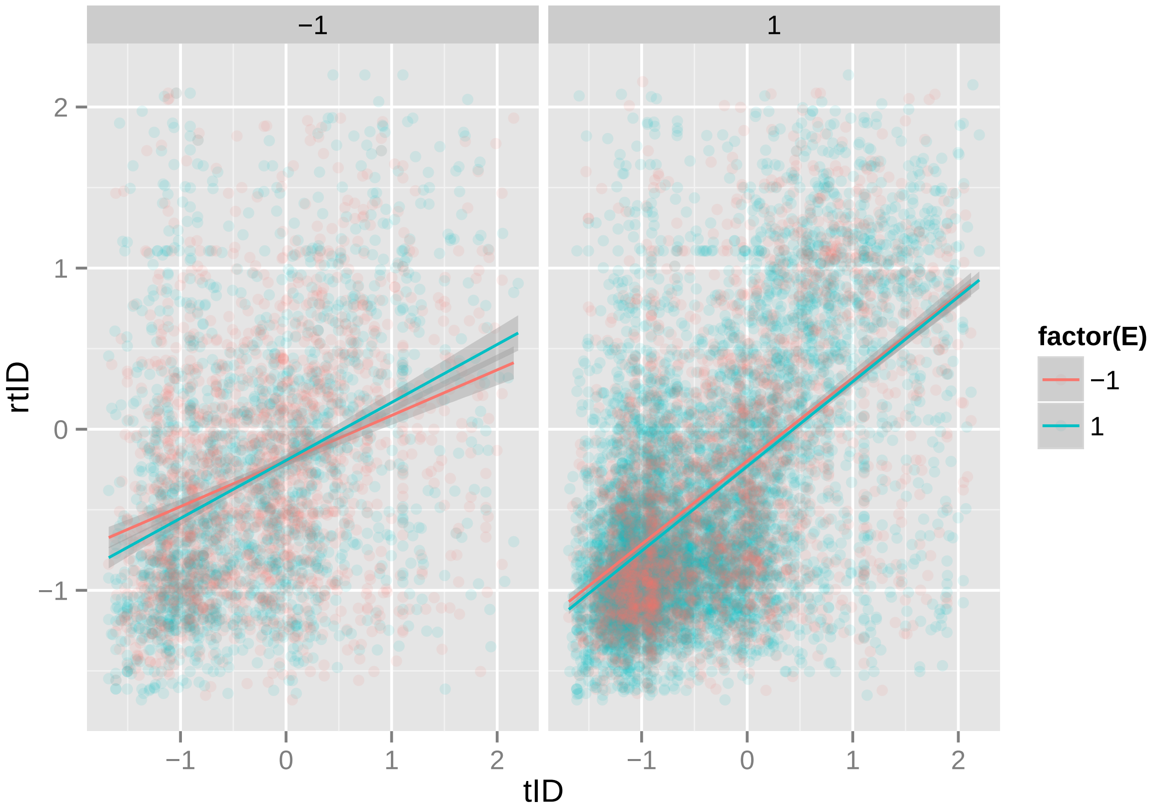

---

### Morality drives ideological clustering more than emotion (though emotion helps in nonmoral context)

```
ggplot(dCjoin2,aes(x=tID,y=rtID,color=factor(E))) +
  geom_point(alpha=.1) +
  geom_smooth(method="loess") +
  facet_grid(~M)
```



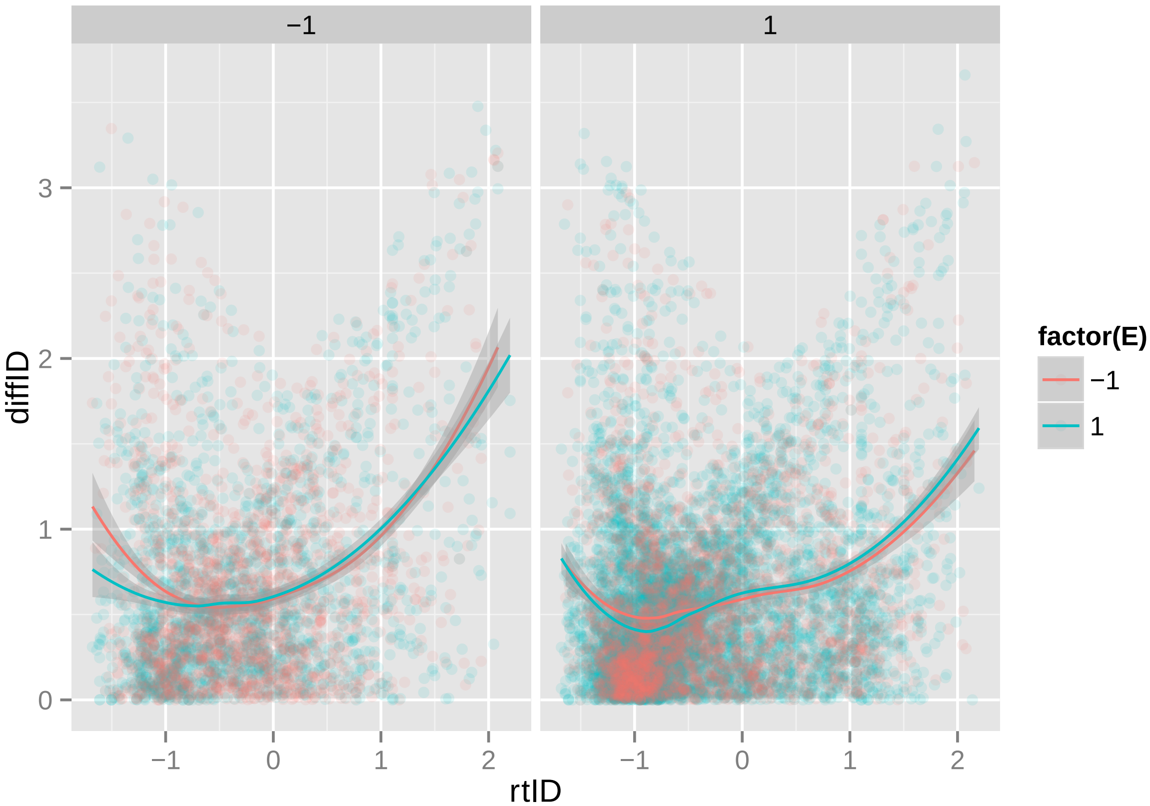

```
ggplot(dCjoin2,aes(x=tID,y=rtID,color=factor(E))) +
  geom_point(alpha=.1) +
  geom_smooth(method="lm") +
  facet_grid(~M)
```




---

**Narrower clusters for moral tweets, especially for liberals when they are emotional**

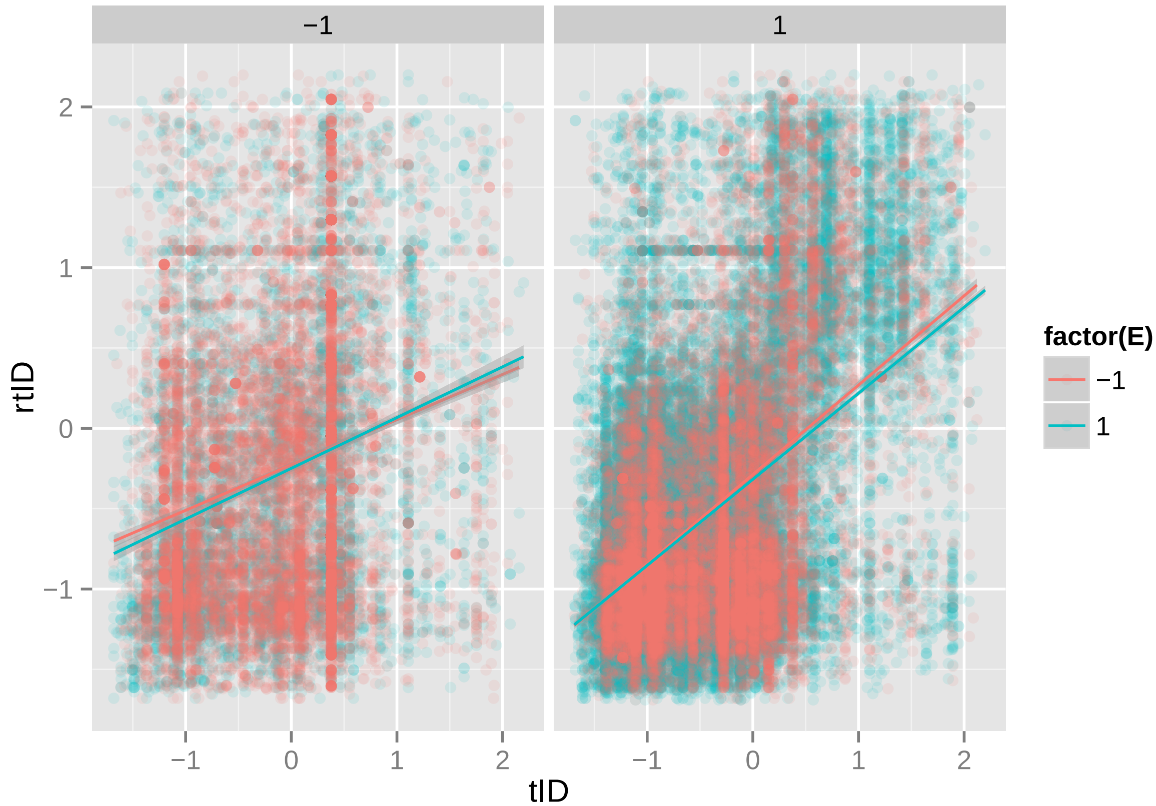
```
ggplot(dCjoin2,aes(x=rtID,y=diffID,color=factor(E))) +
  geom_point(alpha=.1) +
  geom_smooth(method="loess") +
  facet_grid(~M)
```




---

**On raw data, before clustering retweeters w/ original author. Vert/Horz lines reflect active users?**

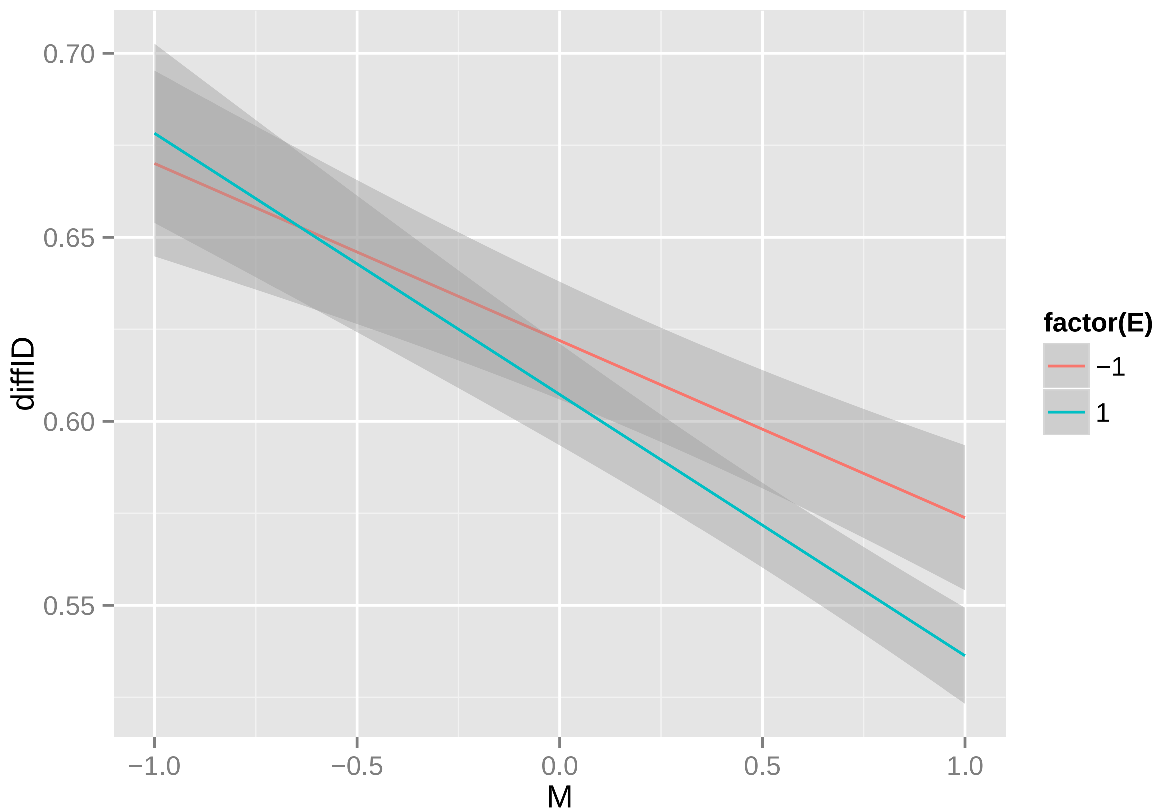
```
ggplot(dCjoin,aes(x=tID,y=rtID,color=factor(E))) +
  geom_point(alpha=.1) +
  geom_smooth(method="lm") +
  facet_grid(~M)
```




***

### In a moral context, emotional tweets accrue even narrower support

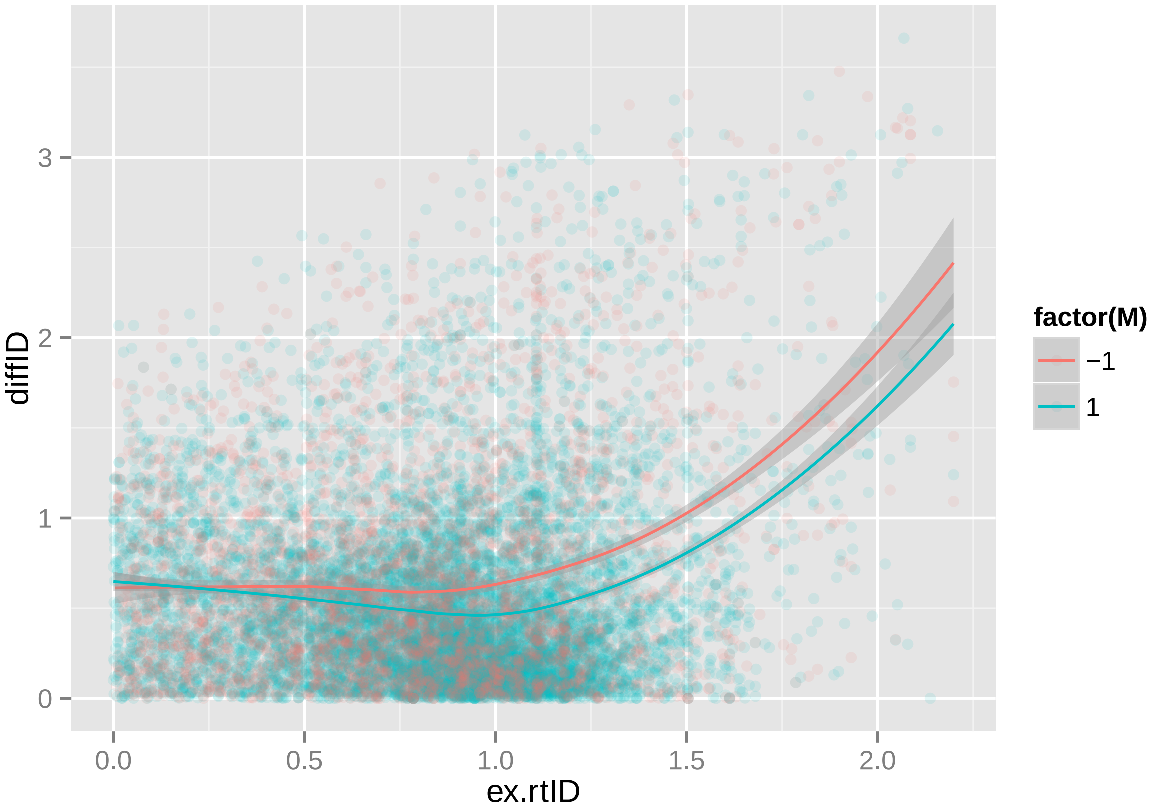
```
ggplot(dCjoin2,aes(x=M,y=diffID,color=factor(E))) +
  geom_smooth(method="lm")
```



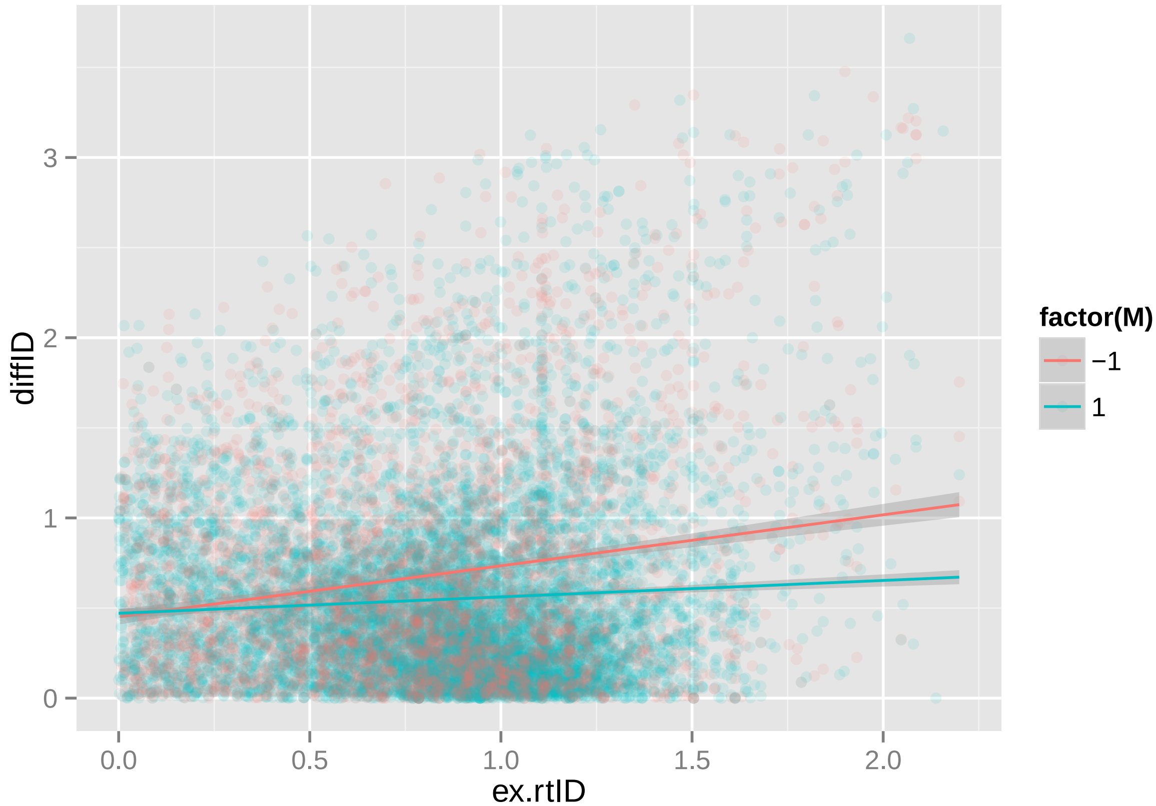

---

### Political extremists retweet same ideology if tweet has moral content

```
ggplot(dCjoin2,aes(x=ex.rtID,y=diffID,color=factor(M))) +
  geom_point(alpha=.1) +
  geom_smooth(method="loess")
```

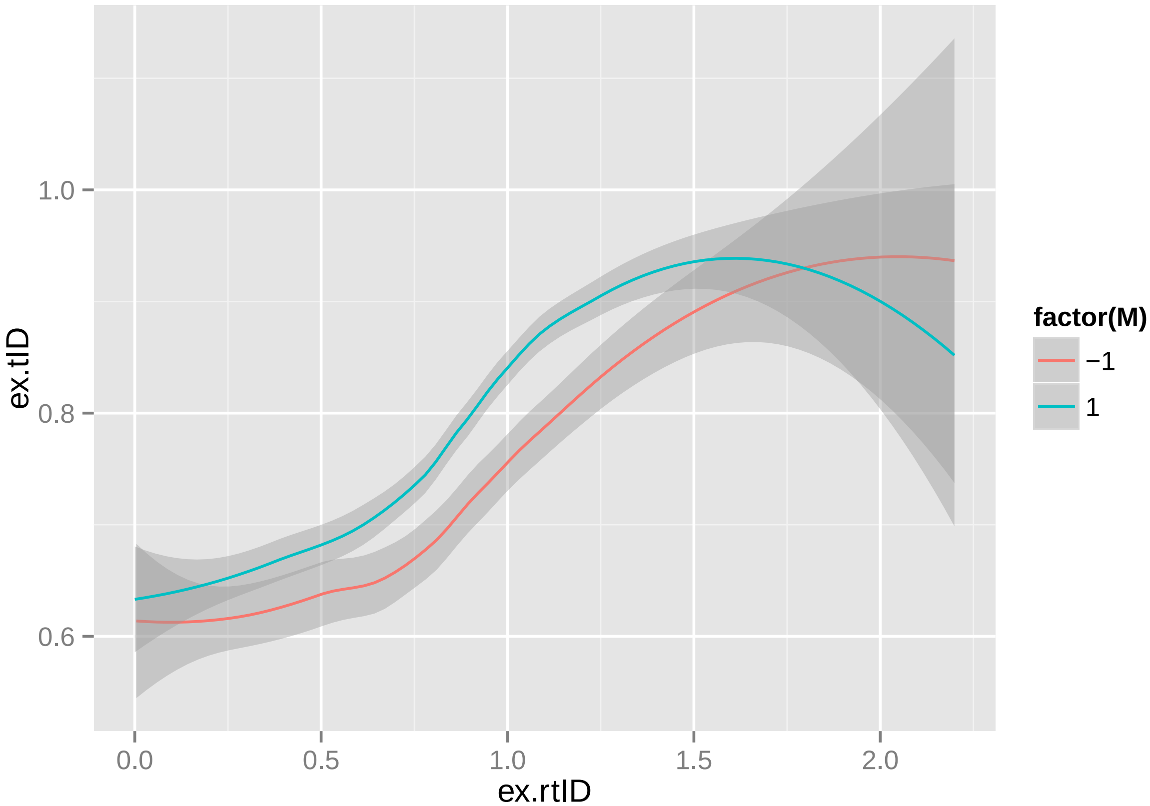



```
ggplot(dCjoin2,aes(x=ex.rtID,y=diffID,color=factor(M))) +
  geom_point(alpha=.1) +
  geom_smooth(method="lm")
```

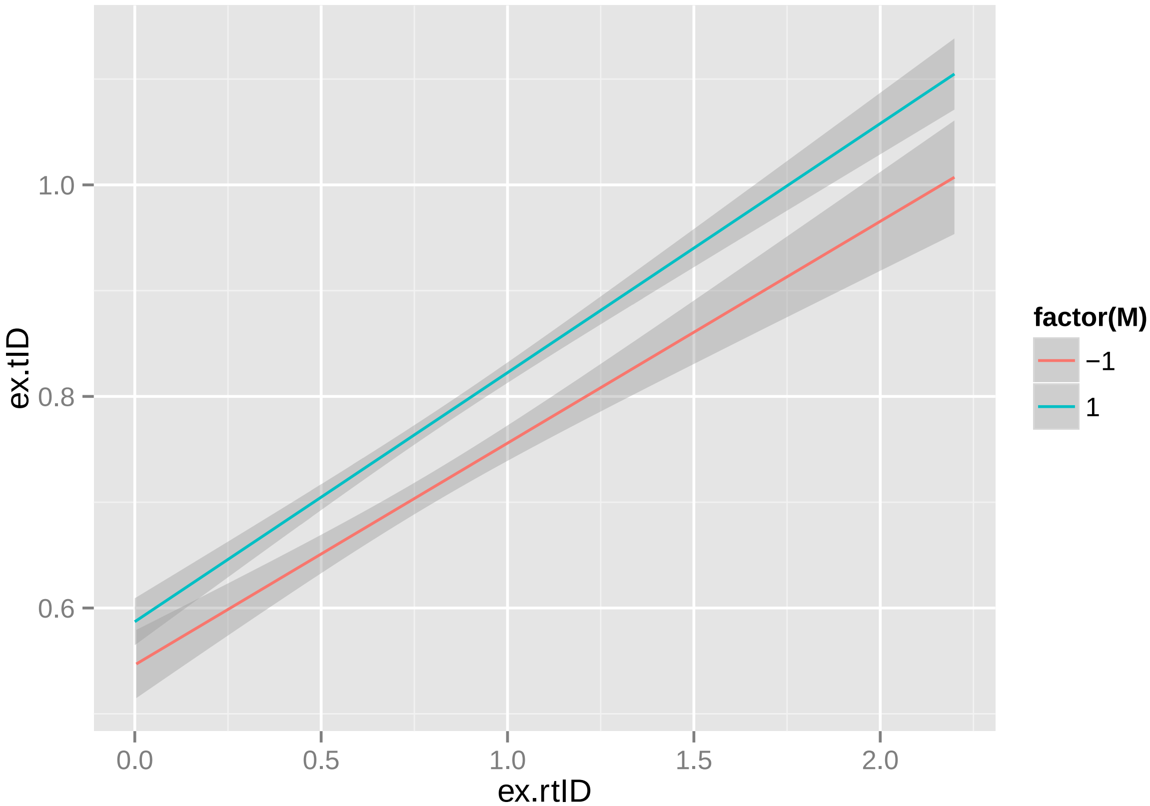



### More extremism (author and retweeter) when content is moral

```
ggplot(dCjoin2,aes(x=ex.rtID,y=ex.tID,color=factor(M))) +
  geom_smooth(method="loess")
```




```
ggplot(dCjoin2,aes(x=ex.rtID,y=ex.tID,color=factor(M))) +
  geom_smooth(method="lm")
```




## Code for heatmaps

```
# # Heat maps -----------------------------------------------------------------
#
# ## functions to construct heatmaps
# min <- -3.5
# max <- 3.5
# breaks <- 0.25
#
# expand_data <- function(df, breaks=0.10, min=-4, max=4){
#   x <- df$rtid %>% as.numeric()
#   y <- df$tid %>% as.numeric()
#   x <- (round((x - min) / breaks, 0) * breaks) + min
#   y <- (round((y - min) / breaks, 0) * breaks) + min
#   tab <- table(x, y)
#   tab <- melt(tab)
#   tab$prop <- tab$value/sum(tab$value)
#   return(tab)
# }
#
# ideoHeatMap <- function(df) {
#   new.xy.me <- expand_data(df %>% filter(cond=="ME"),breaks=0.25) %>%  mutate(cond="ME")
#   new.xy.nme <- expand_data(df %>% filter(cond=="NME"),breaks=0.25) %>% mutate(cond="NME")
```

```r
#   new.xy.mne <- expand_data(df %>% filter(cond=="MNE"),breaks=0.25) %>% mutate(cond="MNE")
#   new.xy.nmne <- expand_data(df %>% filter(cond=="NMNE"),breaks=0.25) %>%  mutate(cond="NMNE")
#   return (rbind(new.xy.me,new.xy.nme,new.xy.mne,new.xy.nmne))
# }
```

## Ideology plots