analysis-benign-final

Yixin Zou

September 13, 2018

```r
#Recode and format variables for the model
#Start with cateogircal variables with regards to participants
#For gender, we code 0 (N/A) and non-binary as other, and exclude other in the final analysis because t
data_all$DQ_Gender[data_all$DQ_Gender=="0" | data_all$DQ_Gender=="Non binary/third gender"] <- "Other"
data_all$gender.f <- factor(data_all$DQ_Gender,levels=c("Female","Male","Other"), exclude = "Other")

#For education, we divide it into no bachelor's degree, bachelor's degree, and graduate degree
data_all$DQ_Educ[data_all$DQ_Educ=="0"] <- "Other"
data_all$DQ_Educ[data_all$DQ_Educ=="High school or equivalent (e.g., GED)" |
      data_all$DQ_Educ=="Some college but no degree"] <- "No Bachelor's degree"
data_all$DQ_Educ[data_all$DQ_Educ=="Associate's degree" |
      data_all$DQ_Educ=="Trade, technical, or vocational training" |
      data_all$DQ_Educ=="Bachelor's degree"] <- "Bachelor's degree"
data_all$DQ_Educ[data_all$DQ_Educ=="Professional degree (JD, MD etc.)" |
      data_all$DQ_Educ=="Doctoral's degree" |
        data_all$DQ_Educ=="Master's degree"] <- "Graduate degree"
data_all$education.f <- factor(data_all$DQ_Educ,
      levels=c("No Bachelor's degree","Bachelor's degree","Graduate degree"), exclude = "Other")

#For occupation, we divide it into technical vs. non-technical
data_all$DQ_Occ[data_all$DQ_Occ == "Computers (Hardware, Desktop Software)"|
        data_all$DQ_Occ == "Engineering / Architecture" |
          data_all$DQ_Occ == "Internet"] <- "Technical occupations"
data_all$DQ_Occ[data_all$DQ_Occ != "Technical occupations"] <- "Non-technical occupations"
data_all$occupation.f <- factor(data_all$DQ_Occ,
                      levels=c("Technical occupations","Non-technical occupations"))

#For each email and link, we categorize if it's a phish or not, and factor the related variables
data_all$p_email.f <- factor(data_all$p_email)
data_all$p_link.f <- factor(data_all$p_link)

#Factor condition groups and placement, give them labels
data_all$placement.f <- factor(data_all$placement,
          level=c("inmail","banner","browser","no warning"))
data_all$condition_group.f <- factor(data_all$condition_group,
          levels=c(1,2,3,4,5,6,0),
          labels=c("on-load, no forced","on-load, forced",
                    "on-click, no forced","on-click, forced",
                    "banner","browser","no warning"))

#Factor dependent variables: click and hover actions
data_all$click_action.f <- factor(data_all$click_action, labels=c("no","yes"))
data_all$hover_action.f <- factor(data_all$hover_action, labels=c("no","yes"))

#Standardize continuous variables
library(standardize)
data_all$age_scaled <- scale(as.numeric(data_all$DQ_Age))[, 1]
```

```r
data_all$cyber_quiz_score_scaled <- scale(data_all$cyber_quiz_score)[, 1]
data_all$PE_score_scaled <- scale(data_all$PE_score)[, 1]
data_all$brand_usage_scaled <- scale(data_all$brand_usage)[, 1]
data_all$hover_time_scaled <- scale(data_all$hover_time)[, 1]
#rename time series variable
colnames(data_all)[which(names(data_all) == "time_series")] <- "warnings_seen"
data_all$warnings_seen_scaled <- scale(data_all$warnings_seen)[, 1]

#Factor random effect variables
data_all$ref_id.f <- factor(data_all$ref_id)
data_all$username.f <- factor(data_all$username)
data_all$adj_link_id.f <- factor(data_all$adj_link_id)
```

```r
#Filter the dataset to include benign link entries only
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
```

```r
data_benign_only <- data_all %>%
  filter(data_all$p_link == "False")
```

```r
#We conduct a one-way ANOVA on benign links CTR between groups
library(dplyr)
data_benign_ctr <- data_benign_only %>%
  select(username, non_phish_ctr, condition_group.f) %>%
  distinct(username, non_phish_ctr, condition_group.f)

#descriptive stats for each group, look into normality
library(pastecs)
```

```
##
## Attaching package: 'pastecs'

## The following objects are masked from 'package:dplyr':
##
##     first, last
```

```r
stat.desc(data_benign_ctr$non_phish_ctr, norm = TRUE)
```

```
##                           nbr.val                            nbr.null
## 701.000000000000000000000000000 153.000000000000000000000000000
##                            nbr.na                                 min
##   0.000000000000000000000000000   0.000000000000000000000000000
##                               max                               range
##   1.000000000000000000000000000   1.000000000000000000000000000
##                               sum                              median
## 352.464285714285724062026944011   0.571428571428571396850770725O
##                              mean                             SE.mean
```

```
##    0.50280211942123498314316520920   0.01464552097641516614778822271
##                        CI.mean.0.95                           var
##    0.02875441118763983130990034709   0.15035839055410224474051972227
##                             std.dev                      coef.var
##    0.38776073879920108788965649180   0.77119949145310762173011198680
##                            skewness                      skew.2SE
##   -0.13905436531115433762550992470  -0.75312046795404676461507091684
##                            kurtosis                      kurt.2SE
##   -1.64518919421927978241626533420  -4.61471901376322435339716321300
##                           normtest.W                    normtest.p
##    0.84647075546113281774296410730   0.00000000000000000000000001712
```

```
by(data_benign_ctr$non_phish_ctr, data_benign_ctr$condition_group.f, stat.desc, norm = TRUE)
```

```
## data_benign_ctr$condition_group.f: on-load, no forced
##          nbr.val         nbr.null           nbr.na              min
## 103.00000000000   21.00000000000    0.00000000000    0.00000000000
##              max            range              sum           median
##    1.00000000000    1.00000000000   55.39285714286    0.60714285714
##             mean          SE.mean     CI.mean.0.95              var
##    0.53779472954    0.03732278805    0.07402957313    0.14347802228
##          std.dev         coef.var         skewness         skew.2SE
##    0.37878492879    0.70432993851   -0.29936873959   -0.62908947454
##         kurtosis         kurt.2SE       normtest.W       normtest.p
##   -1.52927871572   -1.62127893773    0.85506364074    0.00000001273
## -----------------------------------------------------------
## data_benign_ctr$condition_group.f: on-load, forced
##          nbr.val         nbr.null           nbr.na              min             max
## 98.0000000000   13.0000000000    0.0000000000    0.0000000000    1.0000000000
##            range              sum           median             mean          SE.mean
##  1.0000000000   52.5357142857    0.6071428571    0.5360787172    0.0381748876
##     CI.mean.0.95              var          std.dev         coef.var         skewness
##  0.0757665845    0.1428175599    0.3779121061    0.7049563692   -0.1831481746
##        skew.2SE         kurtosis         kurt.2SE       normtest.W       normtest.p
## -0.3756747183   -1.6366497443   -1.6943849168    0.8549486419    0.0000000232
## -----------------------------------------------------------
## data_benign_ctr$condition_group.f: on-click, no forced
##          nbr.val         nbr.null           nbr.na              min             max
## 101.0000000000   27.0000000000    0.0000000000    0.0000000000    1.0000000000
##            range              sum           median             mean          SE.mean
##  1.0000000000   49.7142857143    0.5714285714    0.4922206506    0.0394368443
##     CI.mean.0.95              var          std.dev         coef.var         skewness
##  0.0782415759    0.1570817337    0.3963353803    0.8051986031   -0.1102878085
##        skew.2SE         kurtosis         kurt.2SE       normtest.W       normtest.p
## -0.2295596492   -1.6861960079   -1.7709732068    0.8371629626    0.0000000036
## -----------------------------------------------------------
## data_benign_ctr$condition_group.f: on-click, forced
##              nbr.val            nbr.null              nbr.na                 min
## 100.0000000000000   25.0000000000000    0.0000000000000    0.0000000000000
##                  max               range                 sum              median
##    1.0000000000000    1.0000000000000   48.8571428571429    0.5535714285714
##                 mean             SE.mean        CI.mean.0.95                 var
##    0.4885714285714    0.0415284659211    0.0824014860539    0.1724613481756
##              std.dev            coef.var            skewness            skew.2SE
##    0.4152846592106    0.8499978404895   -0.0312577834094   -0.0647481399098
```

3

```
##         kurtosis          kurt.2SE          normtest.W         normtest.p
## -1.8020868681700 -1.8837231614919    0.8047705618197    0.0000000003478
## ------------------------------------------------------------
## data_benign_ctr$condition_group.f: banner
##           nbr.val          nbr.null           nbr.na                 min
## 100.000000000000  24.000000000000    0.000000000000    0.000000000000
##               max             range              sum              median
##   1.000000000000   1.000000000000   50.750000000000    0.589285714286
##              mean           SE.mean       CI.mean.0.95                 var
##   0.507500000000   0.039618149419    0.078611003667    0.156959776335
##           std.dev          coef.var          skewness            skew.2SE
##   0.396181494185   0.780653190513   -0.174532170351   -0.361530222302
##          kurtosis          kurt.2SE         normtest.W          normtest.p
##  -1.687450150857  -1.763893289041    0.831937022794    0.000000002695
## ------------------------------------------------------------
## data_benign_ctr$condition_group.f: browser
##          nbr.val       nbr.null         nbr.na            min            max
## 99.0000000000 20.0000000000   0.0000000000   0.0000000000   1.0000000000
##            range            sum         median           mean        SE.mean
##   1.0000000000 44.6428571429   0.4642857143   0.4509379509   0.0366538149
##      CI.mean.0.95            var        std.dev       coef.var       skewness
##   0.0727383027   0.1330067123   0.3647008531   0.8087606119   0.0757739482
##          skew.2SE       kurtosis       kurt.2SE      normtest.W      normtest.p
##   0.1561958941  -1.5666559456  -1.6297934981    0.8777874779    0.0000001641
## ------------------------------------------------------------
## data_benign_ctr$condition_group.f: no warning
##           nbr.val          nbr.null           nbr.na                 min
## 100.000000000000  23.000000000000    0.000000000000    0.000000000000
##               max             range              sum              median
##   1.000000000000   1.000000000000   50.571428571429    0.660714285714
##              mean           SE.mean       CI.mean.0.95                 var
##   0.505714285714   0.038758165896    0.076904609783    0.150219542362
##           std.dev          coef.var          skewness            skew.2SE
##   0.387581658960   0.766404410373   -0.241538363236   -0.500328495197
##          kurtosis          kurt.2SE         normtest.W          normtest.p
##  -1.655109016848  -1.730087070107    0.832416412524    0.000000002799
```

```r
boxplot(data_benign_ctr$non_phish_ctr~data_benign_ctr$condition_group.f)
```
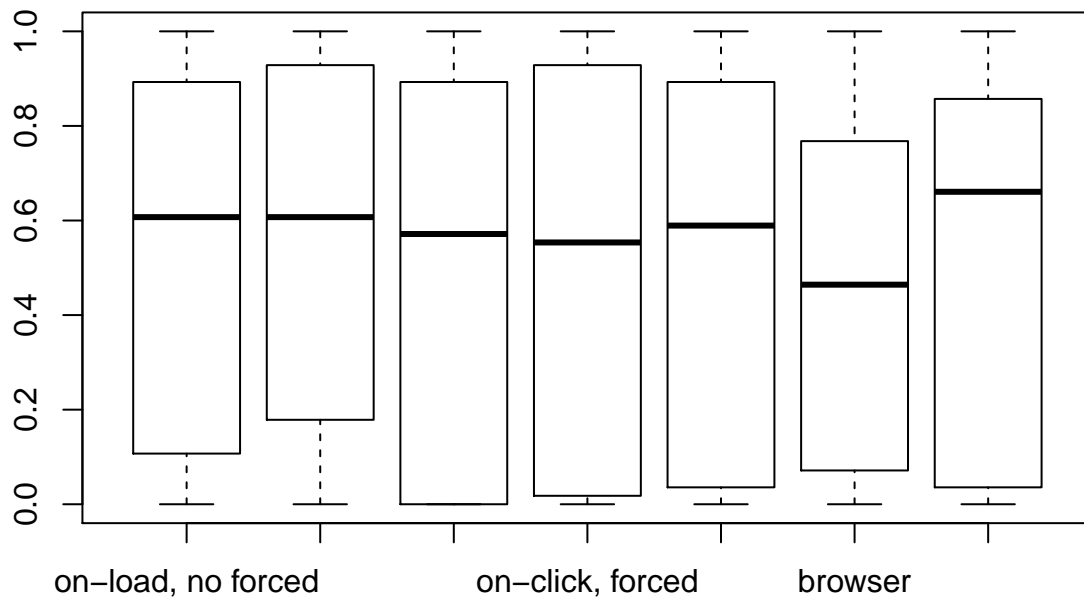
```
#one way anova
benign_ctr_model <- lm(non_phish_ctr~condition_group.f, data=data_benign_ctr)
summary(benign_ctr_model)
```

```
##
## Call:
## lm(formula = non_phish_ctr ~ condition_group.f, data = data_benign_ctr)
##
## Residuals:
##    Min     1Q  Median     3Q    Max
## -0.5378 -0.4509  0.0693  0.3854  0.5491
##
## Coefficients:
##                                  Estimate Std. Error t value
## (Intercept)                       0.53779    0.03827   14.05
## condition_group.fon-load, forced -0.00172    0.05481   -0.03
## condition_group.fon-click, no forced -0.04557 0.05440   -0.84
## condition_group.fon-click, forced -0.04922    0.05453   -0.90
## condition_group.fbanner          -0.03029    0.05453   -0.56
## condition_group.fbrowser         -0.08686    0.05467   -1.59
## condition_group.fno warning      -0.03208    0.05453   -0.59
##                                              Pr(>|t|)
## (Intercept)                      <0.0000000000000002 ***
## condition_group.fon-load, forced                0.98
## condition_group.fon-click, no forced            0.40
## condition_group.fon-click, forced               0.37
```

```
## condition_group.fbanner                               0.58
## condition_group.fbrowser                              0.11
## condition_group.fno warning                           0.56
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.388 on 694 degrees of freedom
## Multiple R-squared:  0.00509,    Adjusted R-squared:  -0.00351
## F-statistic: 0.592 on 6 and 694 DF,  p-value: 0.737
```

```r
anova(benign_ctr_model)
```

```
## Analysis of Variance Table
##
## Response: non_phish_ctr
##                    Df Sum Sq Mean Sq F value Pr(>F)
## condition_group.f   6    0.5  0.0893    0.59   0.74
## Residuals         694  104.7  0.1509
```

```r
#calculate the effect size
library(heplots)
```

```
## Loading required package: car
```

```
## Loading required package: carData
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```r
etasq(benign_ctr_model, anova = TRUE)
```

```
## Anova Table (Type II tests)
##
## Response: non_phish_ctr
##                   Partial eta^2 Sum Sq  Df F value Pr(>F)
## condition_group.f       0.00509    0.5   6    0.59   0.74
## Residuals                        104.7 694
```

```r
lsr::etaSquared(benign_ctr_model)
```

```
##                      eta.sq eta.sq.part
## condition_group.f 0.005088    0.005088
```

```r
#Mixed-effect logistic regression model on benign links click action
#For all group comparison we include placement as the only warning-related predictor
library(lme4)
```

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

```r
glm_click_benign <- glmer(click_action.f ~
    placement.f #warning-related factors
    + warnings_seen_scaled #number of warnings seen before and during clicking on the link
    + cyber_quiz_score_scaled + PE_score_scaled + brand_usage_scaled # phishing-related individual cha
    + gender.f + education.f + age_scaled + occupation.f #participants demographics
    + (1 | ref_id.f) + (1 | adj_link_id.f), #random effect, dropped username.f for benign link analysi
    data=data_benign_only,
```

```
    family=binomial(link=logit),
    control=glmerControl(optCtrl=list(maxfun=2e4)))
```

singular fit

```
summary(glm_click_benign)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## click_action.f ~ placement.f + warnings_seen_scaled + cyber_quiz_score_scaled +
##     PE_score_scaled + brand_usage_scaled + gender.f + education.f +
##     age_scaled + occupation.f + (1 | ref_id.f) + (1 | adj_link_id.f)
##    Data: data_benign_only
## Control: glmerControl(optCtrl = list(maxfun = 20000))
##
##      AIC      BIC   logLik deviance df.resid
##    28928    29048   -14449    28898    22161
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -2.104 -0.914 -0.435  0.916  3.511
##
## Random effects:
##  Groups        Name        Variance      Std.Dev.
##  adj_link_id.f (Intercept) 0.19252550588 0.4387773
##  ref_id.f      (Intercept) 0.00000000171 0.0000414
## Number of obs: 22176, groups:  adj_link_id.f, 32; ref_id.f, 10
##
## Fixed effects:
##                                   Estimate Std. Error z value
## (Intercept)                       -0.120661   0.091594   -1.32
## placement.fbanner                  0.070512   0.041799    1.69
## placement.fbrowser                -0.071999   0.041918   -1.72
## placement.fno warning              0.199706   0.046933    4.26
## warnings_seen_scaled               0.027780   0.015946    1.74
## cyber_quiz_score_scaled            0.302462   0.014912   20.28
## PE_score_scaled                   -0.073057   0.014696   -4.97
## brand_usage_scaled                 0.021880   0.021656    1.01
## gender.fMale                      -0.485671   0.029028  -16.73
## education.fBachelor's degree      -0.000205   0.031162   -0.01
## education.fGraduate degree        -0.405813   0.049070   -8.27
## age_scaled                         0.053651   0.014276    3.76
## occupation.fNon-technical occupations  0.414676   0.039901   10.39
##                                                Pr(>|z|)
## (Intercept)                                     0.18772
## placement.fbanner                               0.09162 .
## placement.fbrowser                              0.08587 .
## placement.fno warning                        0.00002090 ***
## warnings_seen_scaled                            0.08148 .
## cyber_quiz_score_scaled         < 0.0000000000000002 ***
## PE_score_scaled                              0.00000067 ***
## brand_usage_scaled                              0.31231
```

```
## gender.fMale                       < 0.0000000000000002 ***
## education.fBachelor's degree                         0.99474
## education.fGraduate degree         < 0.0000000000000002 ***
## age_scaled                                           0.00017 ***
## occupation.fNon-technical occupations < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
##
## convergence code: 0
## singular fit
```

```r
#Print correlation tables
print(glm_click_benign, correlation=TRUE)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
## Family: binomial  ( logit )
## Formula:
## click_action.f ~ placement.f + warnings_seen_scaled + cyber_quiz_score_scaled +
##     PE_score_scaled + brand_usage_scaled + gender.f + education.f +
##     age_scaled + occupation.f + (1 | ref_id.f) + (1 | adj_link_id.f)
##    Data: data_benign_only
##     AIC      BIC   logLik deviance df.resid
##   28928    29048   -14449    28898    22161
## Random effects:
##  Groups        Name        Std.Dev.
##  adj_link_id.f (Intercept) 0.4387773
##  ref_id.f      (Intercept) 0.0000414
## Number of obs: 22176, groups:  adj_link_id.f, 32; ref_id.f, 10
## Fixed Effects:
##                         (Intercept)
##                           -0.120661
##                     placement.fbanner
##                            0.070512
##                    placement.fbrowser
##                           -0.071999
##                  placement.fno warning
##                            0.199706
##                 warnings_seen_scaled
##                            0.027780
##               cyber_quiz_score_scaled
##                            0.302462
##                      PE_score_scaled
##                           -0.073057
##                   brand_usage_scaled
##                            0.021880
##                         gender.fMale
##                           -0.485671
##          education.fBachelor's degree
##                           -0.000205
##            education.fGraduate degree
```

```
##                                          -0.405813
##                          age_scaled
##                           0.053651
## occupation.fNon-technical occupations
##                           0.414676
## convergence code 0; 1 optimizer warnings; 0 lme4 warnings
```

```r
#Get confidence intervals
se_click_benign <- sqrt(diag(vcov(glm_click_benign)))
# table of estimates with 95% CI
(tab_click_benign <- cbind(Est = fixef(glm_click_benign),
            LL = fixef(glm_click_benign) - 1.96 * se_click_benign,
            UL = fixef(glm_click_benign) + 1.96 * se_click_benign))
```

```
##                                           Est         LL        UL
## (Intercept)                          -0.1206612 -0.300186  0.05886
## placement.fbanner                     0.0705118 -0.011414  0.15244
## placement.fbrowser                   -0.0719986 -0.154157  0.01016
## placement.fno warning                 0.1997059  0.107716  0.29170
## warnings_seen_scaled                  0.0277799 -0.003474  0.05903
## cyber_quiz_score_scaled               0.3024623  0.273235  0.33169
## PE_score_scaled                      -0.0730570 -0.101862 -0.04425
## brand_usage_scaled                    0.0218804 -0.020564  0.06433
## gender.fMale                         -0.4856712 -0.542566 -0.42878
## education.fBachelor's degree         -0.0002053 -0.061282  0.06087
## education.fGraduate degree           -0.4058129 -0.501991 -0.30964
## age_scaled                            0.0536509  0.025670  0.08163
## occupation.fNon-technical occupations 0.4146765  0.336470  0.49288
```

```r
#Odds ratio
exp(tab_click_benign)
```

```
##                                           Est     LL     UL
## (Intercept)                            0.8863 0.7407 1.0606
## placement.fbanner                      1.0731 0.9887 1.1647
## placement.fbrowser                     0.9305 0.8571 1.0102
## placement.fno warning                  1.2210 1.1137 1.3387
## warnings_seen_scaled                   1.0282 0.9965 1.0608
## cyber_quiz_score_scaled                1.3532 1.3142 1.3933
## PE_score_scaled                        0.9295 0.9032 0.9567
## brand_usage_scaled                     1.0221 0.9796 1.0664
## gender.fMale                           0.6153 0.5813 0.6513
## education.fBachelor's degree           0.9998 0.9406 1.0628
## education.fGraduate degree             0.6664 0.6053 0.7337
## age_scaled                             1.0551 1.0260 1.0851
## occupation.fNon-technical occupations  1.5139 1.4000 1.6370
```

```r
#Logistic regression on hover actions regarding benign links
glm_hover_benign <- glmer(hover_action.f ~
    placement.f #warning-related factors
    + warnings_seen_scaled #number of warnings seen before and during hovering on the link
    + cyber_quiz_score_scaled + PE_score_scaled + brand_usage_scaled # phishing-related individual cha
    + gender.f + education.f + age_scaled + occupation.f #participants demographics
    + (1 | ref_id.f) + (1 | adj_link_id.f), #random effect, dropped username.f for benign link analysi
    data=data_benign_only,
    family=binomial(link=logit),
```

```
    control=glmerControl(optCtrl=list(maxfun=2e4)))
```

singular fit

```
summary(glm_hover_benign)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## hover_action.f ~ placement.f + warnings_seen_scaled + cyber_quiz_score_scaled +
##     PE_score_scaled + brand_usage_scaled + gender.f + education.f +
##     age_scaled + occupation.f + (1 | ref_id.f) + (1 | adj_link_id.f)
##    Data: data_benign_only
## Control: glmerControl(optCtrl = list(maxfun = 20000))
##
##      AIC      BIC   logLik deviance df.resid
##    17416    17536    -8693    17386    22161
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -11.316   0.166   0.286   0.463   3.364
##
## Random effects:
##  Groups        Name        Variance      Std.Dev.
##  adj_link_id.f (Intercept) 0.87910259801 0.9376047
##  ref_id.f      (Intercept) 0.00000000736 0.0000858
## Number of obs: 22176, groups:  adj_link_id.f, 32; ref_id.f, 10
##
## Fixed effects:
##                                   Estimate Std. Error z value
## (Intercept)                        1.66511    0.17866    9.32
## placement.fbanner                  0.47275    0.06152    7.68
## placement.fbrowser                 0.16194    0.05694    2.84
## placement.fno warning              0.31517    0.06346    4.97
## warnings_seen_scaled               0.14822    0.02215    6.69
## cyber_quiz_score_scaled            0.67011    0.01979   33.87
## PE_score_scaled                   -0.14764    0.01824   -8.10
## brand_usage_scaled                 0.00403    0.02268    0.18
## gender.fMale                      -0.01565    0.03958   -0.40
## education.fBachelor's degree      -0.05156    0.04416   -1.17
## education.fGraduate degree        -0.12479    0.06475   -1.93
## age_scaled                        -0.06554    0.01935   -3.39
## occupation.fNon-technical occupations  0.20222    0.05433    3.72
##                                                   Pr(>|z|)
## (Intercept)                  < 0.0000000000000002 ***
## placement.fbanner              0.00000000000001537 ***
## placement.fbrowser                         0.00446 **
## placement.fno warning          0.00000068268762344 ***
## warnings_seen_scaled           0.00000000002221070 ***
## cyber_quiz_score_scaled      < 0.0000000000000002 ***
## PE_score_scaled                0.0000000000000057 ***
## brand_usage_scaled                         0.85885
## gender.fMale                               0.69249
```

```
## education.fBachelor's degree                     0.24298
## education.fGraduate degree                       0.05394 .
## age_scaled                                       0.00071 ***
## occupation.fNon-technical occupations            0.00020 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

## convergence code: 0
## singular fit
```

```
#Print correlation tables
print(glm_hover_benign, correlation=TRUE)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## hover_action.f ~ placement.f + warnings_seen_scaled + cyber_quiz_score_scaled +
##     PE_score_scaled + brand_usage_scaled + gender.f + education.f +
##     age_scaled + occupation.f + (1 | ref_id.f) + (1 | adj_link_id.f)
##   Data: data_benign_only
##     AIC      BIC   logLik deviance df.resid
##   17416    17536    -8693    17386    22161
## Random effects:
##  Groups        Name        Std.Dev.
##  adj_link_id.f (Intercept) 0.9376047
##  ref_id.f      (Intercept) 0.0000858
## Number of obs: 22176, groups:  adj_link_id.f, 32; ref_id.f, 10
## Fixed Effects:
##                         (Intercept)
##                             1.66511
##                   placement.fbanner
##                             0.47275
##                  placement.fbrowser
##                             0.16194
##                placement.fno warning
##                             0.31517
##                warnings_seen_scaled
##                             0.14822
##             cyber_quiz_score_scaled
##                             0.67011
##                     PE_score_scaled
##                            -0.14764
##                  brand_usage_scaled
##                             0.00403
##                         gender.fMale
##                            -0.01565
##          education.fBachelor's degree
##                            -0.05156
##            education.fGraduate degree
##                            -0.12479
```

```
##                               age_scaled
##                                 -0.06554
## occupation.fNon-technical occupations
##                                  0.20222
## convergence code 0; 1 optimizer warnings; 0 lme4 warnings
```

```r
#Get confidence intervals
se_hover_benign <- sqrt(diag(vcov(glm_hover_benign)))
# table of estimates with 95% CI
(tab_hover_benign <- cbind(Est = fixef(glm_hover_benign),
            LL = fixef(glm_hover_benign) - 1.96 * se_hover_benign,
            UL = fixef(glm_hover_benign) + 1.96 * se_hover_benign))
```

```
##                                           Est        LL        UL
## (Intercept)                          1.665109   1.31493   2.015287
## placement.fbanner                    0.472745   0.35217   0.593325
## placement.fbrowser                   0.161941   0.05033   0.273551
## placement.fno warning                0.315173   0.19079   0.439562
## warnings_seen_scaled                 0.148223   0.10480   0.191644
## cyber_quiz_score_scaled              0.670113   0.63133   0.708893
## PE_score_scaled                     -0.147638  -0.18338  -0.111893
## brand_usage_scaled                   0.004033  -0.04042   0.048486
## gender.fMale                        -0.015653  -0.09323   0.061925
## education.fBachelor's degree        -0.051565  -0.13813   0.034997
## education.fGraduate degree          -0.124789  -0.25170   0.002118
## age_scaled                          -0.065541  -0.10347  -0.027608
## occupation.fNon-technical occupations 0.202222   0.09573   0.308714
```

```r
#Odds ratio
exp(tab_hover_benign)
```

```
##                                        Est     LL     UL
## (Intercept)                         5.2862 3.7245 7.5029
## placement.fbanner                   1.6044 1.4221 1.8100
## placement.fbrowser                  1.1758 1.0516 1.3146
## placement.fno warning               1.3705 1.2102 1.5520
## warnings_seen_scaled                1.1598 1.1105 1.2112
## cyber_quiz_score_scaled             1.9545 1.8801 2.0317
## PE_score_scaled                     0.8627 0.8324 0.8941
## brand_usage_scaled                  1.0040 0.9604 1.0497
## gender.fMale                        0.9845 0.9110 1.0639
## education.fBachelor's degree        0.9497 0.8710 1.0356
## education.fGraduate degree          0.8827 0.7775 1.0021
## age_scaled                          0.9366 0.9017 0.9728
## occupation.fNon-technical occupations 1.2241 1.1005 1.3617
```

```r
#RUN THIS MODEL
#Linear regression on hover time regarding benign links
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
##
##     lmer
```

```
## The following object is masked from 'package:stats':
##
##     step

glm_hovertime_benign <- lmerTest::lmer(hover_time_scaled ~
    placement.f #warning-related factors
    + warnings_seen_scaled #number of warnings seen before and during hovering on the link
    + cyber_quiz_score_scaled + PE_score_scaled + brand_usage_scaled # phishing-related individual cha
    + gender.f + education.f + age_scaled + occupation.f #participants demographics
    + (1 | ref_id.f) + (1 | adj_link_id.f), #random effect, dropped username.f for benign link analysi
    data=data_benign_only)
summary(glm_hovertime_benign)

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## hover_time_scaled ~ placement.f + warnings_seen_scaled + cyber_quiz_score_scaled +
##     PE_score_scaled + brand_usage_scaled + gender.f + education.f +
##     age_scaled + occupation.f + (1 | ref_id.f) + (1 | adj_link_id.f)
##    Data: data_benign_only
##
## REML criterion at convergence: 63232
##
## Scaled residuals:
##    Min    1Q Median    3Q    Max
##  -0.94  -0.32  -0.15   0.08  74.87
##
## Random effects:
##  Groups        Name        Variance Std.Dev.
##  adj_link_id.f (Intercept) 0.01135  0.1065
##  ref_id.f      (Intercept) 0.00322  0.0567
##  Residual                  1.00674  1.0034
## Number of obs: 22176, groups:  adj_link_id.f, 32; ref_id.f, 10
##
## Fixed effects:
##                                      Estimate  Std. Error         df
## (Intercept)                          -0.01585     0.03574   22.61511
## placement.fbanner                     0.02753     0.02009 22132.13278
## placement.fbrowser                   -0.02066     0.02007 22132.26858
## placement.fno warning                -0.05192     0.02247 22140.04386
## warnings_seen_scaled                 -0.05420     0.00766 22148.37806
## cyber_quiz_score_scaled               0.07316     0.00698 22132.12644
## PE_score_scaled                      -0.00157     0.00689 22132.00215
## brand_usage_scaled                   -0.00690     0.01002 22132.03350
## gender.fMale                         -0.02697     0.01392 22132.00062
## education.fBachelor's degree         -0.01318     0.01504 22132.00116
## education.fGraduate degree           -0.07806     0.02323 22132.00791
## age_scaled                            0.05939     0.00689 22132.04153
## occupation.fNon-technical occupations 0.07414     0.01894 22132.00223
##                                      t value        Pr(>|t|)
## (Intercept)                            -0.44         0.66154
## placement.fbanner                       1.37         0.17061
## placement.fbrowser                     -1.03         0.30331
## placement.fno warning                  -2.31         0.02084 *
## warnings_seen_scaled                   -7.08  0.0000000000015 ***
```

```
## cyber_quiz_score_scaled              10.48 < 0.0000000000000002 ***
## PE_score_scaled                      -0.23              0.82017
## brand_usage_scaled                   -0.69              0.49137
## gender.fMale                         -1.94              0.05268 .
## education.fBachelor's degree         -0.88              0.38066
## education.fGraduate degree           -3.36              0.00078 ***
## age_scaled                            8.63 < 0.0000000000000002 ***
## occupation.fNon-technical occupations 3.91        0.0000909974090 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)         if you need it
```

```r
#Calculate effect size
r2.corr.mer <- function(m) {
  lmfit <-  lm(model.response(model.frame(m)) ~ fitted(m))
  summary(lmfit)$r.squared
}
r2.corr.mer(glm_hovertime_benign)
```

```
## [1] 0.02748
```

```r
#Now we further filter the dataset to only include entries in the link-focused groups, and examine the
library(dplyr)
benign_inemail_only <- data_benign_only %>%
  filter(data_benign_only$condition_group == 1 | data_benign_only$condition_group == 2 | data_benign_onl
#Factor the activation and forced attention variables
benign_inemail_only$activation.f <- factor(benign_inemail_only$activation,
          level=c("on load","on click"))
benign_inemail_only$forced_attention.f <- factor(benign_inemail_only$forced_attention,
          level=c("no","yes"))
#Rescale continuous variables
library(standardize)
benign_inemail_only$age_rescaled <- scale(as.numeric(benign_inemail_only$DQ_Age))[, 1]
benign_inemail_only$cyber_quiz_score_rescaled <- scale(benign_inemail_only$cyber_quiz_score)[, 1]
benign_inemail_only$PE_score_rescaled <- scale(benign_inemail_only$PE_score)[, 1]
benign_inemail_only$brand_usage_rescaled <- scale(benign_inemail_only$brand_usage)[, 1]
benign_inemail_only$hover_time_rescaled <- scale(benign_inemail_only$hover_time)[, 1]
benign_inemail_only$warnings_seen_rescaled <- scale(benign_inemail_only$warnings_seen)[, 1]
```

```r
#Now we include forced attention and activation as well as their interactions for in-email conditions i
#Logistics regression on benign link click action for link-focused groups
library(lme4)
glm_click_benign_inemail <- glmer(click_action.f ~
    activation.f * forced_attention.f #warning-related factors
    + warnings_seen_rescaled #number of warnings seen before and during clicking on the link
    + cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled # phishing-related individu
    + gender.f + education.f + age_rescaled + occupation.f #participants demographics
    + (1 | ref_id.f) + (1 | adj_link_id.f), #random effect, dropped username.f for benign link analysi
    data=benign_inemail_only,
    family=binomial(link=logit),
    control=glmerControl(optCtrl=list(maxfun=2e4)))
```

```
## singular fit
```

```
summary(glm_click_benign_inemail)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## click_action.f ~ activation.f * forced_attention.f + warnings_seen_rescaled +
##      cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled +
##      gender.f + education.f + age_rescaled + occupation.f + (1 |
##      ref_id.f) + (1 | adj_link_id.f)
##    Data: benign_inemail_only
## Control: glmerControl(optCtrl = list(maxfun = 20000))
##
##      AIC      BIC   logLik deviance df.resid
##    16407    16519    -8189    16377    12657
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -2.154 -0.900 -0.419  0.890  3.428
##
## Random effects:
##  Groups        Name         Variance      Std.Dev.
##  adj_link_id.f (Intercept) 0.23591184417 0.4857076
##  ref_id.f      (Intercept) 0.00000000224 0.0000473
## Number of obs: 12672, groups:  adj_link_id.f, 32; ref_id.f, 10
##
## Fixed effects:
##                                           Estimate Std. Error z value
## (Intercept)                               -0.00519    0.11135   -0.05
## activation.fon click                      -0.07290    0.05335   -1.37
## forced_attention.fyes                     -0.07905    0.05281   -1.50
## warnings_seen_rescaled                    -0.01374    0.01890   -0.73
## cyber_quiz_score_rescaled                  0.26027    0.01950   13.34
## PE_score_rescaled                         -0.01579    0.01918   -0.82
## brand_usage_rescaled                       0.02760    0.01965    1.40
## gender.fMale                              -0.62290    0.03862  -16.13
## education.fBachelor's degree              -0.12820    0.04074   -3.15
## education.fGraduate degree                -0.57122    0.06607   -8.65
## age_rescaled                               0.01764    0.01922    0.92
## occupation.fNon-technical occupations      0.54792    0.05504    9.95
## activation.fon click:forced_attention.fyes 0.12256   0.07516    1.63
##                                                     Pr(>|z|)
## (Intercept)                                           0.9628
## activation.fon click                                  0.1718
## forced_attention.fyes                                 0.1344
## warnings_seen_rescaled                                0.4673
## cyber_quiz_score_rescaled               <0.0000000000000002 ***
## PE_score_rescaled                                     0.4103
## brand_usage_rescaled                                  0.1603
## gender.fMale                            <0.0000000000000002 ***
## education.fBachelor's degree                          0.0016 **
## education.fGraduate degree              <0.0000000000000002 ***
## age_rescaled                                          0.3587
```

```
## occupation.fNon-technical occupations       <0.0000000000000002 ***
## activation.fon click:forced_attention.fyes              0.1030
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##      vcov(x)         if you need it

## convergence code: 0
## singular fit
```

```
#Print correlation tables
print(glm_click_benign_inemail, correlation=TRUE)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
## Family: binomial  ( logit )
## Formula:
## click_action.f ~ activation.f * forced_attention.f + warnings_seen_rescaled +
##     cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled +
##     gender.f + education.f + age_rescaled + occupation.f + (1 |
##     ref_id.f) + (1 | adj_link_id.f)
##    Data: benign_inemail_only
##      AIC      BIC   logLik deviance df.resid
##    16407    16519    -8189    16377    12657
## Random effects:
##  Groups        Name        Std.Dev.
##  adj_link_id.f (Intercept) 0.4857076
##  ref_id.f      (Intercept) 0.0000473
## Number of obs: 12672, groups:  adj_link_id.f, 32; ref_id.f, 10
## Fixed Effects:
##                               (Intercept)
##                                  -0.00519
##                         activation.fon click
##                                  -0.07290
##                         forced_attention.fyes
##                                  -0.07905
##                         warnings_seen_rescaled
##                                  -0.01374
##                      cyber_quiz_score_rescaled
##                                   0.26027
##                              PE_score_rescaled
##                                  -0.01579
##                           brand_usage_rescaled
##                                   0.02760
##                                 gender.fMale
##                                  -0.62290
##                     education.fBachelor's degree
##                                  -0.12820
##                      education.fGraduate degree
##                                  -0.57122
##                                 age_rescaled
##                                   0.01764
##      occupation.fNon-technical occupations
```

```
##                                  0.54792
## activation.fon click:forced_attention.fyes
##                                  0.12256
## convergence code 0; 1 optimizer warnings; 0 lme4 warnings
```

```r
#Get confidence intervals
se_click_benign <- sqrt(diag(vcov(glm_click_benign_inemail)))
# table of estimates with 95% CI
(tab_click_benign <- cbind(Est = fixef(glm_click_benign_inemail),
            LL = fixef(glm_click_benign_inemail) - 1.96 * se_click_benign,
            UL = fixef(glm_click_benign_inemail) + 1.96 * se_click_benign))
```

```
##                                                Est        LL        UL
## (Intercept)                              -0.005187 -0.22343   0.21305
## activation.fon click                     -0.072897 -0.17746   0.03166
## forced_attention.fyes                    -0.079046 -0.18255   0.02446
## warnings_seen_rescaled                   -0.013741 -0.05079   0.02331
## cyber_quiz_score_rescaled                 0.260273  0.22204   0.29850
## PE_score_rescaled                        -0.015791 -0.05338   0.02180
## brand_usage_rescaled                      0.027596 -0.01092   0.06611
## gender.fMale                             -0.622904 -0.69860  -0.54721
## education.fBachelor's degree             -0.128199 -0.20804  -0.04836
## education.fGraduate degree               -0.571215 -0.70072  -0.44171
## age_rescaled                              0.017642 -0.02003   0.05532
## occupation.fNon-technical occupations     0.547923  0.44004   0.65581
## activation.fon click:forced_attention.fyes  0.122562 -0.02475   0.26988
```

```r
#Odds ratio
exp(tab_click_benign)
```

```
##                                               Est     LL     UL
## (Intercept)                                0.9948 0.7998 1.2374
## activation.fon click                       0.9297 0.8374 1.0322
## forced_attention.fyes                      0.9240 0.8331 1.0248
## warnings_seen_rescaled                     0.9864 0.9505 1.0236
## cyber_quiz_score_rescaled                  1.2973 1.2486 1.3478
## PE_score_rescaled                          0.9843 0.9480 1.0220
## brand_usage_rescaled                       1.0280 0.9891 1.0683
## gender.fMale                               0.5364 0.4973 0.5786
## education.fBachelor's degree               0.8797 0.8122 0.9528
## education.fGraduate degree                 0.5648 0.4962 0.6429
## age_rescaled                               1.0178 0.9802 1.0569
## occupation.fNon-technical occupations      1.7297 1.5528 1.9267
## activation.fon click:forced_attention.fyes 1.1304 0.9756 1.3098
```

```r
#Logistic regression on benign link hover actions for link-focused groups
library(optimx)
glm_hover_benign_inemail <- glmer(hover_action.f ~
    activation.f * forced_attention.f #warning-related factors
    + warnings_seen_rescaled #number of warnings seen before and during hovering on the link
    + cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled # phishing-related individu
    + gender.f + education.f + age_rescaled + occupation.f #participants demographics
    + (1 | ref_id.f) + (1 | adj_link_id.f), #random effect, dropped username.f for benign link analysi
    data=benign_inemail_only,
    family=binomial(link=logit),
    REML = FALSE,
```

```
    control = glmerControl(optimizer ='optimx', optCtrl=list(method='L-BFGS-B')))
```

Warning: extra argument(s) 'REML' disregarded

Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
control$checkConv, : Model failed to converge with max|grad| = 0.00841722
(tol = 0.001, component 1)

```
summary(glm_hover_benign_inemail)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## hover_action.f ~ activation.f * forced_attention.f + warnings_seen_rescaled +
##     cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled +
##     gender.f + education.f + age_rescaled + occupation.f + (1 |
##     ref_id.f) + (1 | adj_link_id.f)
##    Data: benign_inemail_only
## Control:
## glmerControl(optimizer = "optimx", optCtrl = list(method = "L-BFGS-B"))
##
##      AIC      BIC   logLik deviance df.resid
##    10011    10123    -4990     9981    12657
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -8.108  0.153  0.278  0.463  3.431
##
## Random effects:
##  Groups        Name        Variance Std.Dev.
##  adj_link_id.f (Intercept) 0.9354   0.967
##  ref_id.f      (Intercept) 0.0877   0.296
## Number of obs: 12672, groups:  adj_link_id.f, 32; ref_id.f, 10
##
## Fixed effects:
##                                          Estimate Std. Error z value
## (Intercept)                               1.86789    0.22939    8.14
## activation.fon click                      0.40438    0.07419    5.45
## forced_attention.fyes                    -0.15226    0.06910   -2.20
## warnings_seen_rescaled                    0.18529    0.02630    7.05
## cyber_quiz_score_rescaled                 0.69337    0.02603   26.64
## PE_score_rescaled                        -0.12118    0.02412   -5.02
## brand_usage_rescaled                      0.00801    0.01966    0.41
## gender.fMale                              0.01323    0.05183    0.26
## education.fBachelor's degree             -0.17773    0.05695   -3.12
## education.fGraduate degree               -0.26587    0.08408   -3.16
## age_rescaled                             -0.15043    0.02547   -5.91
## occupation.fNon-technical occupations     0.06059    0.07616    0.80
## activation.fon click:forced_attention.fyes -0.09411  0.10266   -0.92
##                                          Pr(>|z|)
## (Intercept)                              0.0000000000000039 ***
## activation.fon click                     0.00000005010812728 ***
## forced_attention.fyes                                 0.0276 *
## warnings_seen_rescaled                   0.0000000000184309 ***
```

18

```
## cyber_quiz_score_rescaled                    < 0.0000000000000002 ***
## PE_score_rescaled                            0.00000050606517823 ***
## brand_usage_rescaled                                       0.6838
## gender.fMale                                               0.7985
## education.fBachelor's degree                               0.0018 **
## education.fGraduate degree                                 0.0016 **
## age_rescaled                                   0.00000000349619824 ***
## occupation.fNon-technical occupations                      0.4263
## activation.fon click:forced_attention.fyes                0.3593
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)         if you need it

## convergence code: 0
## Model failed to converge with max|grad| = 0.00841722 (tol = 0.001, component 1)
```

```r
#Print correlation tables
print(glm_hover_benign_inemail, correlation=TRUE)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## hover_action.f ~ activation.f * forced_attention.f + warnings_seen_rescaled +
##     cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled +
##     gender.f + education.f + age_rescaled + occupation.f + (1 |
##     ref_id.f) + (1 | adj_link_id.f)
##    Data: benign_inemail_only
##      AIC      BIC   logLik deviance df.resid
##    10011    10123    -4990     9981    12657
## Random effects:
##  Groups        Name        Std.Dev.
##  adj_link_id.f (Intercept) 0.967
##  ref_id.f      (Intercept) 0.296
## Number of obs: 12672, groups:  adj_link_id.f, 32; ref_id.f, 10
## Fixed Effects:
##                             (Intercept)
##                                 1.86789
##                        activation.fon click
##                                 0.40438
##                        forced_attention.fyes
##                                -0.15226
##                        warnings_seen_rescaled
##                                 0.18529
##                     cyber_quiz_score_rescaled
##                                 0.69337
##                            PE_score_rescaled
##                                -0.12118
##                        brand_usage_rescaled
##                                 0.00801
##                                gender.fMale
##                                 0.01323
```

```
##              education.fBachelor's degree
##                            -0.17773
##              education.fGraduate degree
##                            -0.26587
##                         age_rescaled
##                            -0.15043
##      occupation.fNon-technical occupations
##                             0.06059
## activation.fon click:forced_attention.fyes
##                            -0.09411
## convergence code 0; 1 optimizer warnings; 0 lme4 warnings
```

```
#Get confidence intervals
se_hover_phish_inemail <- sqrt(diag(vcov(glm_hover_benign_inemail)))
# table of estimates with 95% CI
(tab_hover_phish_inemail <- cbind(Est = fixef(glm_hover_benign_inemail),
          LL = fixef(glm_hover_benign_inemail) - 1.96 * se_hover_phish_inemail,
          UL = fixef(glm_hover_benign_inemail) + 1.96 * se_hover_phish_inemail))
```

```
##                                             Est       LL       UL
## (Intercept)                            1.867885  1.41829  2.31748
## activation.fon click                   0.404378  0.25897  0.54978
## forced_attention.fyes                 -0.152263 -0.28770 -0.01683
## warnings_seen_rescaled                 0.185287  0.13374  0.23683
## cyber_quiz_score_rescaled              0.693367  0.64236  0.74438
## PE_score_rescaled                     -0.121182 -0.16846 -0.07391
## brand_usage_rescaled                   0.008006 -0.03053  0.04654
## gender.fMale                           0.013231 -0.08835  0.11481
## education.fBachelor's degree          -0.177726 -0.28935 -0.06610
## education.fGraduate degree            -0.265868 -0.43067 -0.10106
## age_rescaled                          -0.150431 -0.20035 -0.10051
## occupation.fNon-technical occupations  0.060587 -0.08868  0.20985
## activation.fon click:forced_attention.fyes -0.094109 -0.29532  0.10710
```

```
#Odds ratio
exp(tab_hover_phish_inemail)
```

```
##                                        Est     LL      UL
## (Intercept)                         6.4746 4.1301 10.1501
## activation.fon click                1.4984 1.2956  1.7329
## forced_attention.fyes               0.8588 0.7500  0.9833
## warnings_seen_rescaled              1.2036 1.1431  1.2672
## cyber_quiz_score_rescaled           2.0004 1.9010  2.1051
## PE_score_rescaled                   0.8859 0.8450  0.9288
## brand_usage_rescaled                1.0080 0.9699  1.0476
## gender.fMale                        1.0133 0.9154  1.1217
## education.fBachelor's degree        0.8372 0.7487  0.9360
## education.fGraduate degree          0.7665 0.6501  0.9039
## age_rescaled                        0.8603 0.8184  0.9044
## occupation.fNon-technical occupations 1.0625 0.9151  1.2335
## activation.fon click:forced_attention.fyes 0.9102 0.7443  1.1130
```

```
#RUN THIS MODEL
#Linear regression on benign link hover time for link-focused groups
library(lmerTest)
glm_hovertime_benign_inemail <- lmerTest::lmer(hover_time_rescaled ~
```

```
    activation.f * forced_attention.f #warning-related factors
    + warnings_seen_rescaled #number of warnings seen before and during hovering on the link
    + cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled # phishing-related individu
    + gender.f + education.f + age_rescaled + occupation.f #participants demographics
    + (1 | ref_id.f) + (1 | adj_link_id.f), #random effect, dropped username.f for benign link analysi
    data=benign_inemail_only)
summary(glm_hovertime_benign_inemail)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## hover_time_rescaled ~ activation.f * forced_attention.f + warnings_seen_rescaled +
##     cyber_quiz_score_rescaled + PE_score_rescaled + brand_usage_rescaled +
##     gender.f + education.f + age_rescaled + occupation.f + (1 |
##     ref_id.f) + (1 | adj_link_id.f)
##    Data: benign_inemail_only
##
## REML criterion at convergence: 35925
##
## Scaled residuals:
##    Min    1Q Median    3Q    Max
##  -0.98  -0.31  -0.15   0.09  73.00
##
## Random effects:
##  Groups        Name        Variance Std.Dev.
##  adj_link_id.f (Intercept) 0.01299  0.1140
##  ref_id.f      (Intercept) 0.00217  0.0466
##  Residual                  0.98673  0.9933
## Number of obs: 12672, groups:  adj_link_id.f, 32; ref_id.f, 10
##
## Fixed effects:
##                                            Estimate  Std. Error
## (Intercept)                                -0.08434     0.04216
## activation.fon click                        0.11951     0.02520
## forced_attention.fyes                       0.01885     0.02497
## warnings_seen_rescaled                     -0.05237     0.00893
## cyber_quiz_score_rescaled                   0.06144     0.00901
## PE_score_rescaled                           0.01751     0.00891
## brand_usage_rescaled                        0.00321     0.00885
## gender.fMale                                0.00289     0.01817
## education.fBachelor's degree               -0.02864     0.01929
## education.fGraduate degree                 -0.02494     0.03065
## age_rescaled                                0.05242     0.00911
## occupation.fNon-technical occupations       0.07933     0.02549
## activation.fon click:forced_attention.fyes -0.12113     0.03549
##                                                  df t value
## (Intercept)                                45.22007   -2.00
## activation.fon click                    12628.00225    4.74
## forced_attention.fyes                   12628.00325    0.75
## warnings_seen_rescaled                  12639.02841   -5.87
## cyber_quiz_score_rescaled               12628.10299    6.82
## PE_score_rescaled                       12628.00715    1.96
## brand_usage_rescaled                    12630.81743    0.36
## gender.fMale                            12628.01970    0.16
```

```
## education.fBachelor's degree              12628.02163    -1.48
## education.fGraduate degree                12628.00187    -0.81
## age_rescaled                              12628.08502     5.75
## occupation.fNon-technical occupations     12628.00207     3.11
## activation.fon click:forced_attention.fyes 12628.01427   -3.41
##                                                 Pr(>|t|)
## (Intercept)                                       0.05147 .
## activation.fon click                        0.0000021310514 ***
## forced_attention.fyes                             0.45036
## warnings_seen_rescaled                      0.0000000045883 ***
## cyber_quiz_score_rescaled                   0.0000000000096 ***
## PE_score_rescaled                                 0.04948 *
## brand_usage_rescaled                              0.71724
## gender.fMale                                      0.87364
## education.fBachelor's degree                      0.13771
## education.fGraduate degree                        0.41573
## age_rescaled                                0.0000000089102 ***
## occupation.fNon-technical occupations             0.00186 **
## activation.fon click:forced_attention.fyes       0.00064 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```r
#Calculate effect size
r2.corr.mer(glm_hovertime_benign_inemail)
```

```
## [1] 0.02957
```