W241Final

```
#setup and load data; remove first row
library(data.table)
library(magrittr)
library(stargazer)
```

```
##
## Please cite as:
```

```
##  Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2. http://CRAN.R-project.org/package=stargazer
```

```
library(likert)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: xtable
```

```
setwd("~/MIDS/W241")
d = data.table(read.csv("MIDS241Actual.csv"))

#fix column names and remove first row
names(d)[1] = 'V1'
names(d)[1:10] = sapply(d[1, .(V1, V2, V3, V4, V5, V6, V7, V8, V9, V10)], as.character)
d = d[-1]

#create treatment variable (1 = treat; 0 = control)
d[,treat:=(as.numeric(c.bc==''))]
#based on treat, create pre.conf variable
d[,pre.conf:=(treat*as.numeric(t.bc)+(1-treat)*as.numeric(c.bc)-1)]

#sanity check
#makes sure everyone answered both questions; and answer them only once
sum(d$q1=='')+sum(d$q2=='') #should be 0
```

```
## [1] 86
```

```
(sum(d$q1!='')+sum(d$q2!=''))/2 #should be 1046
```

```
## [1] 1118
```

```
(sum(d$c.conf1=='')+sum(d$c.conf2=='0'))/2 #should be number treated = 525
```

```
## [1] 603
```

```
(sum(d$tp.conf=='')+sum(d$tn.conf==''))/2 #should be number control = 521
```

```
## [1] 602.5
```

```
#convert variables into numeric
d[,q1:=as.numeric(as.character(q1))]
d[,q2:=as.numeric(as.character(q2))]
d[,c.conf1:=as.numeric(as.character(c.conf1))]
d[,c.conf2:=as.numeric(as.character(c.conf2))]
d[,tp.conf:=as.numeric(as.character(tp.conf))]
d[,tn.conf:=as.numeric(as.character(tn.conf))]
d[,q1.time_2:=as.numeric(as.character(q1.time_2))]
d[,q2.time_2:=as.numeric(as.character(q2.time_2))]
d[,c.conf1.time_2:=as.numeric(as.character(c.conf1.time_2))]
d[,c.conf2.time_2:=as.numeric(as.character(c.conf2.time_2))]
d[,tp.conf.time_2:=as.numeric(as.character(tp.conf.time_2))]
d[,tn.conf.time_2:=as.numeric(as.character(tn.conf.time_2))]

#treat as attrition if no validation code(vc)
d[vc=='', attr:= 1]
d[vc!='', attr:= 0]

#also treat samples that answered questions too quickly (or have NA for timer) as attrition
d[is.na(d$q1.time_2)]$attr = 1
d[is.na(d$q2.time_2)]$attr = 1
d[q1.time_2 <= 3 | q2.time_2 <= 3, attr := 1]

#test for differential attrition
chisq.test(table(d$gender, d$attr)[-3,])
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table(d$gender, d$attr)[-3, ]
## X-squared = 2.6426, df = 1, p-value = 0.104
```

```
chisq.test(table(d$edu, d$attr)[2:8,]) #some cells are too small for this test, but fisher test exceeds
```

```
## Warning in chisq.test(table(d$edu, d$attr)[2:8, ]): Chi-squared
## approximation may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(d$edu, d$attr)[2:8, ]
## X-squared = 16.115, df = 6, p-value = 0.01315
```

```
fisher.test(table(d$race, d$attr)[2:8,])
```

```
##
```

```
##  Fisher's Exact Test for Count Data
##
## data:  table(d$race, d$attr)[2:8, ]
## p-value = 0.02583
## alternative hypothesis: two.sided
```

```
summary(lm(attr~gender, d))
```

```
##
## Call:
## lm(formula = attr ~ gender, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1431 -0.1431 -0.1091 -0.1091  0.8909
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.14307    0.01304  10.969   <2e-16 ***
## gender2     -0.03395    0.01980  -1.715   0.0866 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3343 on 1159 degrees of freedom
## Multiple R-squared:  0.002531,   Adjusted R-squared:  0.00167
## F-statistic: 2.941 on 1 and 1159 DF,  p-value: 0.08665
```

```
summary(lm(attr~edu, d))
```

```
##
## Call:
## lm(formula = attr ~ edu, data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33333 -0.14194 -0.10000 -0.08824  0.91176
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.0000     0.3323   3.009  0.00268 **
## edu1         -0.6667     0.3589  -1.857  0.06351 .
## edu2         -0.8242     0.3341  -2.467  0.01378 *
## edu3         -0.8581     0.3328  -2.578  0.01006 *
## edu4         -0.8125     0.3338  -2.434  0.01508 *
## edu5         -0.9000     0.3327  -2.705  0.00692 **
## edu6         -0.9118     0.3333  -2.736  0.00632 **
## edu7         -0.7619     0.3401  -2.240  0.02528 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3323 on 1153 degrees of freedom
## Multiple R-squared:  0.01967,    Adjusted R-squared:  0.01371
## F-statistic: 3.304 on 7 and 1153 DF,  p-value: 0.001739
```

```r
summary(lm(attr~race, d))
```

```
##
## Call:
## lm(formula = attr ~ race, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.2222 -0.1201 -0.1201 -0.1201  0.9362
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.0000     0.3327   3.006  0.00271 **
## race1        -0.8800     0.3329  -2.643  0.00832 **
## race2        -0.8125     0.3348  -2.427  0.01537 *
## race3        -0.7778     0.3507  -2.218  0.02676 *
## race4        -0.7867     0.3349  -2.349  0.01900 *
## race5        -0.9362     0.3345  -2.799  0.00521 **
## race6        -1.0000     0.3720  -2.688  0.00728 **
## race7        -0.8000     0.3436  -2.328  0.02007 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3327 on 1153 degrees of freedom
## Multiple R-squared:  0.01737,    Adjusted R-squared:  0.0114
## F-statistic: 2.911 on 7 and 1153 DF,  p-value: 0.005064
```

```r
#remove attrition samples for subsequent analysis
d.attr = d[attr ==1]
d = d[attr == 0,]

d[, qrandomc := (as.numeric(as.character(qrandomc)))]
d[, qrandomt := (as.numeric(as.character(qrandomt)))]
d[, randompn := (as.numeric(as.character(randompn)))]
#qrandomc: 1 = q1 first; 0 = q2 first
#qrandomt: 1 = q1 first; 0 = q2 first
#randompn: 1 = positive first; 0 = negative first

#computes within subject effects (effect between positive and negative reinforcement; no control group
#uses randomization inference to estimate p-value
dwith = d[treat==1, .(ate=tp.conf-tn.conf)]
mean(dwith$ate)
```

```
## [1] 0.65286
```

```r
conf.ri <- function(){
  x = sample(c(rep(-1, 246), rep(1, 261)))
  return(mean((dwith*x)$ate) > mean(dwith$ate))
}

mean(replicate (10000, conf.ri()))
```

```
## [1] 0
```

```
#reconstruct dataframe
cq1 = d[treat==0,.(gender, age, edu, race, answer=q1, qtime=q1.time_2, ptreat=0, ntreat=0, first=qrandom
          q1=1, pre.conf, conf=ifelse(qrandomc, c.conf1, c.conf2),  cftime=ifelse(qrandomc, c.conf1.tim
cq1[, crt:=(answer==1)]
cq2 = d[treat==0,.(gender, age, edu, race, answer=q2, qtime=q2.time_2, ptreat=0, ntreat=0, first=(1-qran
          q1=0, pre.conf, conf=ifelse((1-qrandomc), c.conf1, c.conf2),  cftime=ifelse((1-qrandomc), c.
cq2[, crt:=(answer==3)]

tq1a = d[treat==1&qrandomt==1,.(gender, age, edu, race, answer=q1, qtime=q1.time_2, ptreat=randompn, nt
                  conf=ifelse(randompn, tp.conf, tn.conf),  cftime=ifelse(randompn, tp.conf.time_2, tn
tq1a[, crt:=(answer==1)]
tq1b = d[treat==1&qrandomt==0,.(gender, age, edu, race, answer=q1, qtime=q1.time_2, ptreat=(1-randompn)
                  conf=ifelse(randompn, tn.conf, tp.conf),  cftime=ifelse(randompn, tn.conf.time_2, tp
tq1b[, crt:=(answer==1)]

tq2a = d[treat==1&qrandomt==0,.(gender, age, edu, race, answer=q2, qtime=q2.time_2, ptreat=randompn, nt
                  conf=ifelse(randompn, tp.conf, tn.conf),  cftime=ifelse(randompn, tp.conf.time_2, tn
tq2a[, crt:=(answer==3)]
tq2b = d[treat==1&qrandomt==1,.(gender, age, edu, race, answer=q2, qtime=q2.time_2, ptreat=(1-randompn)
                  conf=ifelse(randompn, tn.conf, tp.conf),  cftime=ifelse(randompn, tn.conf.time_2, tp
tq2b[, crt:=(answer==3)]

#df contains basic info: gender, age, edu, race
#q1 indicates whether the questions is q1
#ptreat, ntreat indicates positive and negative reinforcement
#first indicates whether the question is presented first
#answer = selection for the question; qtime is time used
#crt is whether the answer is correct
#conf is confidence level selected for question; cftime is time used
df=rbind(cq1, cq2, tq1a, tq1b, tq2b, tq2a)

#check whether two questions are significantly different in confidence level (singals difficulty differ
summary(lm(conf~q1, df[ptreat==0&ntreat==0]))
```

```
##
## Call:
## lm(formula = conf ~ q1, data = df[ptreat == 0 & ntreat == 0])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8713 -1.7505  0.2495  1.2495  2.2495
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.87129    0.06970  41.193   <2e-16 ***
## q1          -0.12079    0.09858  -1.225    0.221
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.566 on 1008 degrees of freedom
## Multiple R-squared:  0.001487,   Adjusted R-squared:  0.0004968
## F-statistic: 1.502 on 1 and 1008 DF,  p-value: 0.2207
```

```r
summary(lm(conf~q1, df[ptreat!=0 | ntreat!=0]))
```

```
##
## Call:
## lm(formula = conf ~ q1, data = df[ptreat != 0 | ntreat != 0])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.98225 -1.90533  0.09467  1.09467  2.09467
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.98225    0.06605  45.152   <2e-16 ***
## q1          -0.07692    0.09341  -0.824     0.41
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.487 on 1012 degrees of freedom
## Multiple R-squared:  0.0006697,  Adjusted R-squared:  -0.0003178
## F-statistic: 0.6782 on 1 and 1012 DF,  p-value: 0.4104
```

```r
#check whether order of the question and treatment matters
summary(lm(conf~first, df))
```

```
##
## Call:
## lm(formula = conf ~ first, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8864 -1.8686  0.1314  1.1314  2.1314
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.88636    0.04806  60.063   <2e-16 ***
## first       -0.01779    0.06796  -0.262    0.794
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.529 on 2022 degrees of freedom
## Multiple R-squared:  3.387e-05,  Adjusted R-squared:  -0.0004607
## F-statistic: 0.0685 on 1 and 2022 DF,  p-value: 0.7936
```

```r
summary(lm(conf~first*ptreat, df))
```

```
##
## Call:
## lm(formula = conf ~ first * ptreat, data = df)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -2.276 -1.742  0.249  1.258  2.258
```

```
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.750999   0.055193  49.843  < 2e-16 ***
## first        -0.009484   0.077671  -0.122    0.903
## ptreat        0.524863   0.108681   4.829 1.47e-06 ***
## first:ptreat -0.002150   0.155235  -0.014    0.989
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.513 on 2020 degrees of freedom
## Multiple R-squared:  0.02209,    Adjusted R-squared:  0.02064
## F-statistic: 15.21 on 3 and 2020 DF,  p-value: 8.661e-10
```

```
summary(lm(conf~first*ntreat, df))
```

```
##
## Call:
## lm(formula = conf ~ first * ntreat, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.98172 -1.64368  0.05326  1.05326  2.41057
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.98172    0.05499  54.223  < 2e-16 ***
## first        -0.03499    0.07816  -0.448 0.654461
## ntreat       -0.39229    0.11153  -3.517 0.000446 ***
## first:ntreat  0.08923    0.15620   0.571 0.567884
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.522 on 2020 degrees of freedom
## Multiple R-squared:  0.009862,   Adjusted R-squared:  0.008392
## F-statistic: 6.707 on 3 and 2020 DF,  p-value: 0.0001677
```

```
summary(lm(conf~first*ptreat, df[ptreat!=0 | ntreat!=0]))
```

```
##
## Call:
## lm(formula = conf ~ first * ptreat, data = df[ptreat != 0 | ntreat !=
##     0])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2759 -1.5894  0.3563  1.3563  2.4106
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.58943    0.09262  27.958  < 2e-16 ***
## first         0.05425    0.12909   0.420    0.674
## ptreat        0.68643    0.12909   5.318 1.29e-07 ***
```

```
## first:ptreat -0.06588    0.18256  -0.361    0.718
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.453 on 1010 degrees of freedom
## Multiple R-squared:  0.04841,    Adjusted R-squared:  0.04559
## F-statistic: 17.13 on 3 and 1010 DF,  p-value: 7.435e-11
```

```
summary(lm(conf~first*ntreat, df[ptreat!=0 | ntreat!=0]))
```

```
##
## Call:
## lm(formula = conf ~ first * ntreat, data = df[ptreat != 0 | ntreat !=
##     0])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2759 -1.5894  0.3563  1.3563  2.4106
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.27586    0.08992  36.432  < 2e-16 ***
## first        -0.01163    0.12909  -0.090    0.928
## ntreat       -0.68643    0.12909  -5.318 1.29e-07 ***
## first:ntreat  0.06588    0.18256   0.361    0.718
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.453 on 1010 degrees of freedom
## Multiple R-squared:  0.04841,    Adjusted R-squared:  0.04559
## F-statistic: 17.13 on 3 and 1010 DF,  p-value: 7.435e-11
```

```
summary(lm(conf~first*(ptreat+ntreat), df))
```

```
##
## Call:
## lm(formula = conf ~ first * (ptreat + ntreat), data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2759 -1.6437  0.2079  1.2079  2.4106
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.82970    0.06724  42.082  < 2e-16 ***
## first        -0.03762    0.09510  -0.396 0.692411
## ptreat        0.44616    0.11520   3.873 0.000111 ***
## ntreat       -0.24027    0.11749  -2.045 0.040978 *
## first:ptreat  0.02599    0.16454   0.158 0.874512
## first:ntreat  0.09187    0.16454   0.558 0.576670
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.511 on 2018 degrees of freedom
## Multiple R-squared:  0.02491,    Adjusted R-squared:  0.0225
## F-statistic: 10.31 on 5 and 2018 DF,  p-value: 8.854e-10
```

```
#ATE of positive reinforcement
summary(lm(conf~ptreat, df[ntreat==0]))
```

```
##
## Call:
## lm(formula = conf ~ ptreat, data = df[ntreat == 0])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2702 -1.8109  0.1891  1.1891  2.1891
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.81089    0.04774  58.884  < 2e-16 ***
## ptreat       0.45933    0.08257   5.563 3.14e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.517 on 1515 degrees of freedom
## Multiple R-squared:  0.02002,    Adjusted R-squared:  0.01937
## F-statistic: 30.94 on 1 and 1515 DF,  p-value: 3.136e-08
```

```
summary(lm(conf~ptreat, df[ntreat!=0 | ptreat !=0]))
```

```
##
## Call:
## lm(formula = conf ~ ptreat, data = df[ntreat != 0 | ptreat !=
##     0])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2702 -1.6174  0.3826  1.3826  2.3826
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.61736    0.06446  40.606  < 2e-16 ***
## ptreat       0.65286    0.09116   7.162 1.53e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.451 on 1012 degrees of freedom
## Multiple R-squared:  0.04824,    Adjusted R-squared:  0.0473
## F-statistic: 51.29 on 1 and 1012 DF,  p-value: 1.529e-12
```

```
#ATE of negative reinforcement
summary(lm(conf~ntreat, df[ptreat==0]))
```

```
##
```

```
## Call:
## lm(formula = conf ~ ntreat, data = df[ptreat == 0])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8109 -1.6174  0.1891  1.1891  2.3826
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.81089    0.04850  57.962   <2e-16 ***
## ntreat      -0.19353    0.08389  -2.307   0.0212 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.541 on 1515 degrees of freedom
## Multiple R-squared:  0.003501,   Adjusted R-squared:  0.002843
## F-statistic: 5.323 on 1 and 1515 DF,  p-value: 0.02118
```

```
summary(lm(conf~ntreat, df[ntreat!=0 | ptreat !=0]))
```

```
##
## Call:
## lm(formula = conf ~ ntreat, data = df[ntreat != 0 | ptreat !=
##     0])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2702 -1.6174  0.3826  1.3826  2.3826
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.27022    0.06446  50.734  < 2e-16 ***
## ntreat      -0.65286    0.09116  -7.162 1.53e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.451 on 1012 degrees of freedom
## Multiple R-squared:  0.04824,    Adjusted R-squared:  0.0473
## F-statistic: 51.29 on 1 and 1012 DF,  p-value: 1.529e-12
```

```
#combine both
summary(ma<-lm(conf~ptreat+ntreat, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2702 -1.6174  0.1891  1.1891  2.3826
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  2.81089    0.04752  59.156  < 2e-16 ***
## ptreat       0.45933    0.08219   5.588  2.6e-08 ***
## ntreat      -0.19353    0.08219  -2.355   0.0186 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.51 on 2021 degrees of freedom
## Multiple R-squared:  0.02476,    Adjusted R-squared:  0.02379
## F-statistic: 25.65 on 2 and 2021 DF,  p-value: 9.982e-12
```

```
#check whether pre-confidence level determines outcome
summary(lm(conf~ptreat+ntreat+pre.conf, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3989 -0.9694  0.0306  0.9027  3.6795
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.67171    0.09212   7.292 4.37e-13 ***
## ptreat       0.48312    0.07120   6.786 1.51e-11 ***
## ntreat      -0.16974    0.07120  -2.384   0.0172 *
## pre.conf     0.64882    0.02500  25.956  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.308 on 2020 degrees of freedom
## Multiple R-squared:  0.2687, Adjusted R-squared:  0.2676
## F-statistic: 247.4 on 3 and 2020 DF,  p-value: < 2.2e-16
```

```
#check whether getting the answer correct affects outcome
summary(lm(conf~ptreat+ntreat+pre.conf+crt, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1752 -0.8920  0.0788  0.8935  3.9074
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.48307    0.08899   5.428 6.37e-08 ***
## ptreat       0.45455    0.06802   6.683 3.02e-11 ***
## ntreat      -0.15573    0.06800  -2.290   0.0221 *
## pre.conf     0.60953    0.02403  25.361  < 2e-16 ***
## crtTRUE      0.79944    0.05707  14.009  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.249 on 2019 degrees of freedom
## Multiple R-squared:  0.3335, Adjusted R-squared:  0.3321
## F-statistic: 252.5 on 4 and 2019 DF,  p-value: < 2.2e-16
```

```
summary(lm(conf~ptreat+ntreat+pre.conf+crt+qtime, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt + qtime,
##     data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2098 -0.8873  0.0777  0.9107  3.8801
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.5241194  0.0930347   5.634 2.01e-08 ***
## ptreat       0.4527531  0.0680082   6.657 3.58e-11 ***
## ntreat      -0.1583715  0.0679970  -2.329    0.020 *
## pre.conf     0.6106952  0.0240386  25.405  < 2e-16 ***
## crtTRUE      0.7989081  0.0570482  14.004  < 2e-16 ***
## qtime       -0.0014723  0.0009765  -1.508    0.132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.249 on 2018 degrees of freedom
## Multiple R-squared:  0.3342, Adjusted R-squared:  0.3326
## F-statistic: 202.6 on 5 and 2018 DF,  p-value: < 2.2e-16
```

```
summary(lm(conf~ptreat+ntreat+pre.conf+crt+qtime+cftime, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt + qtime +
##     cftime, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2280 -0.8831  0.0784  0.9060  3.8819
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.5442189  0.0941789   5.779 8.71e-09 ***
## ptreat       0.4643474  0.0685248   6.776 1.61e-11 ***
## ntreat      -0.1484246  0.0683739  -2.171   0.0301 *
## pre.conf     0.6088132  0.0240733  25.290  < 2e-16 ***
## crtTRUE      0.7969361  0.0570545  13.968  < 2e-16 ***
## qtime       -0.0013965  0.0009779  -1.428   0.1534
## cftime      -0.0062981  0.0046252  -1.362   0.1734
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.248 on 2017 degrees of freedom
## Multiple R-squared:  0.3348, Adjusted R-squared:  0.3328
## F-statistic: 169.2 on 6 and 2017 DF,  p-value: < 2.2e-16
```

```
summary(lm(conf~pre.conf*(ptreat+ntreat), df))
```

```
##
## Call:
## lm(formula = conf ~ pre.conf * (ptreat + ntreat), data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3385 -0.9392  0.0608  0.9094  3.7328
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.59513    0.12537   4.747 2.21e-06 ***
## pre.conf         0.67205    0.03592  18.711  < 2e-16 ***
## ptreat           0.67289    0.21185   3.176  0.00151 **
## ntreat          -0.06396    0.21185  -0.302  0.76276
## pre.conf:ptreat -0.05794    0.06096  -0.951  0.34195
## pre.conf:ntreat -0.03218    0.06096  -0.528  0.59759
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.308 on 2018 degrees of freedom
## Multiple R-squared:  0.269,  Adjusted R-squared:  0.2672
## F-statistic: 148.5 on 5 and 2018 DF,  p-value: < 2.2e-16
```

```
#big model that runs on covariates
summary(mb<-lm(conf~ptreat+ntreat+pre.conf+crt, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1752 -0.8920  0.0788  0.8935  3.9074
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.48307    0.08899   5.428 6.37e-08 ***
## ptreat       0.45455    0.06802   6.683 3.02e-11 ***
## ntreat      -0.15573    0.06800  -2.290   0.0221 *
## pre.conf     0.60953    0.02403  25.361  < 2e-16 ***
## crtTRUE      0.79944    0.05707  14.009  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.249 on 2019 degrees of freedom
```

```
## Multiple R-squared:  0.3335, Adjusted R-squared:  0.3321
## F-statistic: 252.5 on 4 and 2019 DF,  p-value: < 2.2e-16
```

```
summary(mc<-lm(conf~ptreat+ntreat+pre.conf+crt+edu, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt + edu, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4311 -0.8512  0.0508  0.9407  4.0508
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.05175    0.44829  -0.115  0.90811
## ptreat       0.46241    0.06705   6.897  7.1e-12 ***
## ntreat      -0.15010    0.06702  -2.239  0.02524 *
## pre.conf     0.59422    0.02379  24.975  < 2e-16 ***
## crtTRUE      0.75785    0.05641  13.434  < 2e-16 ***
## edu2         0.11379    0.44723   0.254  0.79918
## edu3         0.40674    0.43879   0.927  0.35405
## edu4         0.51391    0.44509   1.155  0.24838
## edu5         0.70093    0.43769   1.601  0.10944
## edu6         0.88567    0.44091   2.009  0.04470 *
## edu7         1.30374    0.48635   2.681  0.00741 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.229 on 2013 degrees of freedom
## Multiple R-squared:  0.3562, Adjusted R-squared:  0.353
## F-statistic: 111.4 on 10 and 2013 DF,  p-value: < 2.2e-16
```

```
summary(md<-lm(conf~ptreat+ntreat+pre.conf+crt+edu+race+gender, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt + edu +
##     race + gender, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3007 -0.9066  0.0821  0.9207  4.0821
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.1110618  0.4542721   0.244  0.80688
## ptreat       0.4526363  0.0669129   6.765 1.75e-11 ***
## ntreat      -0.1610057  0.0669076  -2.406  0.01620 *
## pre.conf     0.5695798  0.0246902  23.069  < 2e-16 ***
## crtTRUE      0.7364266  0.0565904  13.013  < 2e-16 ***
## edu2         0.1078190  0.4496575   0.240  0.81053
## edu3         0.4067330  0.4413236   0.922  0.35684
```

```
## edu4         0.5117745  0.4480148   1.142  0.25346
## edu5         0.6852415  0.4402883   1.556  0.11978
## edu6         0.8891060  0.4429906   2.007  0.04488 *
## edu7         1.2218666  0.4890948   2.498  0.01256 *
## race2        0.0007871  0.1127554   0.007  0.99443
## race3       -0.4182674  0.3290917  -1.271  0.20388
## race4       -0.1786657  0.1177840  -1.517  0.12945
## race5        0.3167897  0.0985642   3.214  0.00133 **
## race6       -0.4193625  0.4345351  -0.965  0.33462
## race7        0.1697791  0.2547162   0.667  0.50514
## gender2     -0.1694847  0.0575690  -2.944  0.00328 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.224 on 2006 degrees of freedom
## Multiple R-squared:  0.3639, Adjusted R-squared:  0.3585
## F-statistic: 67.51 on 17 and 2006 DF,  p-value: < 2.2e-16
```

```
summary(me<-lm(conf~ptreat+ntreat+pre.conf+crt+edu+race+gender+qtime+cftime, df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt + edu +
##     race + gender + qtime + cftime, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3431 -0.8910  0.0637  0.9090  4.0454
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.1914550  0.4560581   0.420  0.67467
## ptreat       0.4604155  0.0674051   6.831 1.12e-11 ***
## ntreat      -0.1554573  0.0672778  -2.311  0.02095 *
## pre.conf     0.5689714  0.0247338  23.004  < 2e-16 ***
## crtTRUE      0.7347034  0.0565767  12.986  < 2e-16 ***
## edu2         0.0876918  0.4496200   0.195  0.84538
## edu3         0.3844469  0.4413313   0.871  0.38380
## edu4         0.4880817  0.4480523   1.089  0.27613
## edu5         0.6640511  0.4402872   1.508  0.13166
## edu6         0.8641106  0.4430262   1.950  0.05126 .
## edu7         1.2126742  0.4888994   2.480  0.01320 *
## race2        0.0135816  0.1129070   0.120  0.90427
## race3       -0.4267240  0.3289718  -1.297  0.19473
## race4       -0.1643623  0.1180763  -1.392  0.16408
## race5        0.3189158  0.0985555   3.236  0.00123 **
## race6       -0.4065177  0.4344110  -0.936  0.34949
## race7        0.1671965  0.2546293   0.657  0.51150
## gender2     -0.1711346  0.0575634  -2.973  0.00298 **
## qtime       -0.0014364  0.0009616  -1.494  0.13541
## cftime      -0.0053629  0.0045518  -1.178  0.23886
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.224 on 2004 degrees of freedom
## Multiple R-squared:  0.3651, Adjusted R-squared:  0.3591
## F-statistic: 60.66 on 19 and 2004 DF,  p-value: < 2.2e-16
```

```
stargazer(ma, mb, mc, md, me, type = 'text')
```

```
##
## ===============================================================================================
##                                                                Dependent variable:
##                     ------------------------------------------------------------------------
##                                                                       conf
##                               (1)                    (2)                (3)
## -----------------------------------------------------------------------------------------------
## ptreat                      0.459***               0.455***           0.462***
##                             (0.082)                (0.068)            (0.067)
##
## ntreat                      -0.194**               -0.156**           -0.150**
##                             (0.082)                (0.068)            (0.067)
##
## pre.conf                                           0.610***           0.594***
##                                                    (0.024)            (0.024)
##
## crt                                                0.799***           0.758***
##                                                    (0.057)            (0.056)
##
## edu2                                                                   0.114
##                                                                       (0.447)
##
## edu3                                                                   0.407
##                                                                       (0.439)
##
## edu4                                                                   0.514
##                                                                       (0.445)
##
## edu5                                                                   0.701
##                                                                       (0.438)
##
## edu6                                                                  0.886**
##                                                                       (0.441)
##
## edu7                                                                  1.304***
##                                                                       (0.486)
##
## race2
##
##
## race3
##
##
## race4
##
##
## race5
##
```

```
##
## race6
##
##
## race7
##
##
## gender2
##
##
## qtime
##
##
## cftime
##
##
## Constant                    2.811***                0.483***                  -0.052
##                             (0.048)                  (0.089)                 (0.448)
##
## ------------------------------------------------------------------------------------------------------
## Observations                 2,024                    2,024                    2,024
## R2                           0.025                    0.333                    0.356
## Adjusted R2                  0.024                    0.332                    0.353
## Residual Std. Error    1.510 (df = 2021)        1.249 (df = 2019)        1.229 (df = 2013)
## F Statistic         25.650*** (df = 2; 2021) 252.527*** (df = 4; 2019) 111.359*** (df = 10; 2013) 67
## ======================================================================================================
## Note:
```

```
anova(ma, mb, mc, md, me, test = 'F')
```

```
## Analysis of Variance Table
##
## Model 1: conf ~ ptreat + ntreat
## Model 2: conf ~ ptreat + ntreat + pre.conf + crt
## Model 3: conf ~ ptreat + ntreat + pre.conf + crt + edu
## Model 4: conf ~ ptreat + ntreat + pre.conf + crt + edu + race + gender
## Model 5: conf ~ ptreat + ntreat + pre.conf + crt + edu + race + gender +
##     qtime + cftime
##   Res.Df    RSS Df Sum of Sq        F    Pr(>F)
## 1   2021 4608.6
## 2   2019 3149.8  2   1458.85 487.2323 < 2.2e-16 ***
## 3   2013 3042.5  6    107.27  11.9424  3.27e-13 ***
## 4   2006 3005.9  7     36.63   3.4956  0.000988 ***
## 5   2004 3000.1  2      5.72   1.9112  0.148170
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#HTE checks

summary(lm(conf~gender*(ptreat+ntreat), df))
```

```
##
## Call:
```

```
## lm(formula = conf ~ gender * (ptreat + ntreat), data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4759 -1.5410  0.1103  1.1103  2.7465
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     3.04029    0.06367  47.748  < 2e-16 ***
## gender2        -0.49934    0.09394  -5.315 1.18e-07 ***
## ptreat          0.43557    0.10811   4.029 5.81e-05 ***
## ntreat         -0.15064    0.10811  -1.393    0.164
## gender2:ptreat  0.01887    0.16328   0.116    0.908
## gender2:ntreat -0.13685    0.16328  -0.838    0.402
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.488 on 2018 degrees of freedom
## Multiple R-squared:  0.05469,    Adjusted R-squared:  0.05234
## F-statistic: 23.35 on 5 and 2018 DF,  p-value: < 2.2e-16
```

```
summary(lm(conf~ptreat+ntreat+pre.conf+crt+edu+race+gender+crt*(ptreat+ntreat), df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt + edu +
##     race + gender + crt * (ptreat + ntreat), data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3320 -0.8847  0.0892  0.9167  4.1283
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.059521   0.454540   0.131 0.895830
## ptreat          0.539368   0.087733   6.148 9.45e-10 ***
## ntreat         -0.052586   0.085181  -0.617 0.537078
## pre.conf        0.568740   0.024676  23.048  < 2e-16 ***
## crtTRUE         0.859365   0.079485  10.812  < 2e-16 ***
## edu2            0.112123   0.449333   0.250 0.802975
## edu3            0.415968   0.441009   0.943 0.345682
## edu4            0.519707   0.447710   1.161 0.245856
## edu5            0.691219   0.439971   1.571 0.116328
## edu6            0.889449   0.442657   2.009 0.044635 *
## edu7            1.233993   0.488757   2.525 0.011654 *
## race2           0.002815   0.112717   0.025 0.980081
## race3          -0.400477   0.328986  -1.217 0.223631
## race4          -0.182391   0.117751  -1.549 0.121550
## race5           0.328021   0.098617   3.326 0.000896 ***
## race6          -0.463540   0.434690  -1.066 0.286385
## race7           0.165247   0.254534   0.649 0.516274
## gender2        -0.172484   0.057541  -2.998 0.002755 **
## ptreat:crtTRUE -0.211314   0.135380  -1.561 0.118707
## ntreat:crtTRUE -0.281010   0.137442  -2.045 0.041028 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.223 on 2004 degrees of freedom
## Multiple R-squared:  0.3655, Adjusted R-squared:  0.3595
## F-statistic: 60.76 on 19 and 2004 DF,  p-value: < 2.2e-16
```

```
summary(lm(conf~ptreat+ntreat+pre.conf+crt+edu+race+gender+pre.conf*(ptreat+ntreat), df))
```

```
##
## Call:
## lm(formula = conf ~ ptreat + ntreat + pre.conf + crt + edu +
##     race + gender + pre.conf * (ptreat + ntreat), data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2799 -0.8996  0.0867  0.9142  4.1195
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.068041   0.458901   0.148  0.88214
## ptreat           0.585136   0.199097   2.939  0.00333 **
## ntreat          -0.085235   0.199038  -0.428  0.66853
## pre.conf         0.586024   0.034703  16.887  < 2e-16 ***
## crtTRUE          0.735742   0.056620  12.994  < 2e-16 ***
## edu2             0.098638   0.450029   0.219  0.82653
## edu3             0.395552   0.441798   0.895  0.37072
## edu4             0.500738   0.448478   1.117  0.26433
## edu5             0.673648   0.440786   1.528  0.12660
## edu6             0.877868   0.443468   1.980  0.04789 *
## edu7             1.219824   0.489284   2.493  0.01274 *
## race2            0.002901   0.112841   0.026  0.97950
## race3           -0.410344   0.329422  -1.246  0.21304
## race4           -0.181607   0.117908  -1.540  0.12366
## race5            0.316881   0.098601   3.214  0.00133 **
## race6           -0.422769   0.434724  -0.972  0.33092
## race7            0.163108   0.255002   0.640  0.52248
## gender2         -0.169126   0.057593  -2.937  0.00336 **
## ptreat:pre.conf -0.040455   0.057285  -0.706  0.48014
## ntreat:pre.conf -0.023067   0.057278  -0.403  0.68720
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.225 on 2004 degrees of freedom
## Multiple R-squared:  0.3641, Adjusted R-squared:  0.3581
## F-statistic: 60.39 on 19 and 2004 DF,  p-value: < 2.2e-16
```

```
summary(lm(conf~edu*(ptreat+ntreat), df))
```

```
##
## Call:
## lm(formula = conf ~ edu * (ptreat + ntreat), data = df)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6364 -1.5070  0.1622  1.3636  3.1622
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.0000     1.0413   3.841 0.000126 ***
## edu2         -2.1622     1.0553  -2.049 0.040610 *
## edu3         -1.4919     1.0455  -1.427 0.153747
## edu4         -1.4388     1.0519  -1.368 0.171539
## edu5         -1.0024     1.0438  -0.960 0.337020
## edu6         -0.6731     1.0480  -0.642 0.520784
## edu7         -0.3333     1.1248  -0.296 0.766990
## ptreat       -0.6667     1.3444  -0.496 0.620023
## ntreat       -2.3333     1.3444  -1.736 0.082782 .
## edu2:ptreat   1.5920     1.3761   1.157 0.247465
## edu3:ptreat   1.1093     1.3533   0.820 0.412473
## edu4:ptreat   1.3673     1.3715   0.997 0.318910
## edu5:ptreat   1.0947     1.3504   0.811 0.417673
## edu6:ptreat   0.9761     1.3599   0.718 0.472986
## edu7:ptreat   0.9000     1.4849   0.606 0.544517
## edu2:ntreat   2.7587     1.3761   2.005 0.045134 *
## edu3:ntreat   2.3323     1.3533   1.723 0.084960 .
## edu4:ntreat   2.2483     1.3715   1.639 0.101316
## edu5:ntreat   1.8742     1.3504   1.388 0.165336
## edu6:ntreat   2.1103     1.3599   1.552 0.120873
## edu7:ntreat   2.8667     1.4849   1.931 0.053681 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.473 on 2003 degrees of freedom
## Multiple R-squared:  0.08073,    Adjusted R-squared:  0.07156
## F-statistic: 8.796 on 20 and 2003 DF,  p-value: < 2.2e-16
```

```
#produces likert plot
dl1 = df[,.('Pre-Confidence' = factor(pre.conf, labels = c('Not Confident', 'Somewhat Confident', 'Neutr
dl2 = df[ptreat==0 & ntreat==0, .(Control = factor(conf, labels = c('Not Confident', 'Somewhat Confident
dl3 = df[ptreat==1, .('Positive Treatment' = factor(conf, labels = c('Not Confident', 'Somewhat Confider
dl4 = df[ntreat==1, .('Negative Treatment' = factor(conf, labels = c('Not Confident', 'Somewhat Confider

dlc = as.data.frame(cbind(dl1, dl2, dl3, dl4))
```

```
## Warning in data.table::data.table(...): Item 2 is of size 1010 but maximum
## size is 2024 (recycled leaving remainder of 4 items)
```

```
## Warning in data.table::data.table(...): Item 3 is of size 507 but maximum
## size is 2024 (recycled leaving remainder of 503 items)
```

```
## Warning in data.table::data.table(...): Item 4 is of size 507 but maximum
## size is 2024 (recycled leaving remainder of 503 items)
```

```
dl = likert(dlc)
plot(dl, order = FALSE, group.order = names(dlc))
```