R Notebook

Author: Joss Ives, *updated by Joss Ives 2018 May 27, 16:28:05*.

# Setup

## Load libraries

suppressWarnings(library(lme4))

## Loading required package: Matrix

require(lme4)  
suppressWarnings(library(ggplot2))  
require(ggplot2)  
source("jossfunc.R")

## Data files

### Load data

dat.raw <- read.csv("C:/Users/Joss/ownCloud/Shared/DP\_Derived\_Data/dpLogisticDat.csv")

### Add additional calculated values

dat.raw$course.grade.frac <- dat.raw$course.grade/100.  
dat.raw$CRT.medsplit <- trunc(dat.raw$NCRT/2)  
dat.raw$final.grade.LMH <- as.integer(  
 cut(dat.raw$f.tot78, quantile(dat.raw$f.tot78, probs=0:3/3), include.lowest=TRUE)  
 )  
dat.raw$final.grade.fix <- (dat.raw$f.tot78-2.\*dat.raw$QCORRECT)/76.  
dat.raw$final.gradeA.fix <- (dat.raw$f.Atot40-2.\*dat.raw$QCORRECT)/38.  
  
names(dat.raw)

## [1] "ID" "QNUM" "QCORRECT"   
## [4] "TREATMENT" "f.Atot40" "f.Btot38"   
## [7] "f.tot78" "course.grade" "d.version"   
## [10] "f.version" "NCRT" "Gender"   
## [13] "course.grade.frac" "CRT.medsplit" "final.grade.LMH"   
## [16] "final.grade.fix" "final.gradeA.fix"

### Set categorical variables

dat.raw$ID <- factor(dat.raw$ID)  
dat.raw$QNUM <- factor(dat.raw$QNUM)  
dat.raw$TREATMENT <- factor(dat.raw$TREATMENT)

### Make some data subsets

# Keep only those data points where TREATMENT was set  
dat.all <- subset(dat.raw, TREATMENT==0 | TREATMENT==1)  
  
# Look only at the 4 questions that had TREATMENT  
dat.trt <- subset(dat.all, QNUM==5 | QNUM==6 | QNUM==9 | QNUM==10)  
names(dat.trt)

## [1] "ID" "QNUM" "QCORRECT"   
## [4] "TREATMENT" "f.Atot40" "f.Btot38"   
## [7] "f.tot78" "course.grade" "d.version"   
## [10] "f.version" "NCRT" "Gender"   
## [13] "course.grade.frac" "CRT.medsplit" "final.grade.LMH"   
## [16] "final.grade.fix" "final.gradeA.fix"

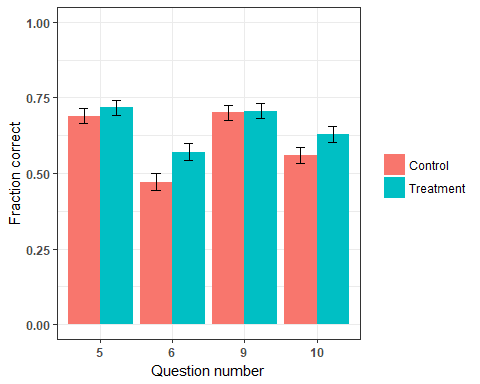
# Initial analyses on 4 EYA questions

### What do the results look like for each question?

dodge=position\_dodge(width=0.9)  
  
corr.by.question <- summarySE(  
 dat.trt, measurevar="QCORRECT", groupvars=c("TREATMENT","QNUM")  
 )

## Loading required package: plyr

limits <- aes(ymax = QCORRECT + binomial.error, ymin = QCORRECT - binomial.error)  
  
ggplot(corr.by.question, aes(x=QNUM, y=QCORRECT, fill=TREATMENT)) +  
 geom\_bar(stat="identity", position=dodge) +   
 geom\_errorbar(limits, position=dodge, width=0.25) +  
 labs(x = "Question number", y = "Fraction correct") +  
 scale\_fill\_discrete(labels=c("Control","Treatment")) +  
 theme\_bw() +   
 theme(axis.text = element\_text(face = "bold")) +   
 #theme(legend.position="none")  
 scale\_y\_continuous(limits = c(0,1)) +  
 guides(fill=guide\_legend(title=NULL))



#ggtitle("All")

### Fisher’s Exact for overall question set

## fisher.2vector  
# 2 vectors of 0s and 1s are passed  
# \* The first is control (top row)  
# \* The second is treatement (bottom row)  
# Adding parenthesis around assigning a variable causes the output to be displayed  
(result <- fisher.2vector(  
 dat.trt$QCORRECT[dat.trt$TREATMENT==0],  
 dat.trt$QCORRECT[dat.trt$TREATMENT==1]  
))

##   
## Fisher's Exact Test for Count Data  
##   
## data: ctable  
## p-value = 0.008497  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
## 1.055850 1.460482  
## sample estimates:  
## odds ratio   
## 1.241638

cat("Cohen's d\n ", cohens.d.from.odds.simple(result$estimate[[1]]),"\n")

## Cohen's d  
## 0.1193251

### Simple logistic regression

m <- glmer(QCORRECT ~ QNUM + TREATMENT + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
print(summary(m))

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3339.2 3374.4 -1663.6 3327.2 2614   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.9040 -0.9880 0.5252 0.6962 1.3870   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.6221 0.7887   
## Number of obs: 2620, groups: ID, 658  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.86542 0.10632 8.140 3.95e-16 \*\*\*  
## QNUM6 -0.89456 0.12503 -7.155 8.39e-13 \*\*\*  
## QNUM9 0.00195 0.12799 0.015 0.98784   
## QNUM10 -0.55054 0.12483 -4.410 1.03e-05 \*\*\*  
## TREATMENT1 0.24798 0.08771 2.827 0.00469 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10  
## QNUM6 -0.641   
## QNUM9 -0.603 0.511   
## QNUM10 -0.632 0.539 0.513   
## TREATMENT1 -0.389 -0.017 0.002 -0.010

#  
se <- sqrt(diag(vcov(m)))  
tab <- cbind(Est = fixef(m), LL = fixef(m) - 1.96 \* se, UL = fixef(m) + 1.96 \*se)  
exp(tab)

## Est LL UL  
## (Intercept) 2.3759910 1.9290626 2.9264644  
## QNUM6 0.4087861 0.3199384 0.5223071  
## QNUM9 1.0019522 0.7796460 1.2876461  
## QNUM10 0.5766358 0.4514868 0.7364754  
## TREATMENT1 1.2814361 1.0790391 1.5217969

### A quick summary of odds ratios so far

odds.table <- matrix(c(result$estimate[[1]],exp(tab[[5]])),ncol=1,byrow=TRUE)  
colnames(odds.table) <- c("Odds Ratio")  
rownames(odds.table) <- c("Fisher's Exact Test","Logistic Regression")  
as.table(odds.table)

## Odds Ratio  
## Fisher's Exact Test 1.241638  
## Logistic Regression 1.281436

# Digging into the logistic regressions a bit more

### QNUM as a random effect

First, let’s look at making QNUM a random effect. Not sure if it is correct, but it makes no difference to the results and cleans up the reporting.

m.2re <- glmer(QCORRECT ~ (1|QNUM) + TREATMENT + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.2re)$coefficients

## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 0.5001450 0.19507001 2.563926 0.010349569  
## TREATMENT1 0.2472692 0.08704356 2.840752 0.004500724

summary(m.2re)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ (1 | QNUM) + TREATMENT + (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3352.2 3375.6 -1672.1 3344.2 2616   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.8220 -0.9899 0.5489 0.7015 1.3533   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.6023 0.7761   
## QNUM (Intercept) 0.1333 0.3652   
## Number of obs: 2620, groups: ID, 658; QNUM, 4  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.50015 0.19507 2.564 0.0103 \*   
## TREATMENT1 0.24727 0.08704 2.841 0.0045 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## TREATMENT1 -0.215

### Fitting using all the questions

This makes the AIC and BIC quadruple and doesn’t provide a significant change in TREATMENT1.

#m.allQs <- glmer(QCORRECT ~ QNUM + TREATMENT + (1|ID),   
# data = dat.all,   
# family = binomial, control=glmerControl(optimizer="bobyqa"))  
#summary(m.allQs)

#m.allQs.2re <- glmer(QCORRECT ~ (1|QNUM) + TREATMENT + (1|ID),   
# data = dat.all,   
# family = binomial, control=glmerControl(optimizer="bobyqa"))  
#summary(m.allQs.2re)

# Looking at other variables with only the treatment questions

## Course grade

Course grade and NCRT both are significant and improve the model slightly, but do not have a significant impact on TREATMENT1

m.grade <- glmer(QCORRECT ~ QNUM + TREATMENT + course.grade.frac + NCRT + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.grade)

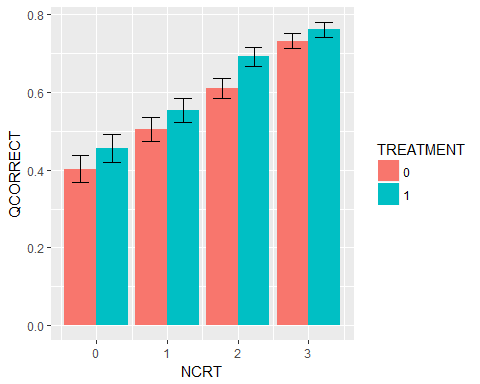
## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + course.grade.frac + NCRT + (1 |   
## ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3041.2 3088.1 -1512.6 3025.2 2612   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.9016 -0.8851 0.4482 0.6771 2.7981   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.0198 0.1407   
## Number of obs: 2620, groups: ID, 658  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.422773 0.284928 -12.013 < 2e-16 \*\*\*  
## QNUM6 -0.904943 0.126218 -7.170 7.52e-13 \*\*\*  
## QNUM9 0.003465 0.130115 0.027 0.97875   
## QNUM10 -0.557258 0.126332 -4.411 1.03e-05 \*\*\*  
## TREATMENT1 0.244800 0.088502 2.766 0.00567 \*\*   
## course.grade.frac 4.993186 0.383692 13.014 < 2e-16 \*\*\*  
## NCRT 0.305673 0.042914 7.123 1.06e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATM crs.g.  
## QNUM6 -0.129   
## QNUM9 -0.231 0.515   
## QNUM10 -0.173 0.544 0.515   
## TREATMENT1 -0.183 -0.007 0.003 0.000   
## crs.grd.frc -0.900 -0.100 0.003 -0.059 0.029   
## NCRT -0.044 -0.054 -0.002 -0.034 0.020 -0.218

# Time to dig into NCRT

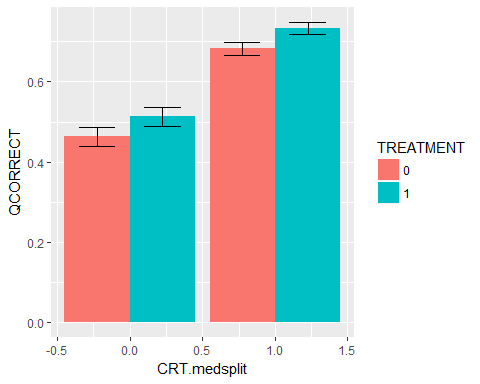
## Graphs

### EYA performance vs NCRT and median split

s.ncrt <- summarySE(dat.trt, measurevar="QCORRECT",  
 groupvars=c("NCRT","TREATMENT"), .drop=FALSE)  
limits <- aes(ymax = QCORRECT + binomial.error, ymin = QCORRECT - binomial.error)  
#options(repr.plot.width=5, repr.plot.height=3)  
  
ggplot(s.ncrt, aes(x=NCRT, y=QCORRECT, fill=TREATMENT)) +   
 geom\_bar(stat="identity", position=dodge) +   
 geom\_errorbar(limits, position=dodge, width=0.5)

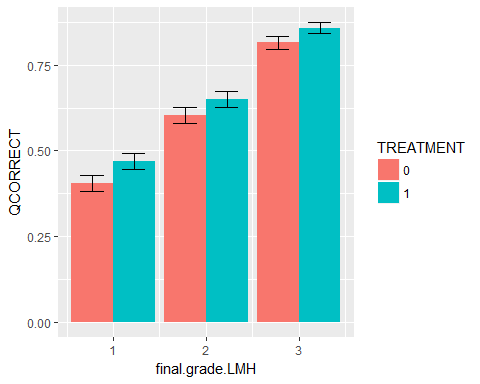


s.crtsplit <- summarySE(dat.trt, measurevar="QCORRECT",  
 groupvars=c("CRT.medsplit","TREATMENT"), .drop=FALSE)  
limits <- aes(ymax = QCORRECT + binomial.error, ymin = QCORRECT - binomial.error)  
#options(repr.plot.width=5, repr.plot.height=3)  
  
ggplot(s.crtsplit, aes(x=CRT.medsplit, y=QCORRECT, fill=TREATMENT)) +   
 geom\_bar(stat="identity", position=dodge) +   
 geom\_errorbar(limits, position=dodge, width=0.5)



# And looking at correlations with final exam grades

s.fLMH <- summarySE(dat.trt, measurevar="QCORRECT",  
 groupvars=c("final.grade.LMH","TREATMENT"), .drop=FALSE)  
limits <- aes(ymax = QCORRECT + binomial.error, ymin = QCORRECT - binomial.error)  
#options(repr.plot.width=5, repr.plot.height=3)  
  
ggplot(s.fLMH, aes(x=final.grade.LMH, y=QCORRECT, fill=TREATMENT)) +   
 geom\_bar(stat="identity", position=dodge) +   
 geom\_errorbar(limits, position=dodge, width=0.5)

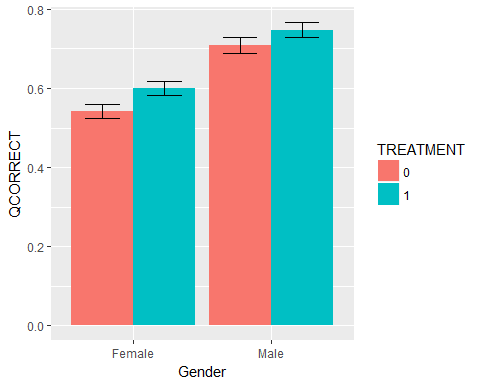


# Finally, let’s look at gender

### Bar graph splitting on gender and looking at the effect of treatment

Splitting on gender shows that there may be a significant difference in who the intervention helps

sum.gender <- summarySE(dat.trt, measurevar="QCORRECT",  
 groupvars=c("Gender","TREATMENT"), .drop=FALSE)  
  
limits <- aes(ymax = QCORRECT + binomial.error, ymin = QCORRECT - binomial.error)  
  
ggplot(sum.gender, aes(x=Gender, y=QCORRECT, fill=TREATMENT)) +   
 geom\_bar(stat="identity", position=dodge) +   
 geom\_errorbar(limits, position=dodge, width=0.5)



### Logistic regression with gender added

Gender is significant.

m.gender0 <- glmer(QCORRECT ~ QNUM + TREATMENT + Gender + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.gender0)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + Gender + (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3288.5 3329.5 -1637.2 3274.5 2613   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.2496 -0.9548 0.5067 0.6997 1.4685   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.4913 0.7009   
## Number of obs: 2620, groups: ID, 658  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.567639 0.110691 5.128 2.93e-07 \*\*\*  
## QNUM6 -0.894988 0.125099 -7.154 8.41e-13 \*\*\*  
## QNUM9 0.000603 0.128207 0.005 0.99625   
## QNUM10 -0.552625 0.124938 -4.423 9.72e-06 \*\*\*  
## TREATMENT1 0.250017 0.087757 2.849 0.00439 \*\*   
## GenderMale 0.793746 0.109970 7.218 5.28e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATM  
## QNUM6 -0.596   
## QNUM9 -0.579 0.512   
## QNUM10 -0.595 0.540 0.513   
## TREATMENT1 -0.381 -0.018 0.002 -0.012   
## GenderMale -0.310 -0.055 -0.001 -0.034 0.024

### Run logistic regressions on the two gender populations seperately

First without course grades included. Treatment is only significant for females, not males.

m.female <- glmer(QCORRECT ~ QNUM + TREATMENT + (1|ID),   
 data = subset(dat.trt, Gender=="Female"),   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.female)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + (1 | ID)  
## Data: subset(dat.trt, Gender == "Female")  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 2161.7 2194.1 -1074.9 2149.7 1624   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -1.6950 -0.9826 0.5855 0.7996 1.5265   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.4497 0.6706   
## Number of obs: 1630, groups: ID, 409  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.54798 0.12555 4.365 1.27e-05 \*\*\*  
## QNUM6 -0.97894 0.15394 -6.359 2.03e-10 \*\*\*  
## QNUM9 -0.01544 0.15396 -0.100 0.92011   
## QNUM10 -0.45561 0.15200 -2.997 0.00272 \*\*   
## TREATMENT1 0.28553 0.10750 2.656 0.00791 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10  
## QNUM6 -0.620   
## QNUM9 -0.614 0.500   
## QNUM10 -0.621 0.521 0.507   
## TREATMENT1 -0.395 -0.045 0.001 -0.029

m<-m.female  
se <- sqrt(diag(vcov(m)))  
tab <- cbind(Est = fixef(m), LL = fixef(m) - 1.96 \* se, UL = fixef(m) + 1.96 \*se)  
exp(tab)

## Est LL UL  
## (Intercept) 1.7297536 1.3524364 2.2123388  
## QNUM6 0.3757092 0.2778549 0.5080255  
## QNUM9 0.9846771 0.7281771 1.3315293  
## QNUM10 0.6340631 0.4707062 0.8541124  
## TREATMENT1 1.3304630 1.0776980 1.6425120

m.male <- glmer(QCORRECT ~ QNUM + TREATMENT + (1|ID),   
 data = subset(dat.trt, Gender=="Male"),   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.male)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + (1 | ID)  
## Data: subset(dat.trt, Gender == "Male")  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 1131.6 1161.0 -559.8 1119.6 984   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.2693 -0.9380 0.4490 0.5734 1.0613   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.6002 0.7747   
## Number of obs: 990, groups: ID, 249  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.42031 0.19600 7.246 4.28e-13 \*\*\*  
## QNUM6 -0.75169 0.21979 -3.420 0.000626 \*\*\*  
## QNUM9 0.03745 0.23347 0.160 0.872557   
## QNUM10 -0.74281 0.21958 -3.383 0.000717 \*\*\*  
## TREATMENT1 0.17833 0.15375 1.160 0.246091   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10  
## QNUM6 -0.674   
## QNUM9 -0.591 0.527   
## QNUM10 -0.672 0.575 0.528   
## TREATMENT1 -0.392 0.032 0.005 0.028

m<-m.male  
se <- sqrt(diag(vcov(m)))  
tab <- cbind(Est = fixef(m), LL = fixef(m) - 1.96 \* se, UL = fixef(m) + 1.96 \*se)  
exp(tab)

## Est LL UL  
## (Intercept) 4.1384133 2.8183316 6.0768095  
## QNUM6 0.4715667 0.3065180 0.7254882  
## QNUM9 1.0381613 0.6569490 1.6405823  
## QNUM10 0.4757749 0.3093769 0.7316696  
## TREATMENT1 1.1952190 0.8842524 1.6155439

When controlling for course grade, nothing changes

m.female.grade <- glmer(QCORRECT ~ QNUM + TREATMENT + course.grade.frac + (1|ID),   
 data = subset(dat.trt,Gender=="Female"),   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.female.grade)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + course.grade.frac + (1 | ID)  
## Data: subset(dat.trt, Gender == "Female")  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 2030.9 2068.7 -1008.5 2016.9 1623   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.1171 -0.9385 0.4891 0.7533 2.4850   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.06234 0.2497   
## Number of obs: 1630, groups: ID, 409  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.2475 0.3539 -9.177 < 2e-16 \*\*\*  
## QNUM6 -0.9817 0.1543 -6.363 1.98e-10 \*\*\*  
## QNUM9 -0.0142 0.1553 -0.091 0.92714   
## QNUM10 -0.4558 0.1528 -2.983 0.00285 \*\*   
## TREATMENT1 0.2801 0.1079 2.596 0.00944 \*\*   
## course.grade.frac 5.2249 0.4699 11.119 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATM  
## QNUM6 -0.114   
## QNUM9 -0.219 0.503   
## QNUM10 -0.174 0.523 0.508   
## TREATMENT1 -0.176 -0.035 0.001 -0.020   
## crs.grd.frc -0.938 -0.114 -0.001 -0.051 0.035

m.male.grade <- glmer(QCORRECT ~ QNUM + TREATMENT + course.grade.frac + (1|ID),   
 data = subset(dat.trt,Gender=="Male"),   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.male.grade)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + course.grade.frac + (1 | ID)  
## Data: subset(dat.trt, Gender == "Male")  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 1032.8 1067.1 -509.4 1018.8 983   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.5741 -0.5821 0.4021 0.5685 1.9337   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.03624 0.1904   
## Number of obs: 990, groups: ID, 249  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.3260 0.5009 -6.640 3.13e-11 \*\*\*  
## QNUM6 -0.7723 0.2232 -3.461 0.000539 \*\*\*  
## QNUM9 0.0467 0.2383 0.196 0.844641   
## QNUM10 -0.7580 0.2230 -3.399 0.000676 \*\*\*  
## TREATMENT1 0.1741 0.1555 1.119 0.263030   
## course.grade.frac 6.1641 0.6521 9.453 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATM  
## QNUM6 -0.166   
## QNUM9 -0.247 0.528   
## QNUM10 -0.170 0.581 0.529   
## TREATMENT1 -0.180 0.044 0.006 0.040   
## crs.grd.frc -0.926 -0.104 0.011 -0.099 0.020

### Looking at interaction term

m.gender <- glmer(QCORRECT ~ QNUM + TREATMENT\*Gender + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.gender)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT \* Gender + (1 | ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3290.1 3337.1 -1637.0 3274.1 2612   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.2060 -0.9461 0.4972 0.6930 1.4834   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0.492 0.7014   
## Number of obs: 2620, groups: ID, 658  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.5505355 0.1143230 4.816 1.47e-06 \*\*\*  
## QNUM6 -0.8972681 0.1251816 -7.168 7.62e-13 \*\*\*  
## QNUM9 0.0004719 0.1282263 0.004 0.99706   
## QNUM10 -0.5546810 0.1250070 -4.437 9.11e-06 \*\*\*  
## TREATMENT1 0.2870489 0.1076585 2.666 0.00767 \*\*   
## GenderMale 0.8476396 0.1425759 5.945 2.76e-09 \*\*\*  
## TREATMENT1:GenderMale -0.1110428 0.1858459 -0.597 0.55017   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATMENT1 GndrMl  
## QNUM6 -0.568   
## QNUM9 -0.561 0.512   
## QNUM10 -0.568 0.540 0.513   
## TREATMENT1 -0.445 -0.034 0.001 -0.026   
## GenderMale -0.389 -0.064 -0.002 -0.045 0.380   
## TREATMENT1: 0.250 0.033 0.002 0.029 -0.579 -0.636

m<-m.gender  
se <- sqrt(diag(vcov(m)))  
tab <- cbind(Est = fixef(m), LL = fixef(m) - 1.96 \* se, UL = fixef(m) + 1.96 \*se)  
exp(tab)

## Est LL UL  
## (Intercept) 1.7341814 1.3860560 2.1697428  
## QNUM6 0.4076819 0.3189809 0.5210485  
## QNUM9 1.0004720 0.7781378 1.2863328  
## QNUM10 0.5742554 0.4494662 0.7336911  
## TREATMENT1 1.3324893 1.0790039 1.6455250  
## GenderMale 2.3341308 1.7650709 3.0866560  
## TREATMENT1:GenderMale 0.8949005 0.6216980 1.2881605

### Control for overall exam score (yes, re-calcuated for each question to not include the specific question in the score)

Treatment does not change significantly, but when controlling for final exam score the gender effect goes down significantly

m.fgradeA <- glmer(QCORRECT ~ QNUM + TREATMENT + final.gradeA.fix + Gender + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.fgradeA)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + final.gradeA.fix + Gender + (1 |   
## ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3143.2 3190.1 -1563.6 3127.2 2612   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.3088 -0.9548 0.4817 0.7213 2.3850   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## ID (Intercept) 0 0   
## Number of obs: 2620, groups: ID, 658  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.42511 0.17934 -7.947 1.92e-15 \*\*\*  
## QNUM6 -0.89729 0.12255 -7.322 2.45e-13 \*\*\*  
## QNUM9 0.00104 0.12683 0.008 0.99346   
## QNUM10 -0.55284 0.12312 -4.490 7.12e-06 \*\*\*  
## TREATMENT1 0.24743 0.08636 2.865 0.00417 \*\*   
## final.gradeA.fix 3.19745 0.24744 12.922 < 2e-16 \*\*\*  
## GenderMale 0.49859 0.09281 5.372 7.79e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATM fnl.A.  
## QNUM6 -0.281   
## QNUM9 -0.355 0.518   
## QNUM10 -0.315 0.540 0.515   
## TREATMENT1 -0.258 -0.008 0.003 -0.003   
## fnl.grdA.fx -0.818 -0.094 0.001 -0.054 0.026   
## GenderMale -0.032 -0.030 -0.002 -0.019 0.015 -0.156

m.fgrade <- glmer(QCORRECT ~ QNUM + TREATMENT + final.grade.fix + Gender + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.fgrade)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT + final.grade.fix + Gender + (1 |   
## ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3104.2 3151.1 -1544.1 3088.2 2612   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.4428 -0.9306 0.4630 0.7092 2.5240   
##   
## Random effects:  
## Groups Name Variance Std.Dev.   
## ID (Intercept) 1.148e-14 1.072e-07  
## Number of obs: 2620, groups: ID, 658  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.812548 0.191717 -9.454 < 2e-16 \*\*\*  
## QNUM6 -0.896874 0.123500 -7.262 3.81e-13 \*\*\*  
## QNUM9 0.001575 0.127780 0.012 0.990   
## QNUM10 -0.552219 0.124039 -4.452 8.51e-06 \*\*\*  
## TREATMENT1 0.244337 0.087036 2.807 0.005 \*\*   
## final.grade.fix 3.729663 0.263766 14.140 < 2e-16 \*\*\*  
## GenderMale 0.505817 0.093251 5.424 5.82e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATM fnl.g.  
## QNUM6 -0.261   
## QNUM9 -0.335 0.517   
## QNUM10 -0.295 0.539 0.515   
## TREATMENT1 -0.245 -0.005 0.003 0.000   
## finl.grd.fx -0.840 -0.092 0.001 -0.053 0.024   
## GenderMale -0.044 -0.031 -0.002 -0.020 0.015 -0.126

m.fgrade.int <- glmer(QCORRECT ~ QNUM + TREATMENT\*Gender + final.grade.fix + (1|ID),   
 data = dat.trt,   
 family = binomial, control=glmerControl(optimizer="bobyqa"))  
summary(m.fgrade.int)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: QCORRECT ~ QNUM + TREATMENT \* Gender + final.grade.fix + (1 |   
## ID)  
## Data: dat.trt  
## Control: glmerControl(optimizer = "bobyqa")  
##   
## AIC BIC logLik deviance df.resid   
## 3105.9 3158.7 -1543.9 3087.9 2611   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.3826 -0.9246 0.4642 0.7075 2.5032   
##   
## Random effects:  
## Groups Name Variance Std.Dev.   
## ID (Intercept) 4.702e-14 2.169e-07  
## Number of obs: 2620, groups: ID, 658  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.82838 0.19383 -9.433 < 2e-16 \*\*\*  
## QNUM6 -0.89881 0.12357 -7.274 3.49e-13 \*\*\*  
## QNUM9 0.00138 0.12780 0.011 0.99138   
## QNUM10 -0.55398 0.12410 -4.464 8.04e-06 \*\*\*  
## TREATMENT1 0.27866 0.10650 2.617 0.00888 \*\*   
## GenderMale 0.55597 0.12951 4.293 1.76e-05 \*\*\*  
## final.grade.fix 3.72950 0.26371 14.142 < 2e-16 \*\*\*  
## TREATMENT1:GenderMale -0.10353 0.18490 -0.560 0.57552   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) QNUM6 QNUM9 QNUM10 TREATMENT1 GndrMl fnl.g.  
## QNUM6 -0.254   
## QNUM9 -0.331 0.517   
## QNUM10 -0.288 0.539 0.515   
## TREATMENT1 -0.283 -0.021 0.001 -0.015   
## GenderMale -0.133 -0.043 -0.003 -0.033 0.405   
## finl.grd.fx -0.831 -0.092 0.001 -0.053 0.021 -0.089   
## TREATMENT1: 0.148 0.030 0.003 0.026 -0.576 -0.694 -0.002

m<-m.fgrade.int  
se <- sqrt(diag(vcov(m)))  
tab <- cbind(Est = fixef(m), LL = fixef(m) - 1.96 \* se, UL = fixef(m) + 1.96 \*se)  
exp(tab)

## Est LL UL  
## (Intercept) 0.1606736 0.1098881 0.2349299  
## QNUM6 0.4070535 0.3194996 0.5186002  
## QNUM9 1.0013810 0.7794947 1.2864283  
## QNUM10 0.5746600 0.4505866 0.7328982  
## TREATMENT1 1.3213578 1.0724332 1.6280608  
## GenderMale 1.7436241 1.3527281 2.2474767  
## final.grade.fix 41.6582357 24.8442523 69.8515124  
## TREATMENT1:GenderMale 0.9016484 0.6275512 1.2954637