1 a 9/9
b 9/9
c 9/9

2 a 9/9
b 9/9
c 9/9

3 9/9

4 a 9/9
b 9/9
c 9/9

5 10/10

Total 100/100

36-763 Homework #5

Abbie Legge

December 18, 2015

Problem 1

(a) First we are interested in the influence of the three main experimental factors (Instrument, Harmony, & Voice) on Classicial ratings. We can use an ANOVA model to examine the influence of the three factors.

9

```
> summary(aov.1a)
                    Df Sum Sq Mean Sq F value   Pr(>F)
factor(Harmony)      3    278    92.8  17.771 2.06e-11 ***
factor(Voice)        2     78    39.2   7.511  0.00056 ***
factor(Instrument)   2   4088  2043.8 391.266  < 2e-16 ***
Residuals         2484  12975     5.2
---
```

The ANOVA analysis indicates that all three factors: Harmony, Voice, and Instrument have variation that is significantly greater than could be explained by chance alone. In other words, the variation in classical ratings across different levels of Harmony, Voice, and Instrument is more significant than could be explained by chance alone. Additionally, partial F tests we used to compare the full model to the reduced model without each of the three experimental factors. The results indicated that the full model is a better fit than any of the reduced models (see appendix). We will fit the full linear model to see how each factor influcences classical ratings.

```
lm(formula = Classical ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument))
                            coef.est coef.se
(Intercept)                  4.34     0.13
as.factor(Harmony)I-V-IV    -0.03     0.13
as.factor(Harmony)I-V-VI     0.77     0.13
as.factor(Harmony)IV-I-V     0.03     0.13
as.factor(Voice)par3rd      -0.40     0.11
as.factor(Voice)par5th      -0.36     0.11
as.factor(Instrument)piano   1.37     0.11
as.factor(Instrument)string  3.12     0.11
---
n = 2492, k = 8
residual sd = 2.29, R-Squared = 0.26
```

The intercept of the full model represents the average rating for a selection with Harmony I-V-vi, contrary motion, and played on an electric guitar. The coefficients represent the change in the overall mean rating for different Harmonies, Voices, and Instruments. Of the harmonies used in the experiement, Harmony I-V-VI had a higher average rating as classical than the other three harmonies. Of the voices, Contrary motion had the highest average classical rating. The largest variation was in the instrument ratings. Piano had a significantly higher average rating that electric guitar and string quartet had a higher average rating than both piano and electric quartet.

(b) Next we are interested in examining the effect of fitting a random intercept for each participant. This model is called a "repeated measures" model.

(i) We can write the model in mathematical terms as a hierarchical linear model.

3

$$
\begin{aligned}
y_i &= \alpha_{0j[i]} + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2) \\
\alpha_{0j} &= \beta_0 + \eta_j, \quad \eta_j \sim N(0, \tau^2)
\end{aligned}
$$

Where x_1 represents Harmony, x_2 represents Voice, and x_3 represents Instrument. β_0 represents the overall average Classical rating.

(ii) One way we can test for the significance of the random effect is by comparing model fit measurements across the different models.

Table 1: Model Comparison - Random Intercept

3

	AIC	BIC
No Random Intercept (part a)	11201.70	11254.09
Random Intercept (part b)	10434.30	10492.49

Including the random intercept reduces both AIC and BIC, so there is evidence that the model including the random effect is a better fit.

We can also use the RLRsim package to do a simulation based test for the random intercept.

```
> model.1.b = lmer(Classical ~ as.factor(Harmony) + as.factor(Voice) +
as.factor(Instrument) + (1|Subject), data=rate)
> exactRLRT(model.1.b)

simulated finite sample distribution of RLRT.

(p-value based on 10000 simulated values)

data:
RLRT = 769.19, p-value < 2.2e-16
```

Based on the results of the test, since the p-value is very small, including the random intercept improves the fit of the model. There is sufficient evidence of a personal bias in music ratings.

(iii) Next we want to reexamine the fixed effects after including the random intercepts.

```
lmer(formula = Classical ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument) + (1 | Subject), data = rate, REML = FALSE)
                            coef.est coef.se
(Intercept)                  4.34     0.19
as.factor(Harmony)I-V-IV    -0.03     0.11
as.factor(Harmony)I-V-VI     0.77     0.11
as.factor(Harmony)IV-I-V     0.03     0.11
as.factor(Voice)par3rd      -0.40     0.09
as.factor(Voice)par5th      -0.36     0.09
```

```
as.factor(Instrument)piano   1.37     0.09
as.factor(Instrument)string  3.12     0.09
```

3

```
Error terms:
 Groups   Name        Std.Dev.
 Subject  (Intercept) 1.29
 Residual             1.88
---
number of obs: 2492, groups: Subject, 70
AIC = 10434.3, DIC = 10414.3
deviance = 10414.3
```

The estimates for the fixed effects have not changed from the model in part (a). Harmony I-V-VI still has the highest classical rating among the harmonies, contrary motion is still the highest classically rated voice and string quartets are still most frequently rated as classical. However, the variance in the estimates of the fixed effects is now smaller and the residual variance is smaller. Accounting for the subjects' personal biases lowered the variance attributed to other elements of the model. The estimated variance across subjects is reasonably high, which suggests that personal biases have a large effect on classical ratings.

(c) Next we are interested in examining the effect of fitting random interactions between the subjects and the experimental factors. We are considering whether personal biases affect one's perception of Insruments, Harmonies, and Voices.

(i) To determine whether the model with random interaction terms is better than the previous models we can compare the AIC and BIC.

Table 2: Model Comparison - Random Interactions

3

	AIC	BIC
No Random Intercept (part a)	11201.70	11254.09
Random Intercept (part b)	10434.30	10492.49
Random Interactions (part c)	10015.50	10085.37

The model with all three random interaction terms has a lower AIC and BIC than both the model from part a and the model from part b, suggesting that it is better than the previous two models.

(ii)

```
lmer(formula = Classical ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
    (1 | Subject:Instrument), data = rate, REML = FALSE)
                            coef.est coef.se
(Intercept)                  4.34     0.21
as.factor(Harmony)I-V-IV    -0.03     0.14
as.factor(Harmony)I-V-VI     0.77     0.14
as.factor(Harmony)IV-I-V     0.04     0.14
as.factor(Voice)par3rd      -0.39     0.08
as.factor(Voice)par5th      -0.36     0.08
as.factor(Instrument)piano   1.36     0.26
as.factor(Instrument)string  3.12     0.26

Error terms:
 Groups             Name        Std.Dev.
```

```
 Subject:Harmony    (Intercept) 0.67
 Subject:Voice      (Intercept) 0.14
 Subject:Instrument (Intercept) 1.47
 Residual                       1.55
---
number of obs: 2492,
groups: Subject:Harmony, 280; Subject:Voice, 210; Subject:Instrument, 210
AIC = 10015.5, DIC = 9991.5, deviance = 9991.5
```

3

The fixed effects represent the overall means and have not changed from the previous models. Harmony I-V-VI is still more likely to be rated as classical, Contrary motion is the most frequent voice to be rated as classical, and string quartet is the most common instrument to be rated as classical. The estimates for the three estimated variance components differ greatly. The largest is the estimate for the student:instrument variance, suggesting that the effect of different instuments on different subjects varies a lot. The effect of different harmonies on different subjects is smaller than the effect of different instruments and the effect of different voices is smaller than both of the other two. All three estimates are smaller than the estimated residual variance, but the subject:instrument estimate is fairly close in size.

(iii) We can write the model in mathematical terms as a hierarchical linear model.

3

$$\begin{aligned}
y_i &= \alpha_0 + \alpha_{Hjk[i]} + \alpha_{Vjl[i]} + \alpha_{Ijm[i]} + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2) \\
\alpha_{Hjk[i]} &= \beta_{H[k]} + \eta_{jH}, \quad \eta_{jH} \sim N(0, \tau_1^2) \text{ for k = 1, 2, 3, 4 (Harmony levels)} \\
\alpha_{Vjl[i]} &= \beta_{V[l]} + \eta_{jV}, \quad \eta_{jV} \sim N(0, \tau_2^2) \text{ for l = 1, 2, 3 (Voice levels)} \\
\alpha_{Ijm[i]} &= \beta_{I[m]} + \eta_{jI}, \quad \eta_{jI} \sim N(0, \tau_3^2) \text{ for m = 1, 2, 3 (Instrument levels)}
\end{aligned}$$

Problem 2

(a) To determine which individual covariates should be included in the model, I used the automated variable selection from the LMERConvenienceFunctions library. I started with a large fixed effects model including many of the optional variables. I chose not to include the variables with large quantities of missing values in order to keep the most I created two big models, one that treated the fixed effects as factors and another that left many of them as integers since there is an ordered structure to many of the (0-5) responses. I ran the variable selection using AIC and BIC on both of the models and compared the selections that it made. Both AIC and BIC selected the same model and the same variables (continuous vs. factors). Mostly for convenience, leaving certain items as continuous makes the model more interpretable so I elected to do that.

9

The code I used to run the selection is in the appendix. The final model selected was:

```
lmer(formula = Classical ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument) + PachListen + as.factor(CollegeMusic) +
    as.factor(APTheory) + Composing + X1990s2000s + (1 | Subject:Harmony) +
    (1 | Subject:Voice) + (1 | Subject:Instrument), data = rate,
    REML = TRUE)
                              coef.est coef.se
(Intercept)                    3.80     0.78
as.factor(Harmony)I-V-IV      -0.03     0.16
as.factor(Harmony)I-V-VI       0.80     0.16
as.factor(Harmony)IV-I-V       0.04     0.16
```

```
as.factor(Voice)par3rd      -0.40     0.09
as.factor(Voice)par5th      -0.36     0.09
as.factor(Instrument)piano   1.43     0.28
as.factor(Instrument)string  3.25     0.28
PachListen                   0.09     0.12
as.factor(CollegeMusic)1    -0.28     0.33
as.factor(APTheory)1         0.23     0.30
Composing                    0.09     0.10
X1990s2000s                  0.04     0.09

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.71
 Subject:Voice      (Intercept) 0.18
 Subject:Instrument (Intercept) 1.45
 Residual                       1.55
---
number of obs: 2108,
groups: Subject:Harmony, 236; Subject:Voice, 177; Subject:Instrument, 177
AIC = 8519.5, DIC = 8433.4
deviance = 8459.4
```

The variables that were selected are: PachListen, College Music, APTheory, Composing, and x1990s2000s.

(b) After selecting the fixed effects, I used the RLRSim package to test for each of the three random effects and the random intercept.

In order, test for (1|subject:instrument), (1|subject:harmony), (1|subject:voice), (1|subject)

```
> exactRLRT(monly2bI, mfull2bI, mnull2bI)
simulated finite sample distribution of RLRT.
(p-value based on 10000 simulated values)
data:  RLRT = 497.82, p-value < 2.2e-16

> exactRLRT(monly2bH, mfull2bH, mnull2bH)
simulated finite sample distribution of RLRT.
(p-value based on 10000 simulated values)
data:  RLRT = 107.78, p-value < 2.2e-16

> exactRLRT(monly2bV, mfull2bV, mnull2bV)
simulated finite sample distribution of RLRT.
(p-value based on 10000 simulated values)
data:  RLRT = 1.2059, p-value = 0.1311 #not significant

> exactRLRT(monly2bInt, mfull2bInt, mnull2bInt)
```

9

```
simulated finite sample distribution of RLRT.
(p-value based on 10000 simulated values)
data:  RLRT = 23.719, p-value < 2.2e-16
```

The tests suggest that the random effect (1:subject:voice) is not significant and can be removed from the model.

(c) Our model so far is:

```
lmer(formula = Classical ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument) + PachListen + as.factor(CollegeMusic) +
    as.factor(APTheory) + Composing + X1990s2000s + (1 | Subject:Harmony) +
    (1 | Subject:Instrument) + (1 | Subject), data = rate)
                              coef.est coef.se
(Intercept)                    3.80     1.05
as.factor(Harmony)I-V-IV      -0.03     0.16
as.factor(Harmony)I-V-VI       0.80     0.16
as.factor(Harmony)IV-I-V       0.04     0.16
as.factor(Voice)par3rd        -0.40     0.08
as.factor(Voice)par5th        -0.36     0.08
as.factor(Instrument)piano     1.43     0.23
as.factor(Instrument)string    3.25     0.23
PachListen                     0.09     0.16
as.factor(CollegeMusic)1      -0.28     0.45
as.factor(APTheory)1           0.23     0.42
Composing                      0.09     0.13
X1990s2000s                    0.04     0.12

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.68
 Subject:Instrument (Intercept) 1.14
 Subject            (Intercept) 1.05
 Residual                       1.56
---
number of obs: 2108, groups: Subject:Harmony, 236; Subject:Instrument, 177; Subject, 59
AIC = 8497, DIC = 8415.6
deviance = 8439.3
```

debatable whether to keep these in the model, but OK.

9

The final model contains the design variables and they have the same effects that they have had all along. The new variables are PachListen, CollegeMusic, APTheory, Composing, and x1990s2000s. The implied associations are that subjects who are more familiar with Pachabel's Canon, who took AP Music Theory in high school, who have done composing, and who listen to pop and rock from the 90s and 00s are more likely to rate the selections as classical, while subjects who took college music classes are less likely to rate the selections as classical. None of the coefficients is statistically significant on its own but might be helping to explain some of the variation in the ratings regardless.

Problem 3

The table below summarizes the responses to "Are You a Musician"

Table 3: Responses to Self-Declare

Self-Declare as Musician	1	2	3	4	5	6
Count of Participants	16	26	12	13	2	1

Based on the results of the table, achieving a perfectly even split will not happen, but grouping participants who answered 1 or 2 as "Non-Musicians" and those who answered 3-6 as "Musicians" seems reasonable.

I focused my analysis on the interactions between the musician indicator and the design variables, instrument, harmony, and voice. The only interaction that reduced the overall model AIC was the interaction between musicians and harmonies (AIC: 8489.7), but I also included the interaction between musicians and instruments.

```
lmer(formula = Classical ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument) + PachListen + as.factor(CollegeMusic) +
    as.factor(APTheory) + Composing + X1990s2000s + as.factor(Harmony):Musician +
    as.factor(Instrument):Musician + Musician + (1 | Subject:Harmony) +
    (1 | Subject:Instrument) + (1 | Subject), data = rate)
                                     coef.est coef.se
(Intercept)                           3.90     1.09
as.factor(Harmony)I-V-IV             -0.07     0.20
as.factor(Harmony)I-V-VI              0.32     0.20
as.factor(Harmony)IV-I-V             -0.01     0.20
as.factor(Voice)par3rd               -0.40     0.08
as.factor(Voice)par5th               -0.36     0.08
as.factor(Instrument)piano            1.64     0.30
as.factor(Instrument)string           3.51     0.30
PachListen                            0.09     0.16
as.factor(CollegeMusic)1             -0.29     0.46
as.factor(APTheory)1                  0.25     0.42
Composing                             0.12     0.16
X1990s2000s                           0.02     0.13
Musician                             -0.16     0.54
as.factor(Harmony)I-V-IV:Musician     0.08     0.31
as.factor(Harmony)I-V-VI:Musician     1.12     0.31
as.factor(Harmony)IV-I-V:Musician     0.12     0.31
as.factor(Instrument)piano:Musician  -0.49     0.46
as.factor(Instrument)string:Musician -0.62     0.46

Error terms:
 Groups             Name        Std.Dev.
9 Subject:Harmony    (Intercept) 0.64
 Subject:Instrument (Intercept) 1.14
 Subject            (Intercept) 1.07
 Residual                       1.56
---
number of obs: 2108, groups: Subject:Harmony, 236; Subject:Instrument, 177; Subject, 59
AIC = 8491.4, DIC = 8393.5
deviance = 8419.4
```

Including the interactions for harmony and instrument with musician shows a few interesting patterns. The interaction with music brings down the coefficients on piano and string, suggesting that musicians are not as influenced by the type of instrument as non-musicians. The coefficients on the harmony interactions increase the likelihood that certain harmonies will be rated as classical, especially harmony I-V-VI.

Problem 4

Now we reconsider the data in terms of the popular ratings.

(a) First we are interested in the influence of the three main experimental factors (Instrument, Harmony, & Voice) on Popular ratings. We can use an ANOVA model to examine the influence of the three factors.

```
> summary(aov.4a)
                      Df Sum Sq Mean Sq F value Pr(>F)
factor(Harmony)        3     29     9.7   1.931  0.122
factor(Voice)          2     14     6.9   1.370  0.254
factor(Instrument)     2   2918  1458.8 290.333 <2e-16 ***
Residuals           2484  12481     5.0
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Unlike the ANOVA analysis on classical ratings, this ANOVA analysis indicates that only the Instrument factor has variation that is significantly greater than what could be explained by chance alone. We will still fit the full model to consider the effect of each level of the factors.

```
Call:
lm(formula = Popular ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument), data = ratep)

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                  6.59327    0.12676  52.012   <2e-16 ***
as.factor(Harmony)I-V-IV    -0.04538    0.12701  -0.357   0.7209
as.factor(Harmony)I-V-VI    -0.26837    0.12695  -2.114   0.0346 *
as.factor(Harmony)IV-I-V    -0.18570    0.12685  -1.464   0.1434
9 as.factor(Voice)par3rd     0.15269    0.11003   1.388   0.1654
as.factor(Voice)par5th       0.16329    0.10993   1.485   0.1376
as.factor(Instrument)piano  -0.96819    0.11030  -8.778   <2e-16 ***
as.factor(Instrument)string -2.61171    0.10961 -23.828   <2e-16 ***
---
Residual standard error: 2.242 on 2484 degrees of freedom
Multiple R-squared:  0.1917,Adjusted R-squared:  0.1894
F-statistic: 84.17 on 7 and 2484 DF,  p-value: < 2.2e-16
```

From the output above, the Harmony I-IV-V was most frequently rated as popular and the Harmony I-V-VI had the lowest popular ratings. Both parallel voice structures were more frequently rated as popular than contrary motion. Fewer of the differents are statistically significant that in the classical ratings, especially those for voice leading.

(b) Again I used the automated variable selection from the LMERConvenienceFunctions package to select my covariates and the final random effects (from the same four that we considered in problem 1). Unlike for classical ratings, the variable selection eliminated both the harmony and voice variables from the model and only kept Instrument. The other predictors that were selected were the same as the predictors in for Classical: PachListen, College Music, APTheory, Composing, and x1990s2000s.

```
lmer(formula = Popular ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument) + PachListen + as.factor(CollegeMusic) +
    as.factor(APTheory) + Composing + X1990s2000s + (1 | Subject:Harmony) +
```

```
    (1 | Subject:Instrument) + (1 | Subject), data = ratep)
                            coef.est coef.se
(Intercept)                  7.06     1.00
as.factor(Harmony)I-V-IV    -0.04     0.15
as.factor(Harmony)I-V-VI    -0.31     0.15
as.factor(Harmony)IV-I-V    -0.22     0.15
as.factor(Voice)par3rd       0.20     0.08
as.factor(Voice)par5th       0.21     0.08
as.factor(Instrument)piano  -1.02     0.21
as.factor(Instrument)string -2.78     0.21
PachListen                  -0.14     0.16
as.factor(CollegeMusic)1     0.03     0.43
as.factor(APTheory)1         0.14     0.40
Composing                    0.15     0.13
X1990s2000s                  0.02     0.12

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.64
 Subject:Instrument (Intercept) 1.05
 Subject            (Intercept) 1.01
 Residual                       1.59
---
number of obs: 2108, groups: Subject:Harmony, 236; Subject:Instrument, 177; Subject, 59
AIC = 8532.9, DIC = 8449.6
deviance = 8474.2
```

9

The final model contains the design variables and they have the same effects that they had before the other variables were introduced. The new variables are PachListen, CollegeMusic, APTheory, Composing, and x1990s2000s. The implied associations are that subjects who are more familiar with Pachabel's Canon are less likely to rate the selections as popular, but subjects who took AP Music Theory in high school, who have done composing, and who listen to pop and rock from the 90s and 00s are more likely to rate the selections as popular. Its interesting to note that the factors increase both classical and popular ratings. Perhaps exposure to music on some level increases the ratings on both scales. None of the coefficients is statistically significant on its own but might be helping to explain some of the variation in the ratings regardless.

(c) As in problem 3, I only consider the interactions between musician and the experimental design variables. When interaction effects were considered, none of them were found to significantly improve the fit of the model. The model with both the interactions for harmony and instrument is presented below.

```
lmer(formula = Popular ~ as.factor(Harmony) + as.factor(Voice) +
    as.factor(Instrument) + PachListen + as.factor(CollegeMusic) +
    as.factor(APTheory) + Composing + X1990s2000s + as.factor(Harmony):Musician +
    as.factor(Instrument):Musician + Musician + (1 | Subject:Harmony) +
    (1 | Subject:Instrument) + (1 | Subject), data = ratep)
                                     coef.est coef.se
(Intercept)                           7.20     1.02
as.factor(Harmony)I-V-IV             -0.12     0.20
as.factor(Harmony)I-V-VI             -0.05     0.20
```

```
as.factor(Harmony)IV-I-V              -0.24     0.20
as.factor(Voice)par3rd                 0.20     0.08
as.factor(Voice)par5th                 0.21     0.08
as.factor(Instrument)piano            -1.21     0.27
as.factor(Instrument)string           -3.10     0.27
PachListen                            -0.14     0.16
as.factor(CollegeMusic)1               0.03     0.43
as.factor(APTheory)1                   0.14     0.40
Composing                              0.15     0.14
X1990s2000s                            0.02     0.12
Musician                              -0.33     0.36
as.factor(Harmony)I-V-IV:Musician      0.20     0.30
9 as.factor(Harmony)I-V-VI:Musician   -0.61     0.30
as.factor(Harmony)IV-I-V:Musician      0.07     0.30
as.factor(Instrument)piano:Musician    0.43     0.41
as.factor(Instrument)string:Musician   0.76     0.41

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.62
 Subject:Instrument (Intercept) 1.04
 Subject            (Intercept) 1.02
 Residual                       1.59
---
number of obs: 2108,
groups: Subject:Harmony, 236; Subject:Instrument, 177; Subject, 59
AIC = 8536, DIC = 8434.2
deviance = 8462.1
```

Similarly to classical ratings, the effect of instruments is smaller for the musicians and the effect of the harmony is larger. Musicians in particular are less likely to rate the harmonic motion I-V-VI as popular than non-musicans.

Problem 5

Introduction

Reasearchers from the University of Pittsburgh are interested in the effect of harmonic motion, instrument, and voice leading on listeners' identification of music as "classical" or "popular." 70 subjects were recruited and given 36 stimuli to rate on a scale from 1 to 10 for how classical the music sounds and how popular the music sounds. Information about the subjects was also collected including their musical background, the types of music they listen to frequently, and whether or not they self-declare as a musician.

Results

Of the three design variables in the experiment, instrument was found to be the most predictive of classical and popular ratings. The string quartet stimuli received the highest classical ratings and the lowest popular ratings. Alternatively, the electric guitar stimuli received the highest popular ratings and the lowest classical ratings. For both classical and popular ratings, the variation across the instruments was significantly greater than what would be expected by chance alone. Harmonic motion and voice leading were not as strong of predictors as instrument, but both were significant in determining ratings. The harmonic motion I-V-VI, *Pachelbel's Canon*, had significantly higher classical ratings and significantly lower popular ratings, but the differences among the other three harmonies were not significant. For voice leading, contrary motion had the highest classical ratings, followed by parallel 3rds and finally followed by parallel 5ths.

The inclusion of variance components improved the overall fit of the model. Treating the model as a repeated measures model with subject level grouping helped to account for personal differences in classical and popular ratings. Additionally, individuals varied in how they were influenced by the different instrumental and harmonic factors. For some individuals a change in the instrument led to large changes in the ratings and for others the ratings only differed slightly. For the variance components that were included in both models, the variation in the effects of different instruments was the largest. The variation in personal biases was slightly smaller, and the variation in the effects of different harmonies was the smallest. The variation in the effect of different voices was not significant.

Automated variable selection methods were used to select additional variables to include in a predictive model. Although none of the predictors was significant given the other variables in the model, including indicators for whether a subject took AP Music Theory in high school, took music classes in college, and variables categorizing how familiar the subject is with Pachelbel's Canon, how much he or she listens to pop and rock from the 90s and 00s, and how much musical compostion he or she has done improves the overall fit and predictive ability of the model. Familiarity with Pachelbel's Canon was associated with higher classical ratings and lower popular ratings and taking music classes in college was association with higher popular ratings and lower classical ratings. The other three variables, whether a subject took music theory, how much musical composition he or she has done, and how much 90's and 00's music he or she listens to were associated with higher ratings of both classical and popular. Perhaps increased exposure to music in different forms has an influence on both types of ratings.

The impact of self-declaring as a musician was also considered in order to see if harmony, voice, and instrument have different influences on musicians than on non-musicians. Overall, the influence of different instruments was smaller for musicians than non musicians and the influence of different harmonies was larger for musicians than non-musicians. Self-declared musicians are less likely to determine ratings based only on instrument. Instead the harmonic motion plays a more significant role in influencing their ratings. In particular, musicians gave significantly higher classical ratings and significantly lower popular ratings to the harmonic motion I-V-VI than non-musicians.

Appendix: Additional R Code and Results

(1a) Partial F tests for each experimental factor in the linear model.

```
> anova(lm.1a.noharmony, full.lm1a)
Analysis of Variance Table

Model 1: Classical ~ as.factor(Voice) + as.factor(Instrument)
Model 2: Classical ~ as.factor(Harmony) + as.factor(Voice) + as.factor(Instrument)
  Res.Df   RSS Df Sum of Sq      F   Pr(>F)
1   2487 13252
2   2484 12975  3    277.16 17.687 2.33e-11 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
> anova(lm.1a.novoice, full.lm1a)
Analysis of Variance Table

Model 1: Classical ~ as.factor(Harmony) + as.factor(Instrument)
Model 2: Classical ~ as.factor(Harmony) + as.factor(Voice) + as.factor(Instrument)
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2486 13055
2   2484 12975  2    79.688 7.6278 0.0004982 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
> anova(lm.1a.noinstrument, full.lm1a)
Analysis of Variance Table

Model 1: Classical ~ as.factor(Harmony) + as.factor(Instrument)
Model 2: Classical ~ as.factor(Harmony) + as.factor(Voice) + as.factor(Instrument)
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2486 13055
2   2484 12975  2    79.688 7.6278 0.0004982 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

(2a) Variable selection for fixed effects in the the classical model.

```
lmer.full3 = lmer(Classical ~ as.factor(Harmony) + as.factor(Voice) +
as.factor(Instrument) +  OMSI + X16.minus.17 + Instr.minus.Notes +
PachListen + ClsListen + as.factor(CollegeMusic) + as.factor(APTheory) +
Composing + PianoPlay + GuitarPlay + X1990s2000s +
(1|Subject:Harmony) + (1|Subject:Voice) + (1|Subject:Instrument), data=rate)

aic.best3 = fitLMER.fnc(lmer.full3, method ="AIC")

bic.best3 = fitLMER.fnc(lmer.full3, method ="BIC")
```

(4b)

```
lmer.full4b = lmer(Popular ~ as.factor(Harmony) + as.factor(Voice) +
as.factor(Instrument) +  OMSI + X16.minus.17 + Instr.minus.Notes +
PachListen + ClsListen + as.factor(CollegeMusic) + as.factor(APTheory) +
Composing + PianoPlay + GuitarPlay + X1990s2000s + (1|Subject), data=ratep)
```

```
aic.best4b = fitLMER.fnc(lmer.full4b, ran.effects = c("(1|Subject:Harmony)",
"(1|Subject:Voice)", "(1|Subject:Instrument)"), method ="AIC")

> display(aic.best4b)
lmer(formula = Popular ~ as.factor(Instrument) + PachListen +
    as.factor(CollegeMusic) + as.factor(APTheory) + Composing +
    X1990s2000s + (1 | Subject) + (1 | Subject:Harmony) + (1 |
    Subject:Instrument), data = ratep, REML = TRUE)
                            coef.est coef.se
(Intercept)                  7.06     1.00
as.factor(Instrument)piano  -1.02     0.21
as.factor(Instrument)string -2.78     0.21
PachListen                  -0.14     0.16
as.factor(CollegeMusic)1     0.03     0.43
as.factor(APTheory)1         0.14     0.40
Composing                    0.15     0.13
X1990s2000s                  0.02     0.12

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.64
 Subject:Instrument (Intercept) 1.05
 Subject            (Intercept) 1.01
 Residual                       1.59
---
number of obs: 2108,
groups: Subject:Harmony, 236; Subject:Instrument, 177; Subject, 59
AIC = 8523.7, DIC = 8476.3
deviance = 8488.0
```