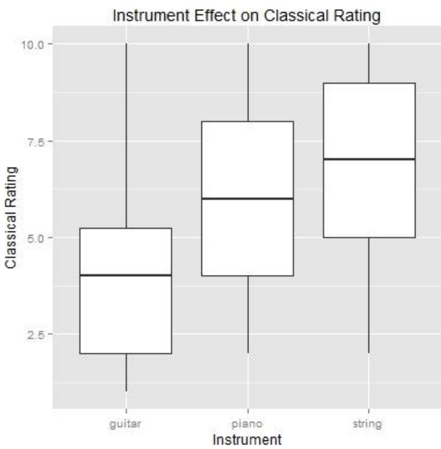
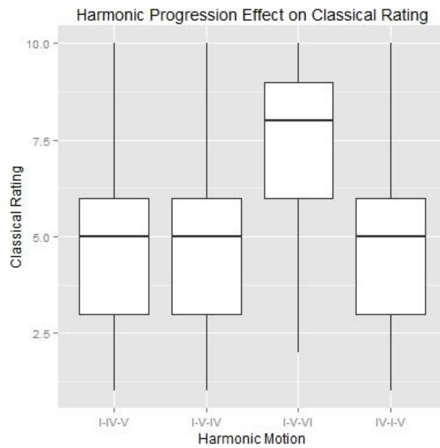
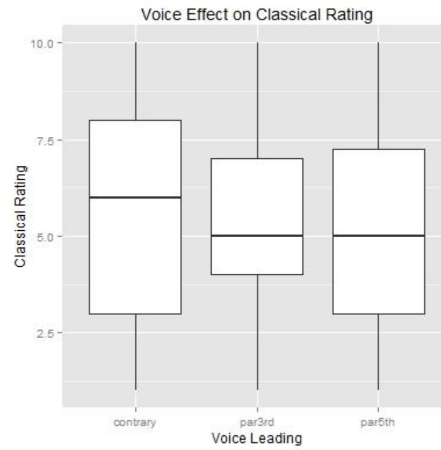
1. The Three Main Experimental Factors
    a. After omitting observations that contained NAs, the data was reduced from 2520 partial observations to 180 complete observations. An initial look at boxplots of the main experimental factors, Instrument, Harmony, and Voice, are displayed below.







The boxplots suggest that Instrument and Harmony probably have a significant effect on classical rating. Voice does not appear to have much of an effect. The boxplots also reaffirm the researcher's three main hypotheses: that instrument will have the largest influence on rating, that the harmonic progression I-V-VI will be rated as the most Classical, and contrary motion will be rated as the most Classical (slightly, as shown by the boxplot).

I fit four models using lm: a full model with all three experimental factors and three subsequent models each with only two experimental factors. Preforming anova on each one of the three reduced models compared to the full model shows us that Harmony and Instrument are significant in predicting Classical rating whereas Voice is not. The anova summaries are shown below:

```
Analysis of Variance Table
Model 1: Classical ~ Instrument + Harmony + Voice
Model 2: Classical ~ Instrument + Harmony
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    172 739.80
2    174 752.21 -2   -12.411 1.4428 0.2391

Analysis of Variance Table
Model 1: Classical ~ Instrument + Harmony + Voice
Model 2: Classical ~ Instrument + Voice
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1    172 739.80
2    175 953.56 -3   -213.76 16.566 1.68e-09 ***

Analysis of Variance Table
Model 1: Classical ~ Instrument + Harmony + Voice
Model 2: Classical ~ Harmony + Voice
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    172 739.80
2    174 948.48 -2   -208.68 24.258 5.242e-10 ***
```

Therefore, the lm model we choose is Classical ~ Harmony + Instrument. The estimated coefficients for each factor of this model are shown in the output below:

```
lm(formula = Classical ~ Instrument + Harmony)
                 coef.est coef.se
(Intercept)       3.64     0.38
Instrumentpiano   1.68     0.38
```

```
Instrumentstring  2.60      0.38
HarmonyI-V-IV    -0.13      0.44
HarmonyI-V-VI     2.36      0.44
HarmonyIV-I-V    -0.31      0.44
```

b. Repeated measures model.

i.

1) b) i)

$$y_i = \alpha_{0j[i]} + \alpha_1 X_i + \alpha_2 X_i + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$\alpha_{0j} = \beta_0 + \eta_j, \quad \eta_j \overset{iid}{\sim} N(0, \tau^2)$$

ii. I first fit the model with the random effect: lmer(Classical ~ Instrument + Harmony + (1 | Subject)) and then compared the AIC of this model to the AIC of the lm model without the random effect. Using the extractAIC(model) command in R, the AIC for the random effect model is 776.37 and the AIC for the previous model is 269.41. AIC alone suggests that the model without the random effect is the better model. Next I used a likelihood ratio test of the two models to test which is the better fit using the exactLRT(m=fullmodel, m0=reducedmodel) command in R. This test yielded a statistic of 7.77 and a p-value of 0.0022. This test suggests that we should reject the null hypothesis (the reduced model) in favor of the full model with the random effect component. To be safe, we will include the random effect in the model.

iii. Using an equivalent method to part (a), I tested each reduced main effect model with random effect against the full model with random effect. I again found that Voice is not needed in the model but Harmony and Instrument are both still significant.

```
Models:
fit5: Classical ~ Instrument + Harmony + (1 | Subject)
fit6: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit5  8 776.37 801.92 -380.19   760.37
fit6 10 777.12 809.05 -378.56   757.12 3.2568      2     0.1962

Models:
fit7: Classical ~ Instrument + Voice + (1 | Subject)
fit6: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit7  7 820.17 842.52 -403.08   806.17
fit6 10 777.12 809.05 -378.56   757.12 49.052      3  1.272e-10 ***

Models:
fit8: Classical ~ Harmony + Voice + (1 | Subject)
fit6: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit8  8 821.15 846.69 -402.57   805.15
fit6 10 777.12 809.05 -378.56   757.12 48.033      2  3.714e-11 ***
```

The final model I select from part (b) is Classical ~ Instrument + Harmony + (1|Subject). The coefficient estimates are displayed below along with a shortened version of the subject random effect estimates.

```
lmer(formula = Classical ~ Instrument + Harmony + (1 | Subject))
                 coef.est coef.se
(Intercept)       3.64     0.47
Instrumentpiano   1.68     0.36
Instrumentstring  2.60     0.36
HarmonyI-V-IV    -0.13     0.42
HarmonyI-V-VI     2.36     0.42
HarmonyIV-I-V    -0.31     0.42
```

```
$Subject
   (Intercept)
29   0.2310343
42   0.3865382
66   0.6086865
81  -0.8130630
94  -0.4131960
```

c. Interaction Random Effects

   i. Based on the significance of two of the experimental factors, Harmony and Instrument, I performed an anova test of the model with a the subject level random intercept, Classical ~ Instrument + Harmony + (1|Subject), against a model with three interaction random effect terms, Classical ~ Instrument + Harmony + (1|Subject:Instrument) + (1|Subject:Harmony) + (1|Subject:Voice). The output from this test is shown below:

```
Models:
fit5: Classical ~ Instrument + Harmony + (1 | Subject)
fit9: Classical ~ Instrument + Harmony + (1 | Subject:Instrument) +
fit9:     (1 | Subject:Harmony) + (1|Subject:Voice)
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit5  8 776.37 801.92 -380.19   760.37
fit9  9 761.76 790.50 -371.88   743.76 16.614      1  4.581e-05 ***
```

   So we conclude that the model with three random effects is better than the model with the single random effect (chisq=16.61, p=4.58e-05). Now we extract the log likelihood scores to compare the three-random-effect model with the simple lm model from part (a). Using the logLik(model) command, we find that the log likelihood of the model with random effects is -370.34 and the log likelihood of the lm model is -384.11. Because we want to maximize the log likelihood, we conclude that the model containing random effects is the superior model.

   ii. In order to reexamine the influence of the three main experimental factors on Classical ratings in a model with all three new random effect terms, we run anovas on each of the reduced models vs. the full model (the model containing all three experimental factors and all three random effects). We find, again, that Voice is not significant but Harmony and Instrument are both significant as evidenced in the output below.

```
Models:
fit9: Classical ~ Instrument + Harmony + (1 | Subject:Instrument) +
fit9:     (1 | Subject:Harmony) + (1|Subject:Voice)
fit10: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
fit10:     (1 | Subject:Harmony) + (1|Subject:Voice)
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit9   9 761.76 790.50 -371.88   743.76
fit10 11 761.55 796.67 -369.78   739.55 4.2084      2     0.1219

Models:
fit11: Classical ~ Harmony + Voice + (1 | Subject:Instrument) + (1 |
fit11:     Subject:Harmony) + (1 | Subject:Voice)
fit10: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
fit10:     (1 | Subject:Harmony) + (1|Subject:Voice)
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit11 10 773.50 805.43 -376.75   753.50
fit10 11 761.55 796.67 -369.78   739.55 13.948      1  0.0001879 ***

Models:
fit12: Classical ~ Harmony + Voice + (1 | Subject:Instrument) + (1 |
fit12:     Subject:Harmony) + (1 | Subject|Voice)
fit10: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
fit10:     (1 | Subject:Harmony) + (1|Subject:Voice)
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit12 10 773.50 805.43 -376.75   753.50
fit10 11 761.55 796.67 -369.78   739.55 13.948      1  0.0001879 ***
```

We also compare the estimated residual variance for each model. We find that all residual variances are about the same, but highest (2.99) for the model that we previously selected containing the main effects of Harmony and Instrument and the three random effects. This reaffirms our correct model choice.

iii.

c)iii) $y_i = \alpha_{0j[i]} + \alpha_1 x_i + \alpha_2 x_i + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

$\alpha_{0j} = \beta_0 + \beta_1 \mu_j + \beta_2 \mu_j + \beta_3 \mu_j + \eta_j, \quad \eta_j \overset{iid}{\sim} N(0, \tau^2)$

2. Individual Covariates

a. From here on out, we use the model with all three main experimental factors: Instrument, Harmony, and Voice, plus the three random effects that we explored in question #1 part (c). This model is significantly better than the model with the three main experimental factors and the single random effect, as evidenced by the following anova test:

```
Models:
fit6: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
fit10: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
fit10:     (1 | Subject:Harmony) + (1|Subject:Voice)
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit6  10 777.12 809.05 -378.56   757.12
fit10 11 761.55 796.67 -369.78   739.55 17.566      1  2.776e-05 ***
```

Using this as our base model, we update this model with each predictor and then anova test the model + predictor vs. the original model to see if that predictor is significant in explaining any additional variation in Classical rating. After fitting 19 such models (that can be seen in my attached R code), we find that KnowAxis is the only covariate that adds any significant predictive power to the original model with a Chi Square value of 5.34 and a corresponding p-value of 0.021, shown below.

```
Models:
fit10: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
fit10:     (1 | Subject:Harmony) + (1 | Subject:Voice)
fit22: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
fit22:     (1 | Subject:Harmony) + (1 | Subject:Voice) + KnowAxis
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit10 12 763.55 801.87 -369.78   739.55
fit22 13 760.22 801.72 -367.11   734.22 5.3352      1     0.0209 *
```

The updated model is now Classical ~ Instrument + Harmony + Voice + KnowAxis + (1|Subject:Instrument) + (1|Subject:Harmony) + (1|Subject:Voice). A summary of this model is given in the output below.

```
lmer(formula = Classical ~ Instrument + Harmony + Voice + (1 |
    Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
    KnowAxis)
                 coef.est coef.se
(Intercept)       3.13     0.73
Instrumentpiano   1.68     0.53
Instrumentstring  2.60     0.53
HarmonyI-V-IV    -0.13     0.69
HarmonyI-V-VI     2.36     0.69
HarmonyIV-I-V    -0.31     0.69
Voicepar3rd      -0.62     0.32
Voicepar5th      -0.15     0.32
KnowAxis          0.26     0.12
```

```
Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.93
 Subject:Voice      (Intercept) 0.14
 Subject:Instrument (Intercept) 0.68
 Residual                       1.72
```

b. We now check to see if there is any change in the random effects of the model. We anova-test the model with four main effects and four random effects (adding (1|Subject:KnowAxis)) against a model with the same four main effects and the single subject-level random effect. We find that the model with the four random effects is significantly better with a Chi Square value of 12.98 and a p-value of 0.0047. The output is shown below:

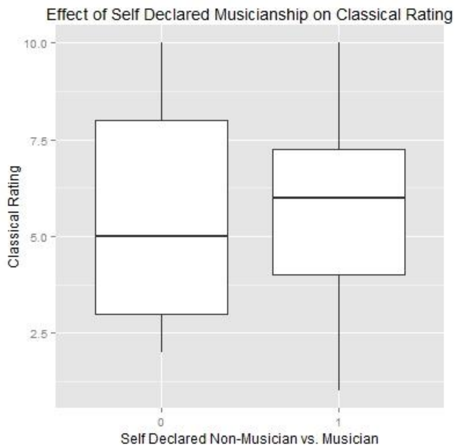
```
Models:
fit33: Classical ~ Harmony + Instrument + Voice + KnowAxis + (1 | Subject)
fit34: Classical ~ Harmony + Instrument + Voice + KnowAxis + (1 | Subject:Harmony)
+  (1 | Subject:Instrument) + (1 | Subject:Voice) + (1 | Subject:KnowAxis)
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit33 11 769.20 804.32 -373.60   747.20
fit34 14 762.22 806.92 -367.11   734.22 12.983      3   0.004674 **
```

Now we check to see if all four of the random effects are needed using the same iterative anova procedure. We find that the only significant random effects out of the four are (1|Subject:Instrument), (p=0.032) and (1|Subject:Harmony), (p=0.00094). Therefore, our new model fit is Classical ~ Harmony + Instrument + Voice + KnowAxis + (1|Subject:Instrument) + (1|Subject:Harmony). We test this fit against the full model with all four random effects to find a Chi Square value equal to approximately zero and a p-value equal to 1. This means that there is no difference between the two models so we are able to use the model with fewer terms without sacrificing any predictive power.

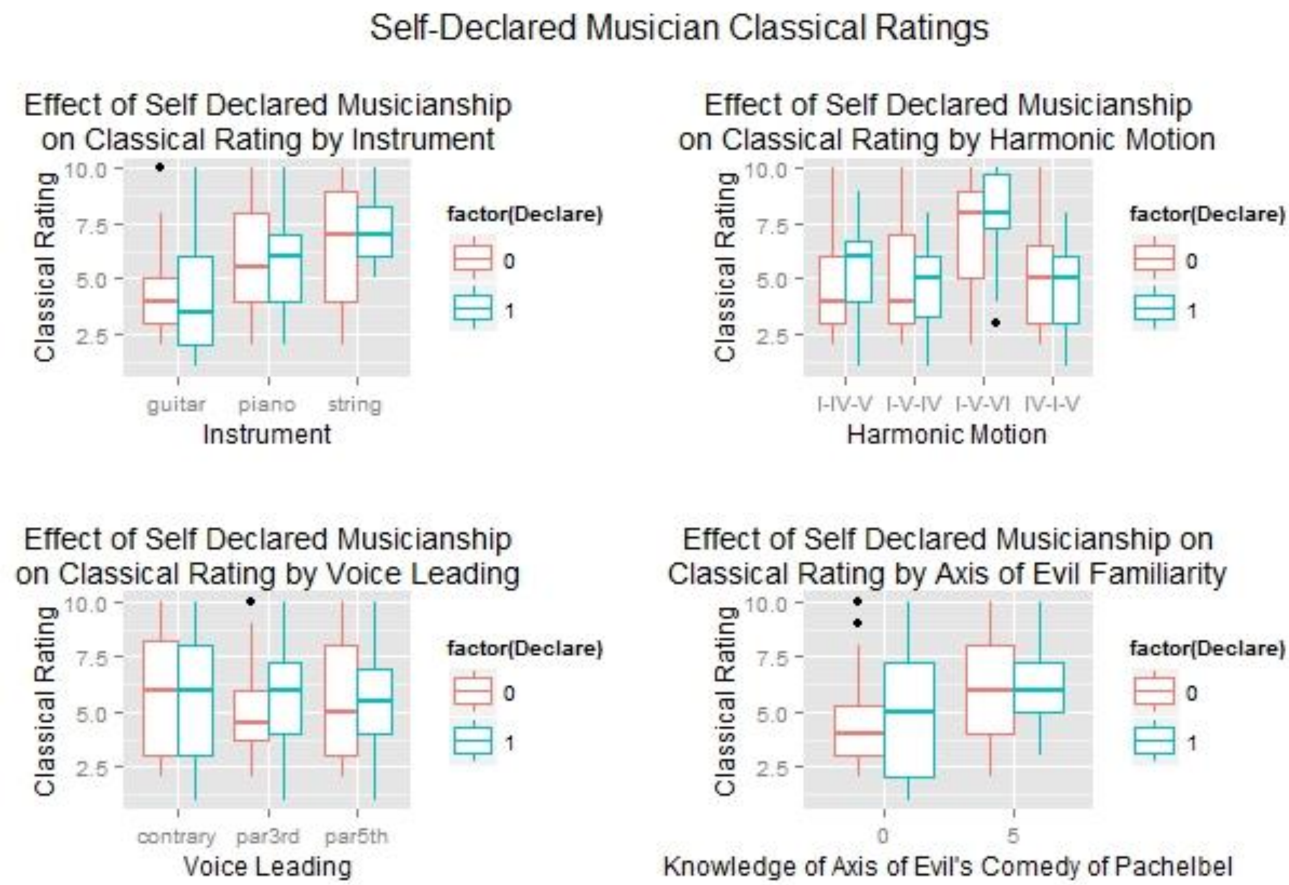
c. Interpretation of effects.
The overall intercept is 3.13. This means that the base Classical rating with a harmony pattern of I-VI-V, electric guitar as the main instrument, contrary motion as the voice leading and no knowledge of the Axis of Evil comedy is 3.13 out of 10. A harmony pattern of I-V-IV decreases Classical rating by an average of 0.13 points and IV-I-V by 0.31 points whereas I-V-VI increases Classical rating by an average of 2.36 points compared to I-VI-V. The use of piano increases Classical rating by an average of 1.68 points and string instruments by 2.60 points compared to electric guitar. Parallel thirds voice structure decreases Classical rating by an average of 0.62 points and parallel fifths by an average of 0.15 points compared to contrary motion. The intercepts vary based on individual subject-harmony and subject-instrument interactions.

3. Musicians vs. Non-musicians

From the boxplot below, we can see that overall, self-declared musicians rate music as more classical than non-declared musicians.



The four following plots show self-declared-musician Classical ratings vs. non-musician classical ratings for each main effect.



The first set of boxplots shows that self-declared musicians rate music about the same as non-musicians based on instrument. The second set of boxplots shows that self-declared musicians rate the I-IV-V progression as much more Classical than non-musicians do. The third set of boxplots shows consistent rankings between musicians and non-musicians, except self-declared musicians generally rate parallel thirds as more classical. The fourth set of boxplots shows consistent answers between musicians and non-musicians regarding classical music and their knowledge of the Axis of Evil's sketch about Pachelbel's chord progression.

We again use anova to test, one-by-one, for any significant interactions between "Self-declare" and any of the other four main effect predictors. Though the boxplots suggested some slight interactions, we do not find any statistically significant interactions between "self-declare" and any of the main effect predictors through anova testing.

4. Classical vs. Popular

a. We fit the same series of lm tests as we did in section (a) of question 1 using Popular as the response variable instead of Classical. Through anova testing, we find that Instrument is the only statistically significant predictor for Popular rating, as shown from the output below.

```
Analysis of Variance Table
Model 1: Popular ~ Harmony + Instrument + Voice
Model 2: Popular ~ Harmony + Voice
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    172 819.80
2    174 916.01 -2   -96.211 10.093 7.168e-05 ***
```

We choose the model Popular ~ Instrument. The coefficients of the model, displayed below, show that Popular rating is 6.05 when electric guitar is the main instrument. This rating decreases by 0.75 points when piano is the main instrument and by 1.78 points when a string quartet is the main instrument.

```
lm(formula = Popular ~ Instrument)
                 coef.est coef.se
(Intercept)       6.05     0.28
Instrumentpiano  -0.75     0.39
Instrumentstring -1.78     0.39
```

b. Our first step is to fit the same model that we decided was the best in 2(c), except using Popular rather than Classical as the predictor. This model is Popular ~ Harmony + Instrument + Voice + KnowAxis + (1|Subject:Instrument) + (1|Subject:Harmony). Looking at the summary (below), Harmony and Voice do not appear to be significant.

```
Fixed effects:
                   Estimate Std. Error t value
(Intercept)       4.867e+00  6.753e-01   7.207
HarmonyI-V-IV    -1.701e-14  6.249e-01   0.000
HarmonyI-V-VI    -2.222e-02  6.249e-01  -0.036
HarmonyIV-I-V    -2.222e-01  6.249e-01  -0.356
Instrumentpiano  -7.500e-01  5.120e-01  -1.465
Instrumentstring -1.783e+00  5.120e-01  -3.483
Voicepar3rd       3.000e-01  3.298e-01   0.910
Voicepar5th       3.167e-01  3.298e-01   0.960
KnowAxis          3.463e-01  1.113e-01   3.110
```

We then change the model to Popular ~Instrument + KnowAxis + (1|Subject:Instrument). All effects in this model appear to be significant for predicting popular rating.

```
Fixed effects:
                 Estimate Std. Error t value
(Intercept)       5.01111    0.47694  10.507
Instrumentpiano  -0.75000    0.55072  -1.362
Instrumentstring -1.78333    0.55072  -3.238
KnowAxis          0.34630    0.09179   3.773
```

c. Using the fit we decided on above, Popular ~ Instrument + KnowAxis + (1|Subject:Instrument), we update separately to include the interaction first between "self-declare" and instrument and second between "self-declare" and KnowAxis. We find that the "self-declare"-instrument interaction is significant here while the "self-declare"-KnowAxis interaction is not. The anova test for the significant interaction model vs. the model without this interaction is shown below.

need to include all experimental factors

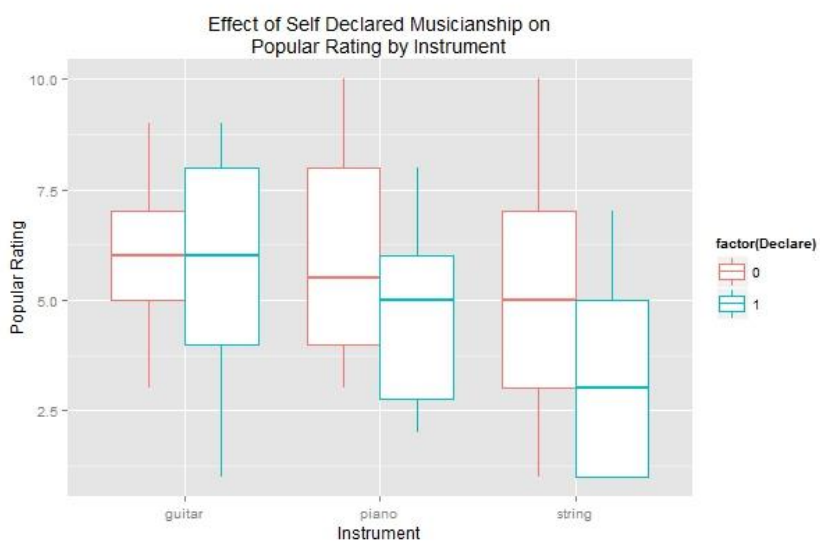
```
Models:
fit50: Popular ~ Instrument + KnowAxis + (1 | Subject:Instrument)
fit51: Popular ~ Instrument + KnowAxis + (1 | Subject:Instrument) +
fit51:     Instrument:factor(Declare)
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
fit50  6 763.08 782.24 -375.54   751.08
fit51  9 760.86 789.60 -371.43   742.86 8.2176      3    0.04172 *
```

The coefficients of the new model including the “self-declare”-instrument interaction are shown in the output below.

```
lmer(formula = Popular ~ Instrument + KnowAxis + (1 | Subject:Instrument) +
    Instrument:factor(Declare))
                                  coef.est coef.se
(Intercept)                        5.10     0.53
Instrumentpiano                   -0.33     0.63
Instrumentstring                  -1.25     0.63
KnowAxis                           0.32     0.08
Instrumentguitar:factor(Declare)1 -0.03     0.71
Instrumentpiano:factor(Declare)1  -1.07     0.71
Instrumentstring:factor(Declare)1 -1.36     0.71
```

why not interact musician with KnowAxis, etc?

We are also able to graphically display this interaction using boxplots, below. Using the boxplots and the output we can see that self-declared musicians rate music with string instruments much less "popular" than people who have not declared themselves to be musicians.



4: 16
5: 20 nicely organized

36

5. Write-up

Influence of Instrument, Harmony, and Voice

When testing the significance of the three main experimental factors on Classical rating, we found that only two of them – Instrument and Harmony – are significant. The coefficients resulting from the Classical ~ Instrument + Harmony model is displayed below. The absent factors: when Instrument is "electric guitar" and Harmony is "I-VI-V" are the baseline variables. This means that a song uses electric guitar as the main instrument and the harmonic motion is I-VI-V, the average Classical rating is the value of the intercept.

Testing similar models as above using Popular rating as the predictor instead of classical yielded only one significant predictor: Instrument. The baseline instrument is the same as above (electric guitar). The coefficients for the Popular ~ Instrument model are also displayed below.

Predictor Effects on Classical Rating

Variable	Coefficient Estimate
Intercept	3.64
Instrument: piano	1.68
Instrument: string	2.60
Harmony I-V-IV	-0.13
Harmony I-V-VI	2.36
Harmony IV-I-V	-0.31

Predictor Effects on Popular Rating

Variable	Coefficient Estimate
Intercept	6.05
Instrument: piano	-0.75
Instrument: string	-1.78

Variance Components

When considering random effects to improve the accuracy of the model, we considered two paths. The first path was adding a single subject-level random effect that would suggest a standard repeated measures model that accounts for "personal biases" in ratings. The second path was adding random effects that account for personal biases that vary with the type of instrument, harmony, and/or voice leading. Ultimately, we choose the random effects that account for personal bias interactions with the main effect variables because of its versatility and statistical significance over the single subject-level random effect.

Individual Covariates

The only covariate that appeared to add significant predictive power as a main effect to our proposed models is KnowAxis. With the new covariate and random effects in mind, and keeping all of the experimental effects in the model, the best model for Classical ratings is Classical ~ Instrument + Harmony + Voice + KnowAxis + (1|Subject:Instrument) + (1|Subject:Harmony) with the coefficients shown in a table below.

A similar analysis using Popular as the predictor would have yielded the same predictors as the Classical model. However, if we choose for a minute to use only the truly significant predictors while testing for but not necessarily including the experimental factors, we find a smaller, simpler model. This model only includes Instrument from the experimental factors but also includes the Instrument-SelfDeclared musician interaction. The model for Popular ratings is denoted by Popular ~ Instrument + KnowAxis + Instrument*SelfDeclare + (1|Subject:Instrument).

Popular Model | Classical Model

Variable	Coefficient Estimate	Variable	Coefficient Estimate
Intercept	5.10	Intercept	3.13
Instrument: piano	-0.33	Instrument: piano	1.68
Instrument: strings	-1.25	Instrument: string	2.60
KnowAxis	0.32	Harmony: I-V-IV	-0.13
Guitar*Musician	-0.03	Harmony: I-V-VI	2.36
Piano*Musician	-1.07	Harmony: IV-I-V	-0.31
String*Musician	-1.36	Voice: parallel 3rds	-0.62
		Voice: parallel 5ths	-0.15
		KnowAxis	0.26

Complete R Code

```
library(arm)
library(foreign)
library(lme4)
library(MASS)
library(ggplot2)
library(arm)
rating = read.csv("ratings.csv", header=T)
rating <- na.omit(rating)
attach(rating)
#omitted ALL observations with NAs, parsed it down from 2520 observations to 180 observations..
ggplot(rating, aes(x=Instrument, y=Classical)) + geom_boxplot() +
  labs(x="Instrument", y="Classical Rating", title="Instrument Effect on Classical Rating")
ggplot(rating, aes(x=Harmony, y=Classical)) + geom_boxplot() +
  labs(x="Harmonic Motion", y="Classical Rating", title="Harmonic Progression Effect on Classical Rating")
ggplot(rating, aes(x=Voice, y=Classical)) + geom_boxplot() +
  labs(x="Voice Leading", y="Classical Rating", title="Voice Effect on Classical Rating")
#from boxplot eda alone we can see that all of the researcher's hypotheses are true

#1

#a
fit1 = lm(Classical ~ Instrument + Harmony)
fit2 = lm(Classical ~ Instrument + Voice)
fit3 = lm(Classical ~ Harmony + Voice)
fit4 = lm(Classical ~ Instrument + Harmony + Voice)
anova(fit4, fit1) #testing for significance of Voice
#Voice does not appear to be significant (F=1.44, p=0.24)
anova(fit4, fit2) #testing for significance of Harmony
#Harmony is significant (F=16.57, p=1.68e-09)
anova(fit4, fit3) #testing for significance of Instrument
#Instrument is significant (F=24.26, F=5.24e-10)
#model: Classical ~ Harmony + Instrument
summary(fit1)
display(fit1)

#b
fit5 = lmer(Classical ~ Instrument + Harmony + (1 | Subject) )
summary(fit5)
ranef(fit5)
fixef(fit5)
#which is better - with or without random effect?
extractAIC(fit1)
extractAIC(fit5)
#the AIC for fit1 is much smaller so the random effect is not needed (for the truncated data set)
library(RLRsim)
exactLRT(m=fit5, m0=fit1)
#a likelihood ratio test tells us that the random effect IS significant. to be safe, we will include it

#b (iii)
fit6 = lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject))
```

```
fit7 = lmer(Classical ~ Instrument + Voice + (1 | Subject))
fit8 = lmer(Classical ~ Harmony + Voice + (1 | Subject))
anova(fit6, fit5) #testing effect of Voice
#Voice is not needed (chisq = 3.26, p=0.20)
anova(fit6, fit7) #testing effect of Harmony
#Harmony is needed (chisq=49.05, p=1.27e-10)
anova(fit6, fit8) #testing effect of Instrument
#Instrument is needed (chisq=48.03, p=3.71e-11)
#The same 2 main experimental factors are significant with or without the random effect
#keep model fit5
display(fit5)
ranef(fit5)
#c
#i model with random effect terms vs with and without random intercept
#fit9=model with 3 random effect terms vs. best model from b
fit9 = lmer(Classical ~ Instrument + Harmony +
        (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1:Subject:Voice))
display(fit9)
anova(fit9, fit1)
anova(fit9, fit5)
#the model with 3 random effects is better than the model with the single random effect (chisq=16.61, p=4.58e-05)
#now, fit9 model vs. best model from a (fit1)
logLik(fit9)
logLik(fit1)
#fit9 - random effects model is better than fit1

#ii
fit10 = lmer(Classical ~ Instrument + Harmony + Voice +
        (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1|Subject:Voice)) # full model
fit11 = lmer(Classical ~ Instrument + Voice +
        (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1|Subject:Voice))
fit12 = fit11 = lmer(Classical ~ Harmony + Voice +
          (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1|Subject:Voice))
anova(fit10,fit9) #is Voice significant?
#no (chisq=4.21, p=0.12)
anova(fit10,fit11) #is Harmony significant?
# yes (chisq=13.95, p=0.00094)
anova(fit10,fit12) #is Instrument significant?
# yes (chisq=13.95, p=0.00094)

#comparing residual variance
attr(VarCorr(fit9),"sc")^2
attr(VarCorr(fit10),"sc")^2
attr(VarCorr(fit11),"sc")^2
attr(VarCorr(fit12),"sc")^2
#all residual variances are about the same, but lowest for fit10 - the model with the most predictors.
#this is because, regardless of whether or not the terms are significant, adding more terms will always
#explain more variation in the response and increase the predictive power of the model


#2
#best model from probelm 1, but including all 3 experimental factors of Voice, Harmony, and Instrument
```

```
#is the best model the one with the single random effect or 3 random effects?
anova(fit10, fit6)
#fit10 (with the 3 random effects) is the better fitting model (chisq=17.56, 2.78e-05)
fit13 = update(fit10, . ~ . + Selfdeclare)
anova(fit10, fit13)
#selfdeclare not significant
fit14 = update(fit10, . ~ . + OMSI)
anova(fit10, fit14)
#omsi not significant
fit15 = update(fit10, . ~ . + X16.minus.17)
anova(fit10, fit15)
#x16.minus.17 not significant
fit16 = update(fit10, . ~ . + ConsInstr)
anova(fit10, fit16)
#ConsInstr not significant
fit17 = update(fit10, . ~ . + ConsNotes)
anova(fit10, fit17)
#ConsNotes not significant
fit18 = update(fit10, . ~ . + Instr.minus.Notes)
anova(fit10, fit18)
#Instr.minus.Notes not significant
fit19 = update(fit10, . ~ . + PachListen)
anova(fit10, fit19)
#?PachListen not significant
fit20 = update(fit10, . ~ . + ClsListen)
anova(fit10, fit20)
#ClsListen not significant
fit21 = update(fit10, . ~ . + KnowRob)
anova(fit10, fit21)
#KnowRob not significant
fit22 = update(fit10, . ~ . + KnowAxis)
anova(fit10, fit22)
#KnowAxis is significant? chisq=5.34, p=0.021
fit23 = update(fit10, . ~ . + X1990s2000s)
anova(fit10, fit23)
#X1990s2000s not significant
fit24 = update(fit10, . ~ . + X1990s2000s.minus.1960s1970s)
anova(fit10, fit24)
#X1990s2000s not significant
fit25 = update(fit10, . ~ . + factor(CollegeMusic))
anova(fit10, fit25)
#CollegeMusic not significant
fit26 = update(fit10, . ~ . + NoClass)
anova(fit10, fit26)
#NoClass not significant
fit27 = update(fit10, . ~ . + factor(APTheory))
anova(fit10, fit27)
#APTheory not significant
fit28 = update(fit10, . ~ . + Composing)
anova(fit10, fit28)
#Composing not significant
fit29 = update(fit10, . ~ . + PianoPlay)
```

```
anova(fit10, fit29)
#PianoPlay not significant
fit30 = update(fit10, . ~ . + GuitarPlay)
anova(fit10, fit30)
#GuitarPlay not significant
fit31 = update(fit10, . ~ . + X1stInstr)
anova(fit10, fit31)
#X1stInstr not significant
fit32 = update(fit10, . ~ . + X2ndInstr)
anova(fit10, fit32)
#X2ndInstr not significant

#The only additional covariate that appears to be significant is KnowAxis, fit22
summary(fit22)
display(fit22)

#b
#fixed effects to include: Harmony, Instrument, Voice, KnowAxis
#random effects to include:
fit33 = lmer(Classical ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject))
fit34 = lmer(Classical ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Subject:Voice) + (1|Subject:KnowAxis))
#1 random effect vs 4-random effects
anova(fit33, fit34)
#4 random effects significantly better than one raneff (chisq 12.98, p=0.0047) but do we need all of them?
fit35 = lmer(Classical ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Subject:Voice))
anova(fit34, fit35)
fit36 = lmer(Classical ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Subject:KnowAxis))
anova(fit34, fit36)
fit37 = lmer(Classical ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject:Harmony)  + (1|Subject:Voice) + (1|Subject:KnowAxis))
anova(fit34, fit37)
# ^ significant, full model explains significantly more than model without (1|Subject:Instrument)
fit38 = lmer(Classical ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject:Instrument) + (1|Subject:Voice) + (1|Subject:KnowAxis))
anova(fit34, fit38)
# ^ significant, full model explains significantly more than model without (1|Subject:Harmony)
fit39 = lmer(Classical ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject:Instrument) + (1|Subject:Harmony))
anova(fit34, fit39)
#^ fit39 is better in terms of AIC, BIC, etc. The random effects Subject:KnowAxis and Subject:Voice
#do not significantly explain any additional variation in the model
#model: fit39: Classical ~ Harmony + Instrument + Voice + KnowAxis + (1|Subject:Instrument) + (1|Subject:Harmony)
display(fit39)
ranef(fit39)

#3
Declare = ifelse(Selfdeclare>3, 1, 0)
ggplot(rating, aes(x=factor(Declare), y=Classical)) + geom_boxplot() +
```

```
  labs(x="Self Declared Non-Musician vs. Musician", y="Classical Rating",
      title="Effect of Self Declared\nMusicianship on Classical Rating")
#overall, self-declared musicians rank music as more Classical
a = ggplot(rating, aes(x=factor(Instrument), y=Classical, colour=factor(Declare))) + geom_boxplot() +
  labs(x="Instrument", y="Classical Rating",
      title="Effect of Self Declared Musicianship\non Classical Rating by Instrument")
b = ggplot(rating, aes(x=factor(Harmony), y=Classical, colour=factor(Declare))) + geom_boxplot() +
  labs(x="Harmonic Motion", y="Classical Rating",
      title="Effect of Self Declared Musicianship\non Classical Rating by Harmonic Motion")
c = ggplot(rating, aes(x=factor(Voice), y=Classical, colour=factor(Declare))) + geom_boxplot() +
  labs(x="Voice Leading", y="Classical Rating",
      title="Effect of Self Declared Musicianship\non Classical Rating by Voice Leading")
d = ggplot(rating, aes(x=factor(KnowAxis), y=Classical, colour=factor(Declare))) + geom_boxplot() +
  labs(x="Knowledge of Axis of Evil's Comedy of Pachelbel", y="Classical Rating",
      title="Effect of Self Declared Musicianship on\nClassical Rating by Axis of Evil Familiarity")
library(gridExtra)
grid.arrange(a,b,c,d, ncol=2, nrow=2, main="Self-Declared Musician Classical Ratings")

display(fit39)
fit40 = update(fit39, . ~ . + (factor(Declare)):Instrument)
anova(fit40, fit39)
#selfdeclare - instrument interaction is not significant
fit41 = update(fit39, . ~ . + (factor(Declare)):Harmony)
anova(fit41, fit39)
#selfdeclare - harmony interaction is not significant
fit42 = update(fit39, . ~ . + (factor(Declare)):Voice)
anova(fit42, fit39)
#selfdeclare-voice interaction is not significant
fit43 = update(fit39, . ~ . + (factor(Declare)):KnowAxis)
anova(fit43, fit39)
#selfdeclare-knowaxis interaction is not significant
#SelfDeclare does not appear to have any significant interactions with the other predictors in the model

#4
#a
fit44 = lm(Popular ~ Harmony + Instrument + Voice)
summary(fit44)
fit45= lm(Popular ~ Harmony + Instrument)
fit46 = lm(Popular ~ Harmony + Voice)
fit47 = lm(Popular ~ Instrument + Voice)
anova(fit44, fit45)
anova(fit44, fit46)
#instrument is imporant factor in model (f=10.09, p=7.17e-5)
anova(fit44, fit47)
fit48 = lm(Popular ~ Instrument)
anova(fit48, fit44)
#Instrument is the only significant predictor of Popular Rating
ggplot(rating, aes(x=Instrument, y=Popular)) + geom_boxplot() +
  labs(x="Instrument", y="Popular Rating", title="Effect of Instrument on Popular Rating")
display(fit48)

#b
```

```
#using the final model we decided upon for Classical in 2c...
fit49 = lmer(Popular ~ Harmony + Instrument + Voice + KnowAxis +
        (1|Subject:Instrument) + (1|Subject:Harmony))
summary(fit49)
# looking at summary, Harmony and Voice do not appear to be significant, changing model to
fit50 = lmer(Popular ~ Instrument + KnowAxis + (1|Subject:Instrument))
summary(fit50)

#c
display(fit50)
fit51 = update(fit50, . ~ . + (factor(Declare)):Instrument)
anova(fit50, fit51)
#selfdeclare - instrument interaction IS significant
fit52 = update(fit50, . ~ . + (factor(Declare)):KnowAxis)
anova(fit50, fit52)
#selfdeclare - knowaxis interaction is not significant
#we choose model fit51: Popular ~ Instrument + KnowAxis + SelfDeclare*Instument + (1|Subject:Instrument)
display(fit51)
ggplot(rating, aes(x=factor(Instrument), y=Popular, colour=factor(Declare))) + geom_boxplot() +
  labs(x="Instrument", y="Popular Rating",
    title="Effect of Self Declared Musicianship on\nPopular Rating by Instrument")
#self declared musicians rate music with string instruments much LESS "popular" than ppl that haven't
#declared themselves musicians

finalclassical = lmer(Classical ~ Instrument + Harmony + Voice + KnowAxis +
    (1|Subject:Instrument) + (1|Subject:Harmony))
display(finalclassical)
```