# The Influence of Instrument, Harmony, and Voice Leading on Listeners' Distinction between Popular and Classical Music

Chu Pan

*chup@andrew.cmu.edu*

*01 December 2019*

**Abstract**

This paper examines the effects of instrument, harmonic motion and voice leading on listeners' distinction between popular and classical music. The experimental factors are analyzed by the use of multi-level models in conjunction with exploratory data analysis and ANOVA. The results show that (i) instrument is more influential than harmonic motion and voice leading for listeners' identification of musical genre regardless of the extent of listeners' musical training; (ii) among the levels of harmonic motion, I-V-vi have the strongest association with classical ratings; the phenomenon is subject to listeners' previous exposure to the Pachelbel rants/comedy bits; (iii) among the levels of voice leading, contrary motion has the strongest association with classical ratings; (iv) musicians and non-musicians responded differently to chord progressions involving I-V-vi; (v) classical and popular ratings are driven by different factors. This paper also addresses on issues such as missing data and personal biases in the experiment.

## 1 Introduction

This paper explores the influence of instrument, harmonic motion, and voice leading on listeners' distinction between popular and classical music. Interest in this question is motivated by the blurring boundaries between the two music genres. Early in the modern era, the distinction between the two music genres were partly socially based. Classical music, performed by professionals to passive audience from elite classes, was understood to be universal and eternal. Popular music, often created in a spontaneous nature and consumed by mass audience, was taken to be regional and ephemeral (Hamm 1995). Nowadays, not only does classical music become accessible to the general public, but popular music are deeply influenced by the classical repertoire. Pachelbel's *Canon in D*—one of the most famous Baroque pieces—has been sampled by over 30 mainstream songs on Spotify. Its beginning harmonic progression, I-V-vi, also appears repeatedly in popular music of the past 20 years. As classical and popular music become intertwined with each other, researchers are interested in learning which factors of music composition are crucial to listeners' identification of musical genre.

This paper takes the data from Ivan Jimenez and student Vincent Rossi and analyzes the influence of three main designed factors: instrument, harmonic motion and voice leading. In particular, this paper will address the following questions:

- What experimental factor, or combination of factors, has the strongest influence on classical ratings?
- Are there differences in the way that musicians and non-musicians identify classical music?
- Are there differences in the factors that drive classical vs popular ratings?

## 2 Method

### 2.1 Experimental design

In 2012, Ivan Jimenez (a composer and musicologist visiting the University of Pittsburgh) and student Vincent Rossi collected the data in a designed experiment intended to measure the influence of instrument, harmonic motion, and voice leading on listeners' identification of music as "classical" or "popular". They presented 36 musical stimuli to 70 listeners, recruited from the population of undergraduates at the University of Pittsburgh, and asked the listeners to rate the music on two different scales:

- How classical does the music sound (1 to 10, 1=not at all, 10=very classical sounding);
- How popular does the music sound (1 to 10, 1=not at all, 10=very popular sounding).

Listeners were told that a piece could be rated as both classical and popular, neither classical nor popular, or mostly classical and not popular (or vice versa), so that the scales should have functioned more or less independently. The 36 stimuli were chosen by completely crossing these factors:

- Instrument:String Quartet, Piano, Electric Guitar
- Harmonic Motion:I-V-vi, I-VI-V, I-V-IV, IV-I-V
- Voice Leading:Contrary Motion, Parallel 3rds, Parallel 5ths

A brief description of all variables in the data set is given in Table 1 on page 3.

### 2.2 Preprocessing

#### 2.2.1 Missing Data

As shown in the data quality diagnosis (Table 2 on page 4), 15 of the 26 variables in this data set contain missing values. Particularly, 87% values in `X2ndInstr` and 60% values in `X1stInstr` are missing. Removing all entries with missing values is tantamount to throwing out at least 87% real data for the study, which is not an ideal condition for modeling. Therefore, it is not appropriate to simply ignore all missing values in this case. According to definitions of the variables, both `X1stInstr` and `X2ndInstr` measure participants' proficiency at musical instruments. Similar information is covered by other variables in the data set. `Selfdeclare` (self identification as musician), `OMSI` (score on a test of musical knowledge), `PianoPlay` could all be used to measure participants' musical training. Since `X1stInstr` and `X2ndInstr` contain lots of missing values and there are other variables that provide similar—if not better—measurements of musical knowledge, I decided to remove these two variables from the data set.

`ConsNotes` also contains 14% missing values. I used MICE (Multivariate Imputation by Chained Equations) to impute the missing values. This imputation method yields the best approximation of the observed data (Figure 2) while preserving uncertainty of the true values. Similar approach was applied to impute `KnowAxis` and `NoClass`.

All the other variables have less than 10% missing values. The trivial proportion would not have a large impact on our estimates. These observations were therefore removed from the data set.

Table 1: Variable definitions for music rating data set

| Number | Variable | Description |
|---|---|---|
| 1 | Subject | Unique subject ID |
| 2 | Harmony | Harmonic Motion (4 levels) |
| 3 | Instrument | Instrument (3 levels) |
| 4 | Voice | Voice Leading (3 levels) |
| 5 | Selfdeclare | Are you a musician? (1-6, 1=not at all) |
| 6 | OMSI | Score on a test of musical knowledge |
| 7 | X16.minus.17 | Auxiliary measure of listener's ability to distinguish classical vs popular music |
| 8 | ConsInstr | How much did you concentrate on the instrument while listening? (0-5, 0=not at all) |
| 9 | ConsNotes | How much did you concentrate on the notes while listening? (0-5, 0=not at all) |
| 10 | Instr.minus.Notes | Difference between prev. two variables |
| 11 | PachListen | How familiar are you with Pachelbel's Canon in D? (0-5, 0=not at all) |
| 12 | ClsListen | How much do you listen to classical music? (0-5, 0=not at all) |
| 13 | KnowRob | Have you heard Rob Paravonian's Pachelbel Rant? (0-5, 0=not at all) |
| 14 | KnowAxis | Have you heard Axis of Evil's Comedy bit on the 4 Pachelbel chords in popular music? (0-5, 0=not at all) |
| 15 | X1990s2000s | How much do you listen to pop and rock from the 90's and 2000's? (0-5, 0=not at all) |
| 16 | X1990s2000s.minus.1960s1970s | Difference between prev variable and a similar variable referring to 60'sand 70's pop and rock |
| 17 | CollegeMusic | Have you taken music classes in college? (0=no, 1=yes) |
| 18 | NoClass | How many music classes have you taken? |
| 19 | APTheory | Did you take AP Music Theory class in High School? (0=no, 1=yes) |
| 20 | Composing | Have you done any music composing? (0-5,0=not at all) |
| 21 | PianoPlay | Do you play piano? (0-5, 0=not at all) |
| 22 | GuitarPlay | Do you play guitar? (0-5, 0=not at all) |
| 23 | X1stInstr | How proficient are you at your first musical instrument? (0-5, 0=not at all) |
| 24 | X2ndInstr | Same, for second musical instrument |
| 25 | Classical | How classical does the stimulus sound? |
| 26 | Popular | How popular does the stimulus sound? |

Table 2: Data Quality Diagnosis

| variables | types | missing_count | missing_percent | unique_count | unique_rate |
|---|---|---|---|---|---|
| Subject | character | 0 | 0.00 | 70 | 0.03 |
| Harmony | character | 0 | 0.00 | 4 | 0.00 |
| Instrument | character | 0 | 0.00 | 3 | 0.00 |
| Voice | character | 0 | 0.00 | 3 | 0.00 |
| Selfdeclare | numeric | 0 | 0.00 | 6 | 0.00 |
| OMSI | numeric | 0 | 0.00 | 60 | 0.02 |
| X16.minus.17 | numeric | 0 | 0.00 | 13 | 0.01 |
| ConsInstr | numeric | 0 | 0.00 | 14 | 0.01 |
| ConsNotes | numeric | 360 | 14.29 | 6 | 0.00 |
| Instr.minus.Notes | numeric | 0 | 0.00 | 20 | 0.01 |
| PachListen | numeric | 72 | 2.86 | 7 | 0.00 |
| ClsListen | numeric | 36 | 1.43 | 6 | 0.00 |
| KnowRob | numeric | 180 | 7.14 | 4 | 0.00 |
| KnowAxis | numeric | 288 | 11.43 | 4 | 0.00 |
| X1990s2000s | numeric | 144 | 5.71 | 6 | 0.00 |
| X1990s2000s.minus.1960s1970s | numeric | 180 | 7.14 | 10 | 0.00 |
| CollegeMusic | numeric | 108 | 4.29 | 3 | 0.00 |
| NoClass | numeric | 288 | 11.43 | 7 | 0.00 |
| APTheory | numeric | 216 | 8.57 | 3 | 0.00 |
| Composing | numeric | 72 | 2.86 | 7 | 0.00 |
| PianoPlay | numeric | 0 | 0.00 | 5 | 0.00 |
| GuitarPlay | numeric | 0 | 0.00 | 5 | 0.00 |
| X1stInstr | numeric | 1512 | 60.00 | 6 | 0.00 |
| X2ndInstr | numeric | 2196 | 87.14 | 6 | 0.00 |
| Classical | numeric | 27 | 1.07 | 17 | 0.01 |
| Popular | numeric | 27 | 1.07 | 17 | 0.01 |

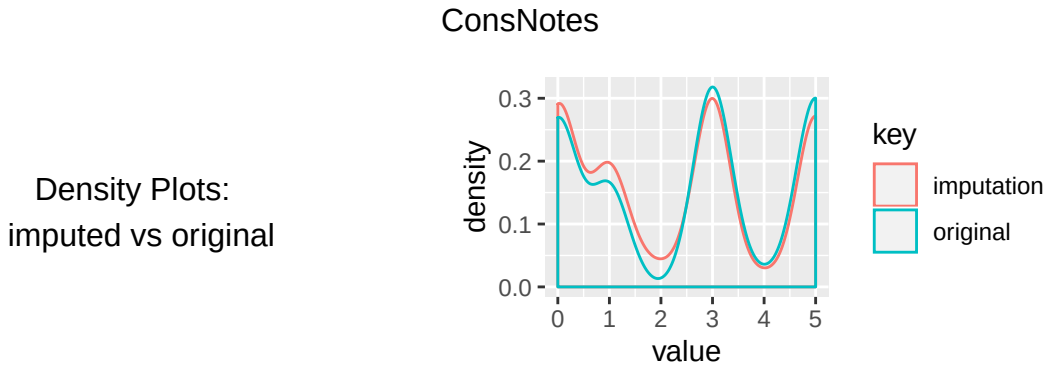



Figure 1: Density plots: imputed vs observed data

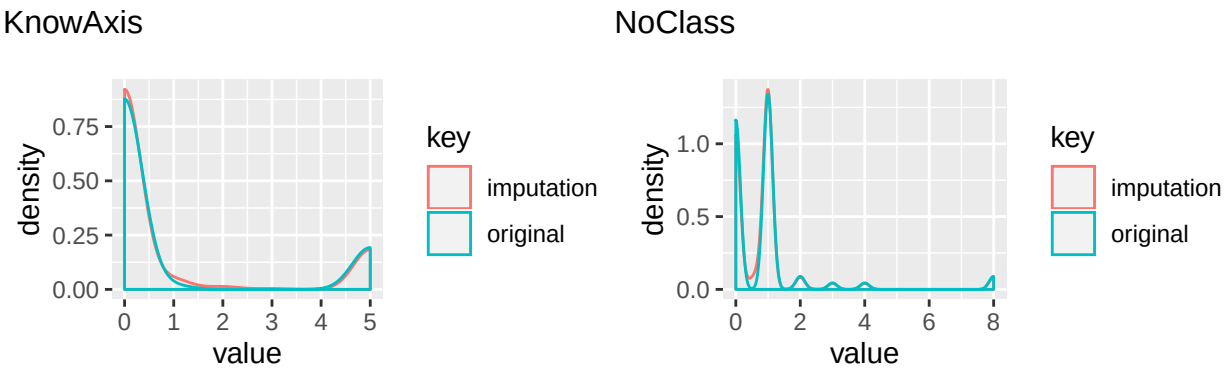



Figure 2: Density plots: imputed vs observed data

#### 2.2.2 Unusual Data Values

`Popular` and `Classical` are defined as 1-10 scaled ratings, yet both variables contain non-integer entries. Also, `Popular` has a minimum of 0 and a maxinum of 19, which look like data entry errors. The "19" is probably meant to be "10" since 9 locates next to 0 on keyboard. I corrected the mistyped entries by replacing all "19"s with "10". "0" appeared frequently in both `Popular` and `Classical`. It is unclear what type of mistake it may be by context. These zero entries were removed from the data set.

### 2.3 Statistical methods

The general plan of procedure in this study is to establish by regression analysis the set of determinants of music genre. Firstly, pairwise correlations between variables were examined using corrplot. The comparison helps to identify unusual correlations that hint at potential sampling problems in the experiment. Details will be discussed in the result section.

Secondly, ANOVAs, linear regressions and multi-level models were used to determine the importance of relationships and to derive regression coefficients. Personal biases in ratings were taken into account by the inclusion of random effects in the model. Natural log transformations were applied to compensate for skewing in the data.

This study used `fitLMER` in `LMERConvenienceFunctions` package to select models that yielded rating estimates with the minimum error. The analysis relied on R language and environment for statistical computing (R Core Team 2017). Concerns of music and meaning of the variables also played a major role in the model selection process.

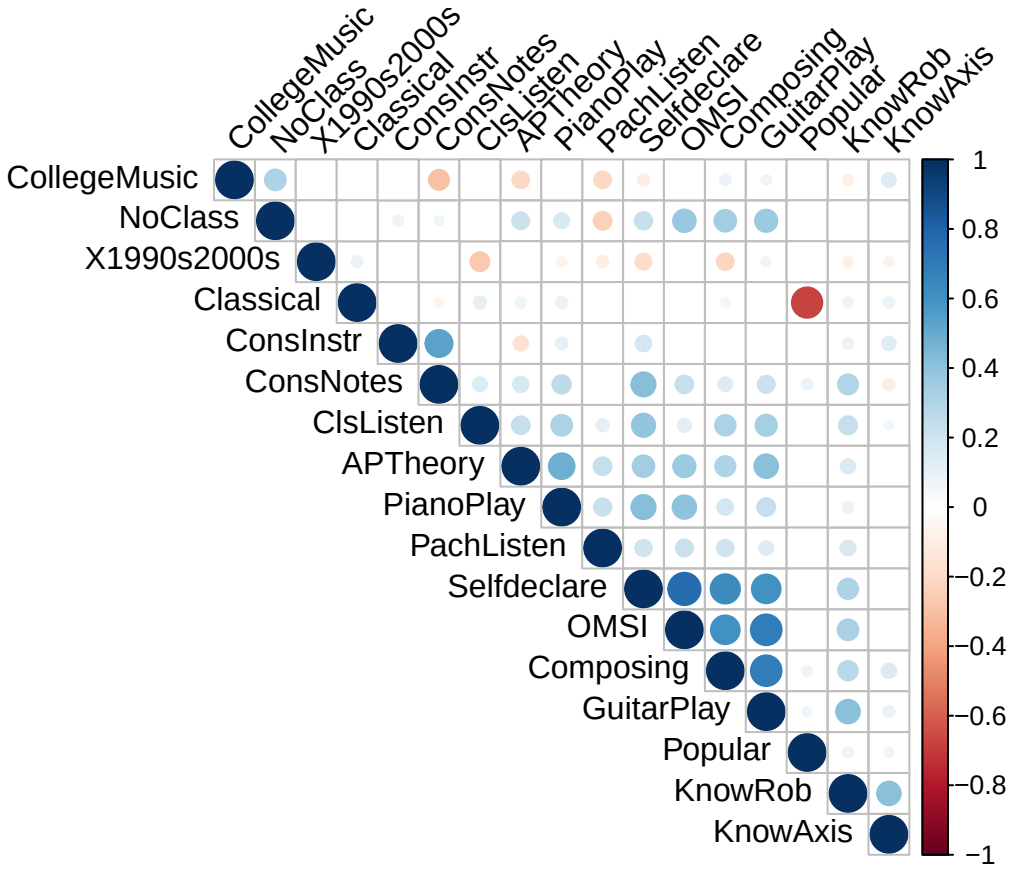



Figure 3: Pairwise relationship between variables

## 3 Result

### 3.1 Pairwise correlations & Transformations

Pairwise relationship between variables are given in the Figure 3. Most correlations align with our expectation. Those who self-identified as musicians have stronger background in instrumental performance, composing, music theory, and they tend to focus on notes while listening.

A surprising finding is that `CollegeMusic` is negatively associated with `ConsNotes`, which means those who have taken music courses in college paid less attention to notes while listening to the stimuli. This variable is also negatively related to `Selfdeclare`,`APTheory` and `PachListen`. The finding is counter-intuitive because we usually expect students to advance on the field after taking courses in college. But the participants in this study exhibited an opposite pattern: those who have taken music courses seem to have less musical training compared to other participants, as they did not take AP music theory and were not familiar with the well-known piece *Canon in D*. This unexpected finding is probably associated with self-selection bias introduced by convenient sampling in this experiment. If the participants were either students taking intro-level music courses or professionals in the UPitt orchestra, then it would make total sense that the amateurs did not know much about music *compared to professionals*.

This study checked skewness of data using QQ-plots. The distribution of `OMSI` is skewed on both sides (Figure 4). A log transformation was applied to fix the skewness. The log transformation also enables interpretation of `OMSI` effects in percentage change.

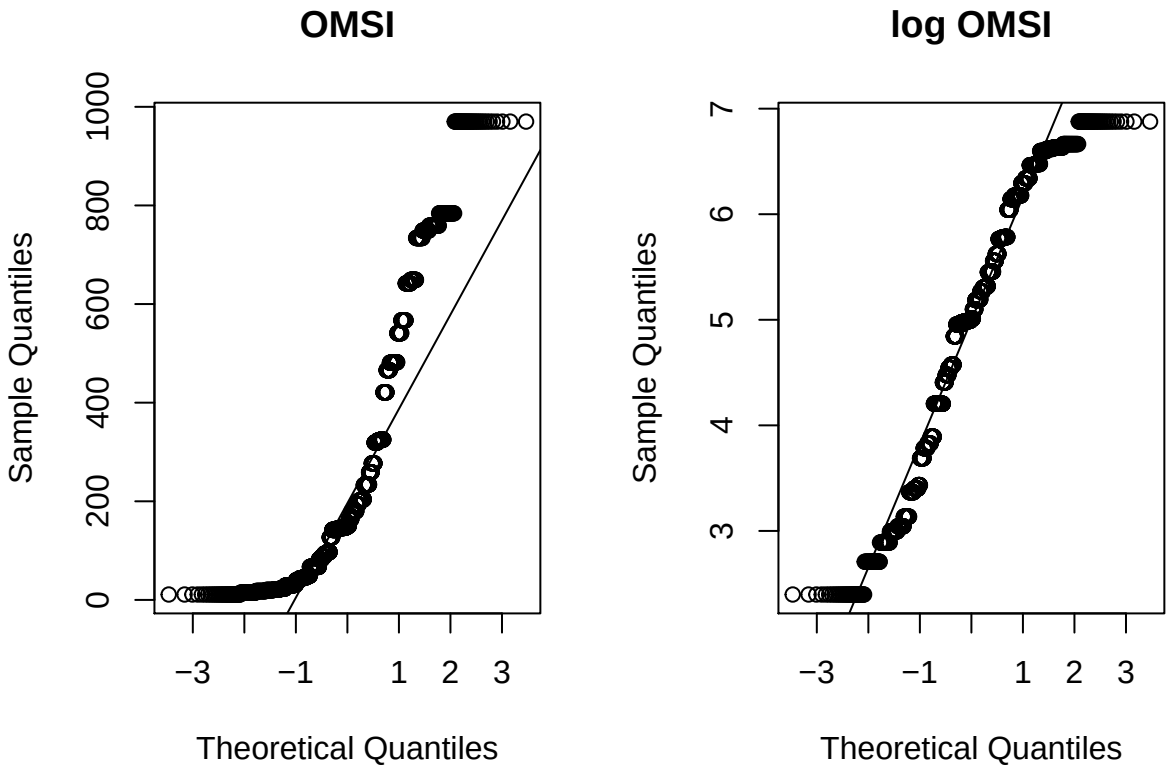



Figure 4: QQ-plots of OMSI: before vs after transformation

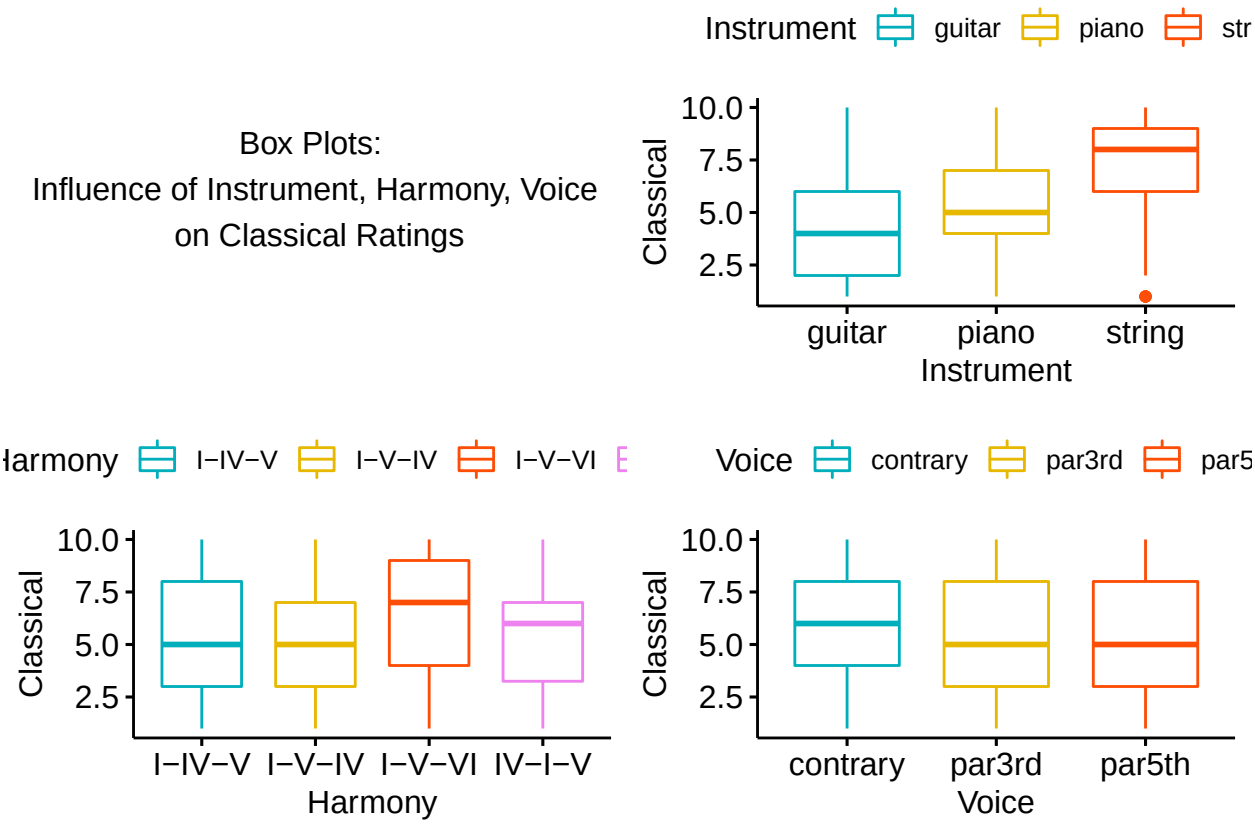



Figure 5: Box plots: influence of instrument, harmony, voice leading on classical ratings

### 3.2 Effects of Experimental Factors on Classical Ratings

Effects of the three main experimental factors on classical ratings were examined by visual comparison of box plots (Figure 5) and ANOVA. The results show that instrument is more influential than harmonic motion and voice leading for classical ratings. Among the three timbres, string quartet received the highest classical ratings on average, whereas guitar was considered least classical sounding.

In terms of harmonic progression, I-V-IV and IV-I-V (two relatively uncommon chord progressions in classical music) received the lowest ratings. I-V-vi was rated as the most classical sounding as expected. The result is subject to respondant's familiarity with the two Pachelbel comedy bits. A multiple regression (Appendix A.4) shows that the average classical ratings for I-V-vi is higher for those who were familar with Pachelbel Rant, lower for those who were familar with Axis's comedy bits.

Voice leading is the least influential factor among the three designed factors. Contrary motion has slightly higher median rating than the other three voice leadings. An interaction between harmony and voice leading is evident, which means the effect of harmonic motion depends on voice leading.

```
##                             Df Sum Sq Mean Sq F value   Pr(>F)
## Instrument                   2   3304  1652.0 339.107  < 2e-16 ***
## Harmony                      3    242    80.5  16.532 1.33e-10 ***
## Voice                        2     51    25.4   5.220 0.005484 **
## KnowRob                      1     67    67.5  13.854 0.000203 ***
## KnowAxis                     1     46    46.1   9.464 0.002126 **
## Instrument:Harmony           6     18     2.9   0.605 0.726228
## Instrument:Voice             4     17     4.3   0.885 0.472008
## Harmony:Voice                6     72    12.0   2.473 0.021915 *
## Harmony:KnowRob              3    113    37.5   7.698 4.12e-05 ***
## Harmony:KnowAxis             3     33    11.0   2.255 0.080154 .
## Instrument:Harmony:Voice    12     67     5.6   1.143 0.319773
## Residuals                 1860   9061     4.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It should be noted that homogeneity of variance is violated in the ANOVA. Levene's Test yields a p-value of 5.626e-05, which means the variance differs significantly across groups.

```
## Levene's Test for Homogeneity of Variance (center = median)
##         Df F value    Pr(>F)
## group   35   2.223 5.626e-05 ***
##       1868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

To determine the best combination of factors predicting classical ratings, the study manually selected the model with the smallest AIC, which gives the best estimates with minimum error while penalizing model complexity. Details can be found in Appendix A.1. Summary of the final model is listed in Table 3.

The model shows that classical ratings differ significantly among subgroups of instruments and harmonic motions. String and piano songs tend to have higher classical ratings than those played by guitar. Stimuli with IV-I-V harmonic motion have been rated as least classical sounding. Voice leading by itself does not have a significant effect on ratings after controlling for other variables, but the effect of harmony depends on voice leading. The interaction term `HarmonyI-V-vi:Voicepar3rd` (with t-statistic=-3.00) suggests that for I-V-vi, contrary motion has been rated as more classical sounding than parallel thirds.

Personal biases in ratings are evident: perhaps person A was more inclined to rate everything as classical, and person B is more inclined to rate everything as popular. The random intercept accounts for such personal biases. Personal biases also exist in the effects of `Instrument`, `Harmony`, `Voice`. That is, personal A probably tends to rate string quartet as more classical than person B. Similar interpretations apply to harmony and

| | coef | SE |
|---|---|---|
| (Intercept) | $3.99^{***}$ | (0.23) |
| Instrumentpiano | $1.44^{***}$ | (0.20) |
| Instrumentstring | $3.25^{***}$ | (0.27) |
| HarmonyI-V-IV | 0.18 | (0.18) |
| HarmonyI-V-vi | $1.26^{***}$ | (0.26) |
| HarmonyIV-I-V | $-0.11$ | (0.17) |
| Voicepar3rd | $-0.21$ | (0.17) |
| Voicepar5th | $-0.18$ | (0.17) |
| HarmonyI-V-IV:Voicepar3rd | $-0.44$ | (0.24) |
| HarmonyI-V-vi:Voicepar3rd | $-0.73^{**}$ | (0.25) |
| HarmonyIV-I-V:Voicepar3rd | $0.50^{*}$ | (0.24) |
| HarmonyI-V-IV:Voicepar5th | $-0.17$ | (0.25) |
| HarmonyI-V-vi:Voicepar5th | $-0.51^{*}$ | (0.25) |
| HarmonyIV-I-V:Voicepar5th | 0.16 | (0.24) |
| AIC | 7577.29 | |
| BIC | 7777.15 | |
| Log Likelihood | -3752.65 | |
| Num. obs. | 1904 | |
| Num. groups: Subject | 54 | |
| Var: Subject (Intercept) | 1.88 | |
| Var: Subject Instrumentpiano | 1.68 | |
| Var: Subject Instrumentstring | 3.66 | |
| Var: Subject HarmonyI-V-IV | 0.08 | |
| Var: Subject HarmonyI-V-vi | 1.91 | |
| Var: Subject HarmonyIV-I-V | 0.04 | |
| Cov: Subject (Intercept) Instrumentpiano | -0.48 | |
| Cov: Subject (Intercept) Instrumentstring | -1.53 | |
| Cov: Subject (Intercept) HarmonyI-V-IV | 0.27 | |
| Cov: Subject (Intercept) HarmonyI-V-vi | -0.05 | |
| Cov: Subject (Intercept) HarmonyIV-I-V | 0.02 | |
| Cov: Subject Instrumentpiano Instrumentstring | 1.56 | |
| Cov: Subject Instrumentpiano HarmonyI-V-IV | -0.26 | |
| Cov: Subject Instrumentpiano HarmonyI-V-vi | -0.47 | |
| Cov: Subject Instrumentpiano HarmonyIV-I-V | -0.12 | |
| Cov: Subject Instrumentstring HarmonyI-V-IV | -0.34 | |
| Cov: Subject Instrumentstring HarmonyI-V-vi | -1.19 | |
| Cov: Subject Instrumentstring HarmonyIV-I-V | 0.01 | |
| Cov: Subject HarmonyI-V-IV HarmonyI-V-vi | 0.11 | |
| Cov: Subject HarmonyI-V-IV HarmonyIV-I-V | 0.04 | |
| Cov: Subject HarmonyI-V-vi HarmonyIV-I-V | 0.04 | |
| Var: Residual | 2.38 | |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3: Effects of the three main experimental factors on classical ratings

voice leading.

Among all random effects, `Instrumentstring` has the largest variance, which means the the effect of string quartet (vs guitar) on classical ratings differs a lot across individuals. `HarmonyIV-I-V` has the smallest variance, which means the effect of IV-I-V (vs I-IV-V) on classical ratings doesn't differ as much as the other random effects across individuals.

Note that the residuals of this model are not perfectly normal (diagnostic plots can be found in Appendix A.1). Adding additional random effects barely improved the situation. Since the distribution is not too crazy, I won't worry too much about it.

### 3.3 Classical Ratings from Musicians vs Non-musicians

Researchers' hypothesis is that people who self-identify as musicians may be influenced by things that do not influence non-musicians. Before hypothesis testing, I dichotomized `Selfdeclare` ("are you a musician?") so that about half the participants were categorized as self-declared musicians, and half not. Using a cutoff at `Selfdeclare=2` (Table 4), we found that `Musician` is correlated with `HarmonyI-V-vi` and the two interaction terms involving I-V-vi. Although I-V-vi was considered most classical sounding in general (Figure 5), musicians and non-musicians rated the harmonic motion differently based on progressions. Particularly, musicians considered I-V-vi contrary (a common progression in both Baroque music and popular genre) as less classical than non-musicians. In contrast, when I-V-vi went with parallel thirds or fifths, musicians gave higher classical ratings than non-musicians. The phenomenon may be explained by small sample bias in the experiment. I-V-vi parallel fifths is a common progression in 1990s and 2000s pop and rock music. However, in this particular data set, `Selfdeclare` is negatively associated with `X1990s2000s` (Figure 3), which means musicians did not listen to 1990s and 2000s pop and rock music as much as non-musicians. Since the musicians in this data set had less exposure to pop music with I-V-vi parallel fifths, they probably assumed the progression was exclusive to classical repertoire and thus identified it as classical sounding. The personal preference of musicians in this data set may not generalize to other samples.

Using a different cutoff at `Selfdeclare=3` (Table 5), `Musician` remains correlated with `HarmonyI-V-vi`. The p-values of the interaction terms for voice leading increased slightly so they are not statistically significant at 0.05 level. Nevertheless, the main takeaways are consistent with that from the previous dichotomization. Therefore we conclude that harmonic motion has different influence for musicians and non-musicians.

### 3.4 Individual Covariates of Classical vs Popular Ratings

Now we move on to the last research question: is there any difference in the factors that drive classical vs popular ratings? Before fitting the model, I converted `CollegeMusic` and `APTheory` to factors since they are both binary. Then I used `fitLMER.fnc` select the best combination of individual covariates. The model selection procedure can be found in Appendix A.2 (for classical ratings) and A.3 (for popular ratings). Summary tables of the final models are displayed in Table 6 and Table 7.

Comparing Table 6 and Table 7, we discovered a few similarities and differences. Among the three designed factors, instrument has the strongest effect on both classical and popular ratings. In terms of harmonic motion, I-V-vi was considered the most classical sounding, whereas I-IV-V was considered most popular sounding. For classical ratings, the effect of harmonic motion is influenced by voice leading, but the interaction is not evident for popular ratings.

In terms of individual covariates, `ConsNotes` and `X16.minus.17` were significant factors of both ratings, but the directions of their influence were not the same. Participants who concentrated on notes and have strong auxiliary abilities tend to give lower classical ratings but higher popular ratings, holding all else constant. An unique factor driving classical ratings was `PlanoPlay`. Participants who played plano rated the stimuli as more classical sounding than those who did not.

Personal biases existed in both classical and popular ratings. Individual differences in the effects of instrument and harmony were clear.

| | coef | SE |
|---|---|---|
| (Intercept) | $3.87^{***}$ | (0.36) |
| MusicianY | 0.20 | (0.46) |
| Instrumentpiano | $1.25^{***}$ | (0.31) |
| Instrumentstring | $2.96^{***}$ | (0.43) |
| HarmonyI-V-IV | 0.48 | (0.27) |
| HarmonyI-V-vi | $2.39^{***}$ | (0.38) |
| HarmonyIV-I-V | −0.18 | (0.27) |
| Voicepar3rd | 0.02 | (0.27) |
| Voicepar5th | −0.11 | (0.27) |
| HarmonyI-V-IV:Voicepar3rd | $-0.83^{*}$ | (0.38) |
| HarmonyI-V-vi:Voicepar3rd | $-1.35^{***}$ | (0.38) |
| HarmonyIV-I-V:Voicepar3rd | 0.61 | (0.38) |
| HarmonyI-V-IV:Voicepar5th | −0.43 | (0.38) |
| HarmonyI-V-vi:Voicepar5th | $-1.16^{**}$ | (0.38) |
| HarmonyIV-I-V:Voicepar5th | 0.46 | (0.38) |
| MusicianY:Instrumentpiano | 0.32 | (0.40) |
| MusicianY:Instrumentstring | 0.49 | (0.55) |
| MusicianY:HarmonyI-V-IV | −0.50 | (0.36) |
| MusicianY:HarmonyI-V-vi | $-1.92^{***}$ | (0.49) |
| MusicianY:HarmonyIV-I-V | 0.12 | (0.35) |
| MusicianY:Voicepar3rd | −0.38 | (0.35) |
| MusicianY:Voicepar5th | −0.13 | (0.35) |
| MusicianY:HarmonyI-V-IV:Voicepar3rd | 0.67 | (0.50) |
| MusicianY:HarmonyI-V-vi:Voicepar3rd | $1.06^{*}$ | (0.50) |
| MusicianY:HarmonyIV-I-V:Voicepar3rd | −0.19 | (0.50) |
| MusicianY:HarmonyI-V-IV:Voicepar5th | 0.44 | (0.50) |
| MusicianY:HarmonyI-V-vi:Voicepar5th | $1.12^{*}$ | (0.50) |
| MusicianY:HarmonyIV-I-V:Voicepar5th | −0.50 | (0.49) |
| AIC | 7582.18 | |
| BIC | 7859.76 | |
| Log Likelihood | -3741.09 | |
| Num. obs. | 1904 | |
| Num. groups: Subject | 54 | |
| Var: Subject (Intercept) | 1.87 | |
| Var: Subject Instrumentpiano | 1.66 | |
| Var: Subject Instrumentstring | 3.60 | |
| Var: Subject HarmonyI-V-IV | 0.06 | |
| Var: Subject HarmonyI-V-vi | 1.53 | |
| Var: Subject HarmonyIV-I-V | 0.02 | |
| Cov: Subject (Intercept) Instrumentpiano | -0.49 | |
| Cov: Subject (Intercept) Instrumentstring | -1.53 | |
| Cov: Subject (Intercept) HarmonyI-V-IV | 0.28 | |
| Cov: Subject (Intercept) HarmonyI-V-vi | -0.01 | |
| Cov: Subject (Intercept) HarmonyIV-I-V | 0.03 | |
| Cov: Subject Instrumentpiano Instrumentstring | 1.52 | |
| Cov: Subject Instrumentpiano HarmonyI-V-IV | -0.25 | |
| Cov: Subject Instrumentpiano HarmonyI-V-vi | -0.38 | |
| Cov: Subject Instrumentpiano HarmonyIV-I-V | -0.11 | |
| Cov: Subject Instrumentstring HarmonyI-V-IV | -0.32 | |
| Cov: Subject Instrumentstring HarmonyI-V-vi | -1.06 | |
| Cov: Subject Instrumentstring HarmonyIV-I-V | 0.02 | |
| Cov: Subject HarmonyI-V-IV HarmonyI-V-vi | 0.02 | |
| Cov: Subject HarmonyI-V-IV HarmonyIV-I-V | 0.02 | |
| Cov: Subject HarmonyI-V-vi HarmonyIV-I-V | -0.02 | |
| Var: Residual | 2.36 | |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$



Table 4: Classical Ratings from Musicians vs Non-musicians (cutoff at Selfdeclare=2)

| | coef |
|---|---|
| (Intercept) | 3.94 (0.46)*** |
| MusicianY | 0.07 (0.53) |
| Instrumentpiano | 1.21 (0.40)** |
| Instrumentstring | 2.60 (0.55)*** |
| HarmonyI-V-IV | 0.71 (0.36)* |
| HarmonyI-V-vi | 2.83 (0.49)*** |
| HarmonyIV-I-V | 0.13 (0.35) |
| Voicepar3rd | 0.03 (0.35) |
| Voicepar5th | 0.02 (0.35) |
| HarmonyI-V-IV:Voicepar3rd | −0.97 (0.49) |
| HarmonyI-V-vi:Voicepar3rd | −1.50 (0.50)** |
| HarmonyIV-I-V:Voicepar3rd | 0.47 (0.49) |
| HarmonyI-V-IV:Voicepar5th | −0.84 (0.50) |
| HarmonyI-V-vi:Voicepar5th | −1.14 (0.50)* |
| HarmonyIV-I-V:Voicepar5th | 0.36 (0.49) |
| MusicianY:Instrumentpiano | 0.31 (0.46) |
| MusicianY:Instrumentstring | 0.86 (0.63) |
| MusicianY:HarmonyI-V-IV | −0.70 (0.41) |
| MusicianY:HarmonyI-V-vi | −2.08 (0.56)*** |
| MusicianY:HarmonyIV-I-V | −0.32 (0.40) |
| MusicianY:Voicepar3rd | −0.31 (0.40) |
| MusicianY:Voicepar5th | −0.27 (0.40) |
| MusicianY:HarmonyI-V-IV:Voicepar3rd | 0.70 (0.57) |
| MusicianY:HarmonyI-V-vi:Voicepar3rd | 1.01 (0.57) |
| MusicianY:HarmonyIV-I-V:Voicepar3rd | 0.03 (0.57) |
| MusicianY:HarmonyI-V-IV:Voicepar5th | 0.88 (0.57) |
| MusicianY:HarmonyI-V-vi:Voicepar5th | 0.84 (0.57) |
| MusicianY:HarmonyIV-I-V:Voicepar5th | −0.27 (0.57) |
| AIC | 7584.80 |
| BIC | 7862.39 |
| Log Likelihood | -3742.40 |
| Num. obs. | 1904 |
| Num. groups: Subject | 54 |
| Var: Subject (Intercept) | 1.87 |
| Var: Subject Instrumentpiano | 1.66 |
| Var: Subject Instrumentstring | 3.52 |
| Var: Subject HarmonyI-V-IV | 0.07 |
| Var: Subject HarmonyI-V-vi | 1.53 |
| Var: Subject HarmonyIV-I-V | 0.03 |
| Cov: Subject (Intercept) Instrumentpiano | -0.48 |
| Cov: Subject (Intercept) Instrumentstring | -1.51 |
| Cov: Subject (Intercept) HarmonyI-V-IV | 0.27 |
| Cov: Subject (Intercept) HarmonyI-V-vi | -0.09 |
| Cov: Subject (Intercept) HarmonyIV-I-V | 0.01 |
| Cov: Subject Instrumentpiano Instrumentstring | 1.51 |
| Cov: Subject Instrumentpiano HarmonyI-V-IV | -0.25 |
| Cov: Subject Instrumentpiano HarmonyI-V-vi | -0.39 |
| Cov: Subject Instrumentpiano HarmonyIV-I-V | -0.10 |
| Cov: Subject Instrumentstring HarmonyI-V-IV | -0.31 |
| Cov: Subject Instrumentstring HarmonyI-V-vi | -0.96 |
| Cov: Subject Instrumentstring HarmonyIV-I-V | 0.07 |
| Cov: Subject HarmonyI-V-IV HarmonyI-V-vi | 0.07 |
| Cov: Subject HarmonyI-V-IV HarmonyIV-I-V | 0.02 |
| Cov: Subject HarmonyI-V-vi HarmonyIV-I-V | -0.06 |
| Var: Residual | 2.36 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$



Table 5: Classical Ratings from Musicians vs Non-musicians (cutoff at Selfdeclare=3)

| | coef |
|---|---|
| (Intercept) | 4.53 (0.28)*** |
| Instrumentpiano | 1.44 (0.20)*** |
| Instrumentstring | 3.25 (0.27)*** |
| HarmonyI-V-IV | 0.19 (0.18) |
| HarmonyI-V-vi | 1.26 (0.26)*** |
| HarmonyIV-I-V | −0.11 (0.17) |
| Voicepar3rd | −0.20 (0.17) |
| Voicepar5th | −0.18 (0.17) |
| X16.minus.17 | −0.15 (0.05)*** |
| ConsNotes | −0.25 (0.07)*** |
| PianoPlay | 0.26 (0.08)*** |
| HarmonyI-V-IV:Voicepar3rd | −0.44 (0.24) |
| HarmonyI-V-vi:Voicepar3rd | −0.75 (0.24)** |
| HarmonyIV-I-V:Voicepar3rd | 0.49 (0.24)* |
| HarmonyI-V-IV:Voicepar5th | −0.19 (0.24) |
| HarmonyI-V-vi:Voicepar5th | −0.51 (0.25)* |
| HarmonyIV-I-V:Voicepar5th | 0.15 (0.24) |
| AIC | 7563.59 |
| BIC | 7780.10 |
| Log Likelihood | -3742.79 |
| Num. obs. | 1904 |
| Num. groups: Subject | 54 |
| Var: Subject (Intercept) | 1.53 |
| Var: Subject Instrumentpiano | 1.68 |
| Var: Subject Instrumentstring | 3.65 |
| Var: Subject HarmonyI-V-IV | 0.11 |
| Var: Subject HarmonyI-V-vi | 1.91 |
| Var: Subject HarmonyIV-I-V | 0.04 |
| Cov: Subject (Intercept) Instrumentpiano | -0.62 |
| Cov: Subject (Intercept) Instrumentstring | -1.66 |
| Cov: Subject (Intercept) HarmonyI-V-IV | 0.35 |
| Cov: Subject (Intercept) HarmonyI-V-vi | 0.23 |
| Cov: Subject (Intercept) HarmonyIV-I-V | 0.08 |
| Cov: Subject Instrumentpiano Instrumentstring | 1.56 |
| Cov: Subject Instrumentpiano HarmonyI-V-IV | -0.25 |
| Cov: Subject Instrumentpiano HarmonyI-V-vi | -0.48 |
| Cov: Subject Instrumentpiano HarmonyIV-I-V | -0.12 |
| Cov: Subject Instrumentstring HarmonyI-V-IV | -0.33 |
| Cov: Subject Instrumentstring HarmonyI-V-vi | -1.18 |
| Cov: Subject Instrumentstring HarmonyIV-I-V | 0.01 |
| Cov: Subject HarmonyI-V-IV HarmonyI-V-vi | 0.13 |
| Cov: Subject HarmonyI-V-IV HarmonyIV-I-V | 0.05 |
| Cov: Subject HarmonyI-V-vi HarmonyIV-I-V | 0.04 |
| Var: Residual | 2.37 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 6: Final model predicting classical ratings

| | coef |
|---|---|
| (Intercept) | 6.25 (0.27)*** |
| Instrumentpiano | −1.05 (0.19)*** |
| Instrumentstring | −2.78 (0.26)*** |
| HarmonyI-V-IV | −0.04 (0.11) |
| HarmonyI-V-vi | −0.37 (0.17)* |
| HarmonyIV-I-V | −0.30 (0.12)* |
| Voicepar3rd | 0.22 (0.09)* |
| Voicepar5th | 0.26 (0.09)** |
| X16.minus.17 | 0.10 (0.05)* |
| ConsNotes | 0.13 (0.07) |
| AIC | 7586.80 |
| BIC | 7764.46 |
| Log Likelihood | -3761.40 |
| Num. obs. | 1904 |
| Num. groups: Subject | 54 |
| Var: Subject (Intercept) | 1.41 |
| Var: Subject Instrumentpiano | 1.50 |
| Var: Subject Instrumentstring | 3.17 |
| Var: Subject HarmonyI-V-IV | 0.15 |
| Var: Subject HarmonyI-V-vi | 1.02 |
| Var: Subject HarmonyIV-I-V | 0.28 |
| Cov: Subject (Intercept) Instrumentpiano | -0.31 |
| Cov: Subject (Intercept) Instrumentstring | -0.89 |
| Cov: Subject (Intercept) HarmonyI-V-IV | 0.21 |
| Cov: Subject (Intercept) HarmonyI-V-vi | -0.03 |
| Cov: Subject (Intercept) HarmonyIV-I-V | -0.20 |
| Cov: Subject Instrumentpiano Instrumentstring | 1.55 |
| Cov: Subject Instrumentpiano HarmonyI-V-IV | -0.19 |
| Cov: Subject Instrumentpiano HarmonyI-V-vi | -0.33 |
| Cov: Subject Instrumentpiano HarmonyIV-I-V | -0.21 |
| Cov: Subject Instrumentstring HarmonyI-V-IV | -0.35 |
| Cov: Subject Instrumentstring HarmonyI-V-vi | -0.50 |
| Cov: Subject Instrumentstring HarmonyIV-I-V | -0.23 |
| Cov: Subject HarmonyI-V-IV HarmonyI-V-vi | -0.05 |
| Cov: Subject HarmonyI-V-IV HarmonyIV-I-V | -0.08 |
| Cov: Subject HarmonyI-V-vi HarmonyIV-I-V | -0.16 |
| Var: Residual | 2.41 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 7: Final model predicting popular ratings

Note that the residuals of neither model is perfectly normal (diagnostic plots can be found in Appendix A.3). Adding additional random effects barely improved the situation. Since the distribution is not too crazy, I won't worry too much about it.

## 4 Discussion

The data collected from Ivan Jimenez suggests that instrument has the strongest influence on classical ratings, followed by harmonic motion and voice leading. String quartet was rated as the most classical sounding, whereas guitar was rated as the least classical sounding. Among the levels of harmonic motions, I-V-vi has the highest classical ratings in general, but that depends on respondant's familiarity with the Pachelbel comedy bits. Respondants who were familar with Pachelbel Rant gave higher average classical ratings for I-V-vi; those who were familar with Axis's comedy bits gave lower average classical ratings for I-V-vi. Voice leading is the least influential experimental factor. Among levels of voice leading, contrary motion received slightly higher ratings than the other two categories. An interaction between harmony and voice leading was evident, which means the effect of harmonic motion depends on voice leading.

The data confirms researchers' hypothesis that people who self-identify as musicians may be influenced by things that do not influence non-musicians. Although I-V-vi harmonic motion received the highest classical ratings in general, musicians tend to gave lower ratings when it came with contrary motion and higher ratings when it came with parallel thirds or fifths. As previously mentioned, the finding may be subject to small sample bias in the experiment.

This study also finds that classical and popular ratings were driven by different sets of factors. The interaction between harmotic motion and voice leading was significant only in classical ratings but not in popular ratings. Another unique factor driving classical ratings was `PlanoPlay`. Participants who played plano tend to give higher classical ratings than those who did not. It should be noted that personal biases existed in both classical and popular ratings. That is, perhaps person A was more inclined to rate everything as classical, and person B was more inclined to rate everything as popular. The effect of instrument and harmony also differed among individuals.

The findings of this study have to be seen in light of some limitations. First, I used MICE to handle missing values, which required the data to be missing at random. That is, we assumed whether an observation was missing had nothing to do with the missing values. Given limited context, we were unable to check if the assumption holds true for this data set. If participants who did not report `NoClass` happened to be those who did not take music classes, then our model would be biased. And our conclusions may not generalize to individuals with little musical training.

Secondly, as briefly discussed in Result 3.1, unexpected associations between variables hint at potential sampling biases in this experiment. The project mentioned that the participants were recruited from the population of undergraduates at the University of Pittsburgh, but we did not know how they were recruited and who were recruited. There is no way we could measure how representative our conclusions may be. Also, the experimental design could be improved. The experiment did a great job in evaluating the listeners' musical training, but it did not specify what type of music the participants were trained to compose or to play. Taking college music course as an example, were the participants taking Baroque music history or popular composition? The choice matters a lot in the music structure they were mostly exposed to and perhaps good at identifying. I would recommend future works to take that into account.

## References

Hamm, C. (1995), *Putting popular music in its place*, Cambridge: Cambridge University Press, p. 3.

R Core Team (2017), *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing.

# Appendix

### A.1 Effects of Instrument, Harmony, and Voice Leading on Classical Ratings

*Fixed Effects.* Instrument is the most influential factor of classical ratings, followed by harmonic motion and voice leading. An interaction between harmony and voice leading is evident, which means the effect of harmonic motion depends on voice leadings.

```
##                            Df Sum Sq Mean Sq F value   Pr(>F)
## Instrument                  2   3304  1652.0 331.119  < 2e-16 ***
## Harmony                     3    242    80.5  16.142 2.32e-10 ***
## Voice                       2     51    25.4   5.098   0.0062 **
## Instrument:Harmony          6     18     2.9   0.586   0.7421
## Instrument:Voice            4     17     4.3   0.859   0.4876
## Harmony:Voice               6     72    12.1   2.421   0.0247 *
## Instrument:Harmony:Voice   12     67     5.6   1.127   0.3331
## Residuals                1868   9320     5.0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It should be noted that homogeneity of variance is violated in the ANOVA. Levene's Test yields a p-value of 5.626e-05, which means the variance differs significantly across groups.

```
## Levene's Test for Homogeneity of Variance (center = median)
##         Df F value    Pr(>F)
## group   35   2.223 5.626e-05 ***
##       1868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*Random Intercept.* The table below compares an lm model without random intercept and an lmer model with random intercept. Both AIC and BIC choose the lmer model, which means the random intercept is needed.

Table 8: Model comparison: fixed model vs random intercept

| Model | AIC | BIC |
|---|---|---|
| lm | 8477.91 | 8561.19 |
| lmer | 8019.58 | 8108.41 |

The conclusion from the random intercept model is consistent with the lm model. There's no huge shift in significance of the predictors. Classical ratings differ significantly among subgroups of instruments and harmonic motions. The effect of harmony depends on voice leading, since the interaction term `HarmonyI-V-vi:Voicepar3rd` has a large t-statistic. Songs played by string and piano tend to have higher classical ratings than those played by guitar. Songs with I-IV-V harmony have the lowest classical ratings. "Personal biases" in ratings are evident: perhaps person A is more inclined to rate everything as classical, and person B is more inclined to rate everything as popular. The random intercept accounts for such personal biases.

```
## lmer(formula = Classical ~ Instrument + Harmony * Voice + (1 |
##     Subject), data = ratings, REML = F)
##                             coef.est coef.se t value
## (Intercept)                  3.98     0.23   17.59
## Instrumentpiano              1.44     0.11   13.57
## Instrumentstring             3.23     0.11   30.19
## HarmonyI-V-IV                0.19     0.21    0.90
## HarmonyI-V-VI                1.28     0.21    5.98
## HarmonyIV-I-V               -0.11     0.21   -0.50
## Voicepar3rd                 -0.19     0.21   -0.91
## Voicepar5th                 -0.18     0.21   -0.83
## HarmonyI-V-IV:Voicepar3rd -0.46     0.30   -1.54
## HarmonyI-V-VI:Voicepar3rd -0.76     0.30   -2.51
## HarmonyIV-I-V:Voicepar3rd  0.48     0.30    1.60
## HarmonyI-V-IV:Voicepar5th -0.19     0.30   -0.62
## HarmonyI-V-VI:Voicepar5th -0.51     0.30   -1.70
## HarmonyIV-I-V:Voicepar5th  0.16     0.30    0.53
##
## Error terms:
##  Groups   Name        Std.Dev.
##  Subject  (Intercept) 1.15
##  Residual             1.90
## ---
## number of obs: 1904, groups: Subject, 54
## AIC = 8019.6, DIC = 7987.6
## deviance = 7987.6
```

*Random Effects.* Both AIC and BIC prefer the lmer model with random effects for `Instrument` and `Harmony`. ANOVA also supports the conclusion. `lmer.2` and `lmer.3` have small p-values, which means they are significant improvements on the previous model. A random effect for `Voice` is not need, since `lmer.4` has a large p-value in the ANOVA.

Table 9: Random effects for classical ratings

| Model | AIC | BIC |
|---|---|---|
| (1 \| Subject) | 8019.58 | 8108.41 |
| (1+Instrument \| Subject) | 7716.97 | 7833.56 |
| (1+Instrument+Harmony \| Subject) | 7577.29 | 7777.15 |
| (1+Instrument+Harmony+Voice \| Subject) | 7586.86 | 7870.00 |

```
## Data: ratings
## Models:
## lmer.1: Classical ~ Instrument + Harmony * Voice + (1 | Subject)
## lmer.2: Classical ~ Instrument + Harmony * Voice + (1 + Instrument |
## lmer.2:     Subject)
## lmer.3: Classical ~ Instrument + Harmony * Voice + (1 + Instrument +
## lmer.3:     Harmony | Subject)
## lmer.4: Classical ~ Instrument + Harmony * Voice + (1 + Instrument +
## lmer.4:     Harmony + Voice | Subject)
##        Df    AIC    BIC  logLik deviance   Chisq Chi Df Pr(>Chisq)
## lmer.1 16 8019.6 8108.4 -3993.8   7987.6
## lmer.2 21 7717.0 7833.6 -3837.5   7675.0 312.613      5     <2e-16 ***
## lmer.3 36 7577.3 7777.2 -3752.6   7505.3 169.680     15     <2e-16 ***
## lmer.4 51 7586.9 7870.0 -3742.4   7484.9  20.426     15     0.1562
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The model shows that classical ratings differ significantly among subgroups of instruments and harmonic motions. String and piano songs tend to have higher classical ratings than those played by guitar. Songs with IV-I-V harmonic motion have been rated as least classical sounding. Voice leading by itself does not have a significant effect on ratings after controlling for other variables, but the effect of harmony depends on voice leading. The interaction term `HarmonyI-V-vi:Voicepar3rd` (with t-statistic=-3.00) suggests that within I-V-vi, contrary motion has been rated as more classical sounding than parallel thirds.

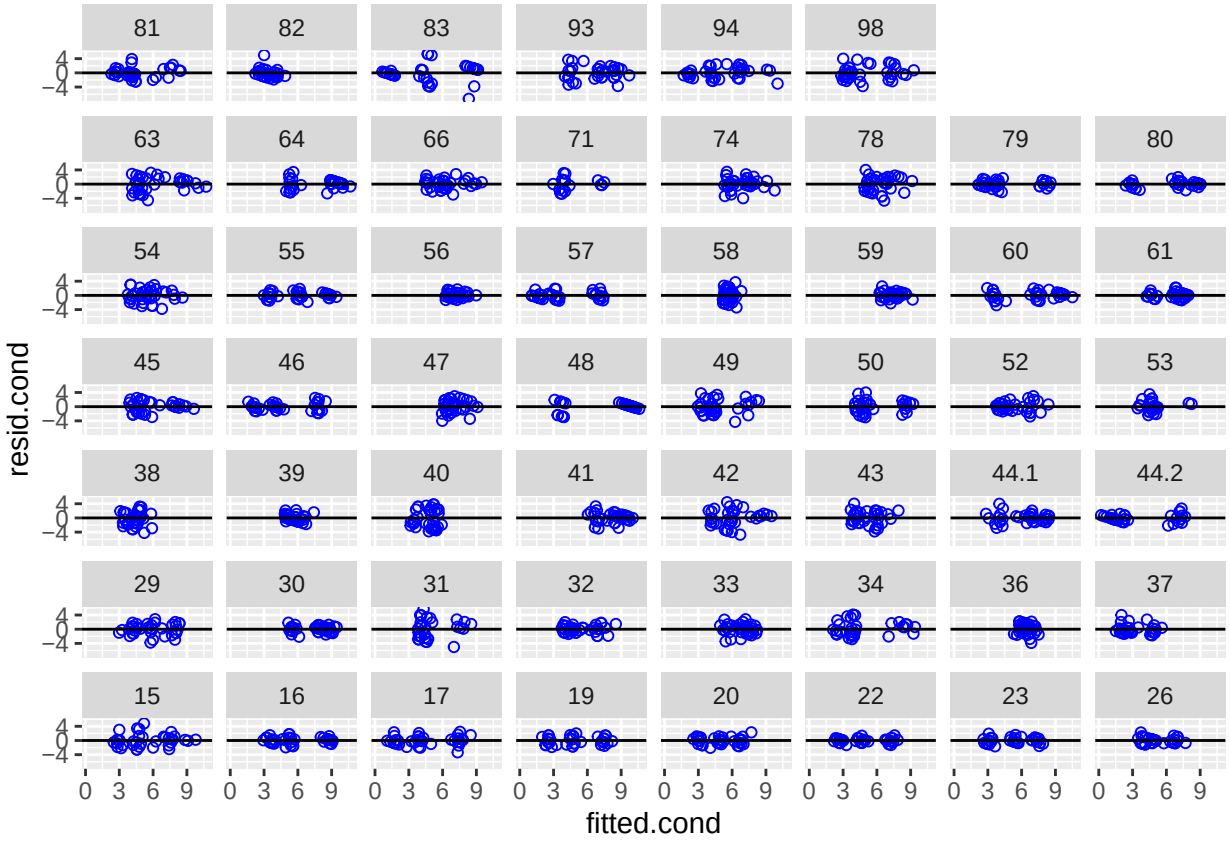
Personal biases in ratings are evident: perhaps person A was more inclined to rate everything as classical, and person B is more inclined to rate everything as popular. The random intercept accounts for such personal biases. Personal biases also exists in the effects of `Instrument`, `Harmony`, `Voice`. That is, personal A probably tends to rate string quartet music as more classical than person B. Similar interpretations apply to harmony and voice leading.

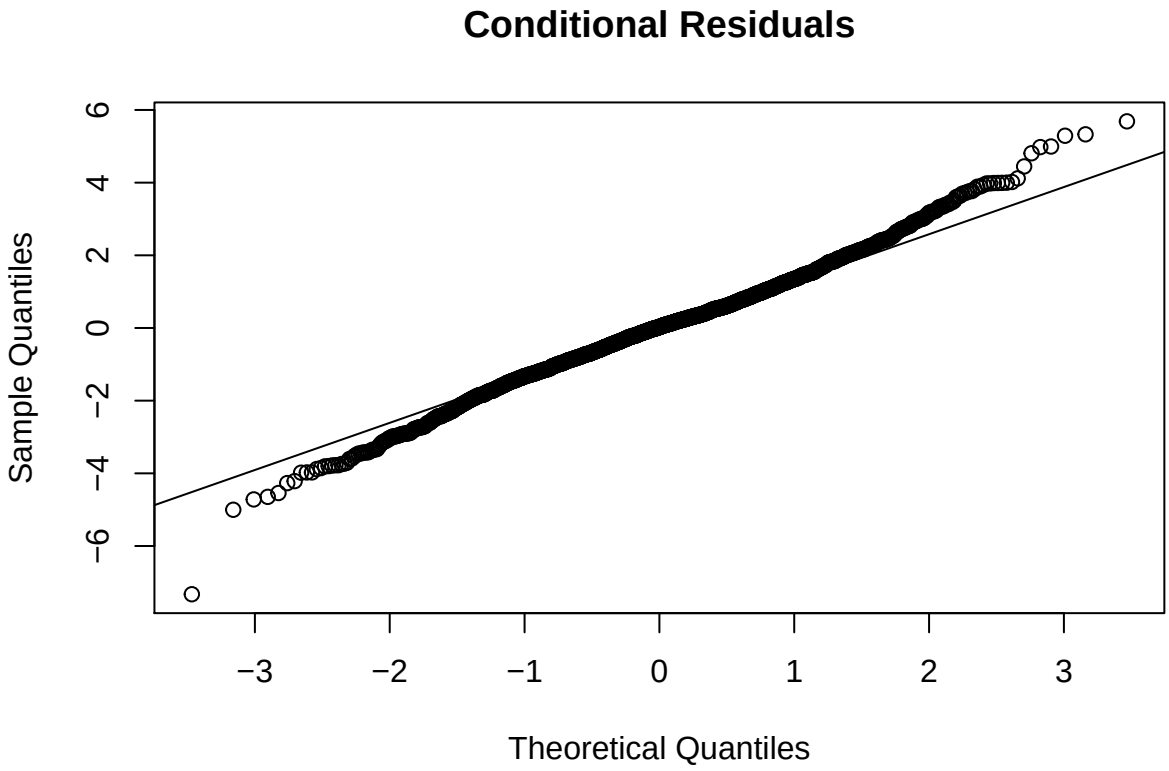
Among all random effects, `Instrumentstring` has the largest variance, which means the the effect of string quartet (vs guitar) on classical ratings differs a lot across individuals. `HarmonyIV-I-V` has the smallest variance, which means the effect of HarmonyIV-I-V (vs HarmonyI-IV-V) on classical ratings doesn't differ as much as the other random effects across individuals.

```
## lmer(formula = Classical ~ Instrument + Harmony * Voice + (1 +
##     Instrument + Harmony | Subject), data = ratings, REML = F)
##                             coef.est coef.se
## (Intercept)                  3.99     0.23
## Instrumentpiano              1.44     0.20
## Instrumentstring             3.25     0.27
## HarmonyI-V-IV                0.18     0.18
## HarmonyI-V-VI                1.26     0.26
## HarmonyIV-I-V               -0.11     0.17
## Voicepar3rd                 -0.21     0.17
## Voicepar5th                 -0.18     0.17
## HarmonyI-V-IV:Voicepar3rd -0.44     0.24
```

```
## HarmonyI-V-VI:Voicepar3rd -0.73     0.25
## HarmonyIV-I-V:Voicepar3rd  0.50     0.24
## HarmonyI-V-IV:Voicepar5th -0.17     0.25
## HarmonyI-V-VI:Voicepar5th -0.51     0.25
## HarmonyIV-I-V:Voicepar5th  0.16     0.24
##
## Error terms:
##  Groups   Name              Std.Dev. Corr
##  Subject  (Intercept)       1.37
##           Instrumentpiano   1.29     -0.27
##           Instrumentstring  1.91     -0.58  0.63
##           HarmonyI-V-IV     0.28      0.72 -0.73 -0.64
##           HarmonyI-V-VI     1.38     -0.03 -0.26 -0.45  0.30
##           HarmonyIV-I-V     0.19      0.09 -0.49  0.03  0.66  0.14
##  Residual                   1.54
## ---
## number of obs: 1904, groups: Subject, 54
## AIC = 7577.3, DIC = 7505.3
## deviance = 7505.3
```

*Conditional Residuals.* The conditional residuals for the final model is mean zero. The skewness in QQ-plot suggests the residuals are not perfectly normal but not too crazy either. I won't worry too much about it.






### A.2 Individual Covariates of Classical Ratings

Before fitting the model, I converted `CollegeMusic` and `APTheory` to factors since they are binary. Then I used `fitLMER.fnc` select the best combination of individual covariates. The result suggests that `X16.minus.17`, `ConsNotes`, `PianoPlay` should be added as fixed effects in the model.

```
## lmer(formula = Classical ~ Instrument + Harmony + Voice + X16.minus.17 +
##     ConsNotes + PianoPlay + (1 + Instrument + Harmony | Subject) +
##     Harmony:Voice, data = ratings, REML = FALSE)
##                             coef.est coef.se
## (Intercept)                  4.53     0.28
## Instrumentpiano              1.44     0.20
## Instrumentstring             3.25     0.27
## HarmonyI-V-IV                0.19     0.18
## HarmonyI-V-VI                1.26     0.26
## HarmonyIV-I-V               -0.11     0.17
## Voicepar3rd                 -0.20     0.17
## Voicepar5th                 -0.18     0.17
## X16.minus.17                -0.15     0.05
## ConsNotes                   -0.25     0.07
## PianoPlay                    0.26     0.08
## HarmonyI-V-IV:Voicepar3rd -0.44     0.24
## HarmonyI-V-VI:Voicepar3rd -0.75     0.24
## HarmonyIV-I-V:Voicepar3rd  0.49     0.24
## HarmonyI-V-IV:Voicepar5th -0.19     0.24
## HarmonyI-V-VI:Voicepar5th -0.51     0.25
```

```
## HarmonyIV-I-V:Voicepar5th  0.15     0.24
##
## Error terms:
##  Groups   Name              Std.Dev. Corr
##  Subject  (Intercept)       1.24
##           Instrumentpiano   1.29     -0.39
##           Instrumentstring  1.91     -0.71  0.63
##           HarmonyI-V-IV     0.33      0.86 -0.59 -0.53
##           HarmonyI-V-VI     1.38      0.14 -0.27 -0.45  0.28
##           HarmonyIV-I-V     0.21      0.29 -0.42  0.03  0.71  0.15
##  Residual                   1.54
## ---
## number of obs: 1904, groups: Subject, 54
## AIC = 7563.6, DIC = 7485.6
## deviance = 7485.6
```

Then I used `fitLMER.fnc` again to check if there is any change in random effects. Note that `fitLMER.fnc` only works for *independent* random effects, such as `(Instrument | Subject)` and `(Harmony | Subject)`. After automated selection, I grouped the selected random effects into one bracket to allow for interactions. That also helped to decrease AIC slightly. It turned out only random effects for harmony and instrument are needed. The final model is given below.

```
## lmer(formula = Classical ~ Instrument + Harmony * Voice + X16.minus.17 +
##     ConsNotes + PianoPlay + (1 + Instrument + Harmony | Subject),
##     data = ratings, REML = FALSE)
##                            coef.est coef.se
## (Intercept)                 4.53     0.28
## Instrumentpiano             1.44     0.20
## Instrumentstring            3.25     0.27
## HarmonyI-V-IV               0.19     0.18
## HarmonyI-V-VI               1.26     0.26
## HarmonyIV-I-V              -0.11     0.17
## Voicepar3rd                -0.20     0.17
## Voicepar5th                -0.18     0.17
## X16.minus.17               -0.15     0.05
## ConsNotes                  -0.25     0.07
## PianoPlay                   0.26     0.08
## HarmonyI-V-IV:Voicepar3rd -0.44     0.24
## HarmonyI-V-VI:Voicepar3rd -0.75     0.24
## HarmonyIV-I-V:Voicepar3rd  0.49     0.24
## HarmonyI-V-IV:Voicepar5th -0.19     0.24
## HarmonyI-V-VI:Voicepar5th -0.51     0.25
## HarmonyIV-I-V:Voicepar5th  0.15     0.24
##
## Error terms:
##  Groups   Name              Std.Dev. Corr
##  Subject  (Intercept)       1.24
##           Instrumentpiano   1.29     -0.39
##           Instrumentstring  1.91     -0.71  0.63
##           HarmonyI-V-IV     0.33      0.86 -0.59 -0.53
##           HarmonyI-V-VI     1.38      0.14 -0.27 -0.45  0.28
##           HarmonyIV-I-V     0.21      0.29 -0.42  0.03  0.71  0.15
##  Residual                   1.54
## ---
## number of obs: 1904, groups: Subject, 54
```

```
## AIC = 7563.6, DIC = 7485.6
## deviance = 7485.6
```

*Conditional Residuals.* The conditional residuals for the final model is mean zero. The skewness in QQ-plot suggests the residuals are not perfectly normal but not too crazy either. I won't worry too much about it.

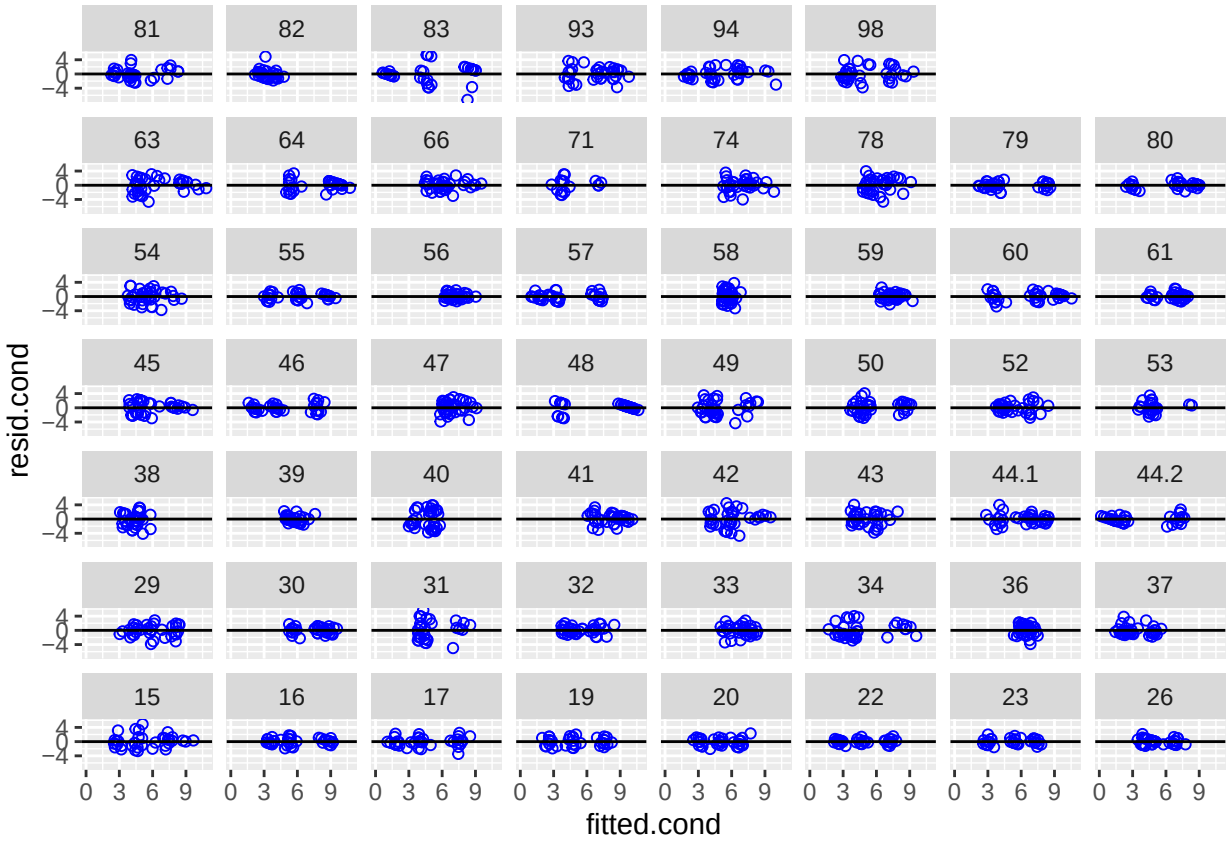
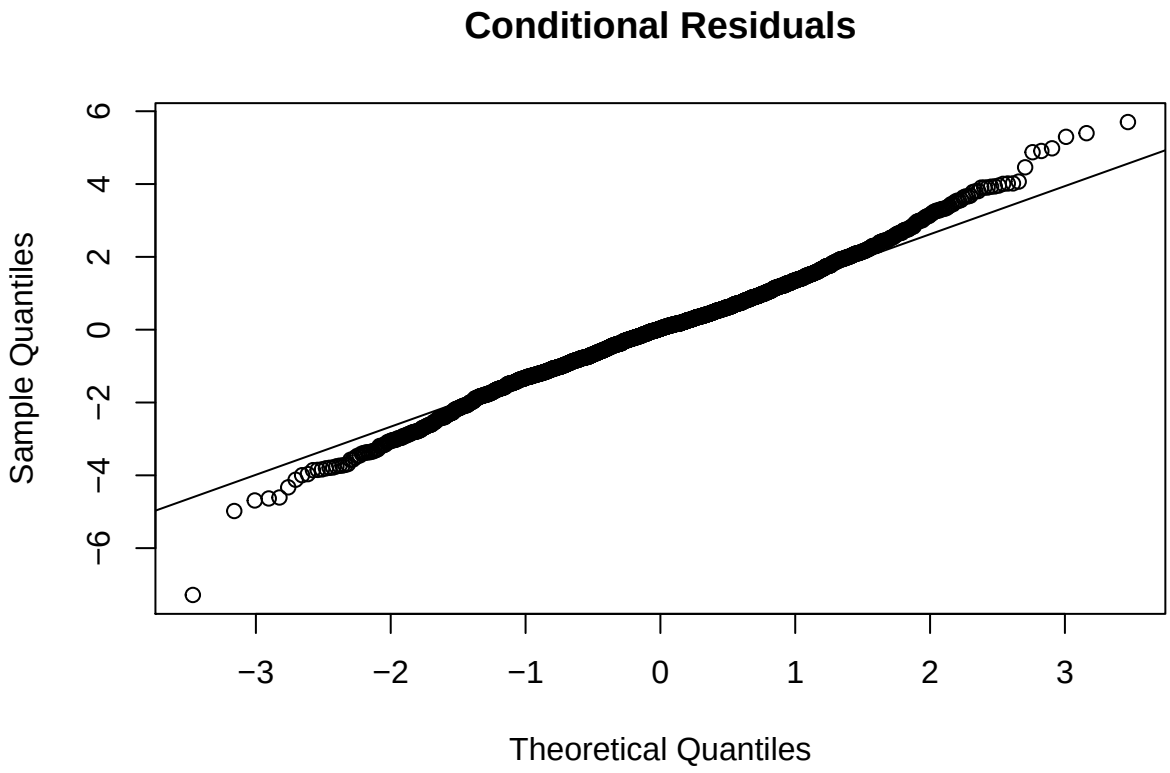


Table 10: Model comparison: with vs without random intercept

| Model | AIC | BIC |
|---|---|---|
| fixed model | 8402.19 | 8452.15 |
| random intercept | 7924.07 | 7979.59 |




### A.3 Model Selection for Popular Ratings

*Main Experimental Factors.* ANOVA shows that `Instrument` has the strongest effect on popular ratings among the three main experimental factors. Guitar music have higher popular ratings than piano and string quartet. Harmonic motion also have an impact on popular ratings. Voice leading does not matter. None of the interactions are needed.

```
##                            Df Sum Sq Mean Sq F value Pr(>F)
## Instrument                  2   2409  1204.4 250.483 <2e-16 ***
## Harmony                     3     49    16.4   3.403 0.0171 *
## Voice                       2     25    12.5   2.593 0.0751 .
## Instrument:Harmony          6     12     2.1   0.427 0.8613
## Instrument:Voice            4     16     3.9   0.809 0.5191
## Harmony:Voice               6     37     6.1   1.272 0.2670
## Instrument:Harmony:Voice   12     65     5.4   1.121 0.3379
## Residuals                1868   8982     4.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*Random Intercept.* Random intercept is needed, as adding it decreases AIC and BIC.

Table 11: Random effects for popular ratings

| Model | AIC | BIC |
|---|---|---|
| (1 \| Subject) | 7924.07 | 7979.59 |
| (1+Instrument \| Subject) | 7673.88 | 7790.46 |
| (1+Instrument+Harmony \| Subject) | 7584.94 | 7784.80 |
| (1+Instrument+Harmony+Voice \| Subject) | 7598.70 | 7881.83 |

*Random Effects.* Both AIC and BIC prefer the lmer model with random effects for `Instrument` and `Harmony`. ANOVA also supports the conclusion. `lmer.2` and `lmer.3` have small p-values, which means they are significant improvements on the previous model. A random effect for `Voice` is not need, since `lmer.4` has a large p-value in the ANOVA.

```
## Data: ratings
## Models:
## lmer.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject)
## lmer.2: Popular ~ Instrument + Harmony * Voice + (1 + Instrument | Subject)
## lmer.3: Popular ~ Instrument + Harmony * Voice + (1 + Instrument + Harmony |
## lmer.3:     Subject)
## lmer.4: Popular ~ Instrument + Harmony * Voice + (1 + Instrument + Harmony +
## lmer.4:     Voice | Subject)
##        Df    AIC    BIC  logLik deviance   Chisq Chi Df Pr(>Chisq)
## lmer.1 10 7924.1 7979.6 -3952.0   7904.1
## lmer.2 21 7673.9 7790.5 -3815.9   7631.9 272.198     11     <2e-16 ***
## lmer.3 36 7584.9 7784.8 -3756.5   7512.9 118.938     15     <2e-16 ***
## lmer.4 51 7598.7 7881.8 -3748.3   7496.7  16.243     15     0.3661
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*Individual Covariates.* Before fitting the model, I converted `CollegeMusic` and `APTheory` to factors since they are binary. Then I used `fitLMER.fnc` select the best combination of individual covariates. The result suggests that `X16.minus.17` and `ConsNotes`should be added as fixed effects in the model.

```
## lmer(formula = Popular ~ Instrument + Harmony + Voice + X16.minus.17 +
##     ConsNotes + (1 + Instrument + Harmony | Subject), data = ratings,
##     REML = FALSE)
##                  coef.est coef.se t value
## (Intercept)         6.25     0.27   22.93
## Instrumentpiano    -1.05     0.19   -5.60
## Instrumentstring   -2.78     0.26  -10.77
## HarmonyI-V-IV      -0.04     0.11   -0.38
## HarmonyI-V-VI      -0.37     0.17   -2.18
## HarmonyIV-I-V      -0.30     0.12   -2.42
## Voicepar3rd         0.22     0.09    2.57
## Voicepar5th         0.26     0.09    3.00
## X16.minus.17        0.10     0.05    2.11
## ConsNotes           0.13     0.07    1.88
##
## Error terms:
##  Groups   Name             Std.Dev. Corr
##  Subject  (Intercept)      1.19
##           Instrumentpiano  1.23     -0.21
##           Instrumentstring 1.78     -0.42  0.71
##           HarmonyI-V-IV    0.39      0.46 -0.40 -0.51
```

```
##           HarmonyI-V-VI    1.01     -0.03 -0.27 -0.28 -0.14
##           HarmonyIV-I-V    0.53     -0.31 -0.32 -0.24 -0.40 -0.30
##  Residual                  1.55
## ---
## number of obs: 1904, groups: Subject, 54
## AIC = 7586.8, DIC = 7522.8
## deviance = 7522.8
```

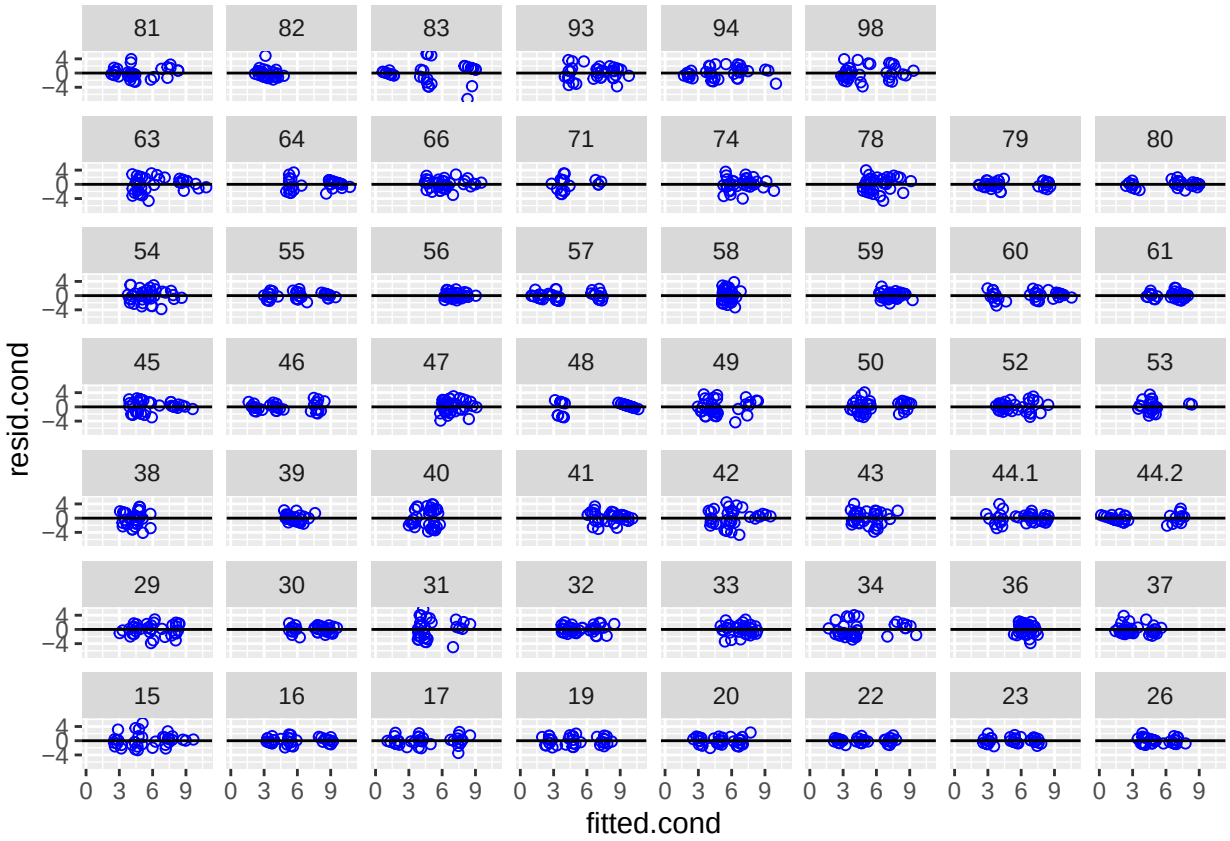
Then I used `fitLMER.fnc` again to check if there is any change in random effects. Note that `fitLMER.fnc` only works for *independent* random effects, such as `(Instrument | Subject)` and `(Harmony | Subject)`. After automated selection, I grouped the selected random effects into one bracket to allow for interactions. It turned out only random effects for harmony and instrument are needed. The final model is given below.

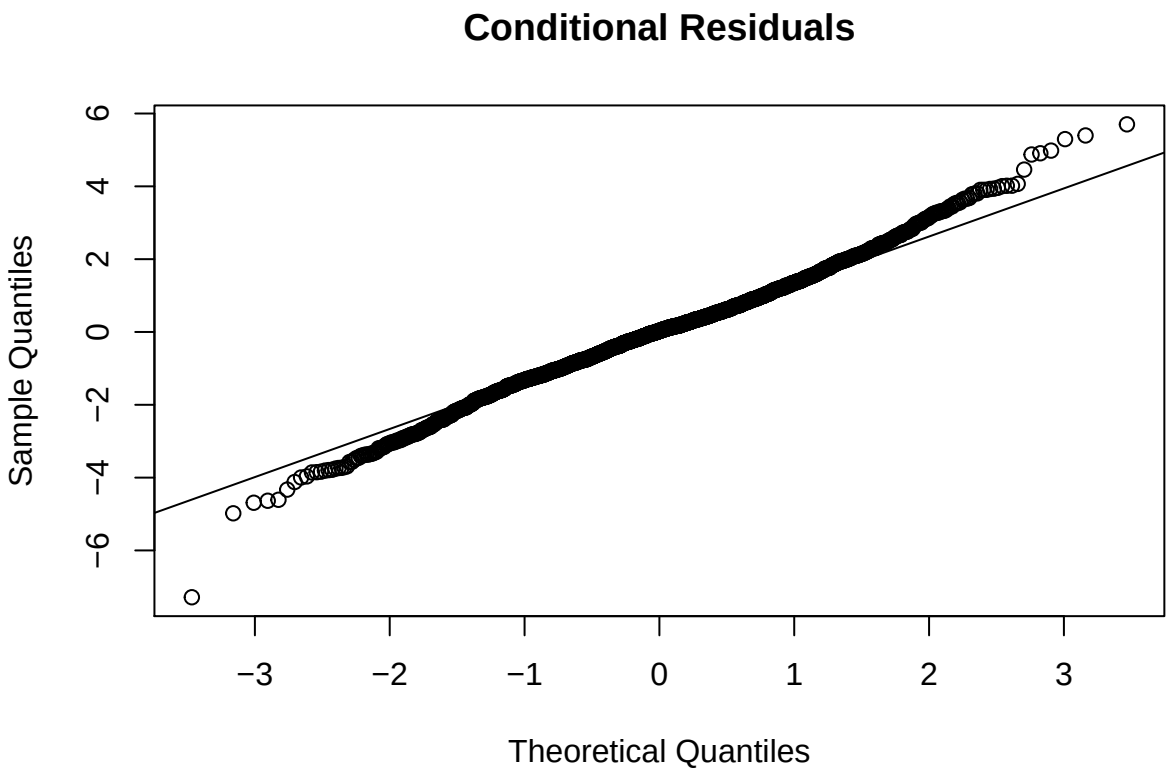
```
## lmer(formula = Popular ~ Instrument + Harmony + Voice + X16.minus.17 +
##     ConsNotes + (1 + Instrument + Harmony | Subject), data = ratings,
##     REML = FALSE)
##                  coef.est coef.se
## (Intercept)       6.25     0.27
## Instrumentpiano  -1.05     0.19
## Instrumentstring -2.78     0.26
## HarmonyI-V-IV    -0.04     0.11
## HarmonyI-V-VI    -0.37     0.17
## HarmonyIV-I-V    -0.30     0.12
## Voicepar3rd       0.22     0.09
## Voicepar5th       0.26     0.09
## X16.minus.17      0.10     0.05
## ConsNotes         0.13     0.07
##
## Error terms:
##  Groups   Name             Std.Dev. Corr
##  Subject  (Intercept)      1.19
##           Instrumentpiano  1.23     -0.21
##           Instrumentstring 1.78     -0.42  0.71
##           HarmonyI-V-IV    0.39      0.46 -0.40 -0.51
##           HarmonyI-V-VI    1.01     -0.03 -0.27 -0.28 -0.14
##           HarmonyIV-I-V    0.53     -0.31 -0.32 -0.24 -0.40 -0.30
##  Residual                  1.55
## ---
## number of obs: 1904, groups: Subject, 54
## AIC = 7586.8, DIC = 7522.8
## deviance = 7522.8
```

Guitar music have higher popular ratings than piano and string quartet. Listener's ability to distinguish classical vs popular music is positively associated with popular ratings, holding all else equal. Listeners who concentrate on notes tend to give higher popular ratings, holding all else equal.

There is personal bias in popular ratings. Perhaps person A was more inclined to rate everything as classical, and person B is more inclined to rate everything as popular. The effect of instrument and harmony on popular ratings differs among individuals.

*Conditional Residuals.* The conditional residuals for the final model is mean zero. The skewness in QQ-plot suggests the residuals are not perfectly normal but not too crazy either. I won't worry too much about it.


81
82
83
93
94
98
63
64
66
71
74
78
79
80
54
55
56
57
58
59
60
61
45
46
47
48
49
50
52
53
38
39
40
41
42
43
44.1
44.2
29
30
31
32
33
34
36
37
15
16
17
19
20
22
23
26
resid.cond
fitted.cond
4
0
−4
0 3 6 9




### A.4 Harmonic motion and Familiarity with Pachelbel Comedy Bits

```
##
## Call:
## lm(formula = Classical ~ Instrument * Harmony * Voice + Harmony *
##     KnowRob + Harmony * KnowAxis, data = ratings)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.2813 -1.6245  0.0641  1.5606  6.2419
##
## Coefficients:
##                                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                               4.04307    0.30917  13.077  < 2e-16
## Instrumentpiano                           1.06100    0.42678   2.486  0.01300
## Instrumentstring                          3.02330    0.43081   7.018 3.15e-12
## HarmonyI-V-IV                            -0.28820    0.43495  -0.663  0.50767
## HarmonyI-V-VI                             1.03011    0.43941   2.344  0.01917
## HarmonyIV-I-V                            -0.33522    0.43712  -0.767  0.44325
## Voicepar3rd                              -0.03138    0.42876  -0.073  0.94167
## Voicepar5th                              -0.40874    0.42876  -0.953  0.34056
## KnowRob                                  -0.05696    0.06617  -0.861  0.38944
## KnowAxis                                  0.16845    0.05480   3.074  0.00214
## Instrumentpiano:HarmonyI-V-IV             0.76968    0.60505   1.272  0.20350
## Instrumentstring:HarmonyI-V-IV            0.80732    0.60788   1.328  0.18431
## Instrumentpiano:HarmonyI-V-VI             0.38874    0.60645   0.641  0.52160
```

```
## Instrumentstring:HarmonyI-V-VI              0.02371    0.61221   0.039  0.96911
## Instrumentpiano:HarmonyIV-I-V               0.64812    0.60355   1.074  0.28303
## Instrumentstring:HarmonyIV-I-V              0.20473    0.60781   0.337  0.73628
## Instrumentpiano:Voicepar3rd                -0.19625    0.60496  -0.324  0.74567
## Instrumentstring:Voicepar3rd               -0.28277    0.60641  -0.466  0.64105
## Instrumentpiano:Voicepar5th                 0.43671    0.60501   0.722  0.47050
## Instrumentstring:Voicepar5th                0.25165    0.60927   0.413  0.67962
## HarmonyI-V-IV:Voicepar3rd                  -0.06958    0.60495  -0.115  0.90844
## HarmonyI-V-VI:Voicepar3rd                  -1.71207    0.60641  -2.823  0.00480
## HarmonyIV-I-V:Voicepar3rd                   0.35213    0.60636   0.581  0.56149
## HarmonyI-V-IV:Voicepar5th                   0.40411    0.60495   0.668  0.50421
## HarmonyI-V-VI:Voicepar5th                  -0.17809    0.60781  -0.293  0.76955
## HarmonyIV-I-V:Voicepar5th                   0.69351    0.60495   1.146  0.25178
## HarmonyI-V-IV:KnowRob                       0.03641    0.09404   0.387  0.69866
## HarmonyI-V-VI:KnowRob                       0.43877    0.09405   4.666 3.30e-06
## HarmonyIV-I-V:KnowRob                       0.04471    0.09355   0.478  0.63273
## HarmonyI-V-IV:KnowAxis                     -0.06252    0.07759  -0.806  0.42048
## HarmonyI-V-VI:KnowAxis                     -0.19720    0.07762  -2.541  0.01115
## HarmonyIV-I-V:KnowAxis                     -0.08291    0.07737  -1.072  0.28406
## Instrumentpiano:HarmonyI-V-IV:Voicepar3rd  -0.66113    0.85562  -0.773  0.43981
## Instrumentstring:HarmonyI-V-IV:Voicepar3rd -0.53730    0.85862  -0.626  0.53154
## Instrumentpiano:HarmonyI-V-VI:Voicepar3rd   1.14151    0.85760   1.331  0.18333
## Instrumentstring:HarmonyI-V-VI:Voicepar3rd  1.76975    0.85969   2.059  0.03967
## Instrumentpiano:HarmonyIV-I-V:Voicepar3rd  -0.08712    0.85658  -0.102  0.91900
## Instrumentstring:HarmonyIV-I-V:Voicepar3rd  0.49596    0.85859   0.578  0.56357
## Instrumentpiano:HarmonyI-V-IV:Voicepar5th  -0.64868    0.85661  -0.757  0.44899
## Instrumentstring:HarmonyI-V-IV:Voicepar5th -1.12553    0.86065  -1.308  0.19111
## Instrumentpiano:HarmonyI-V-VI:Voicepar5th  -0.65471    0.85663  -0.764  0.44479
## Instrumentstring:HarmonyI-V-VI:Voicepar5th -0.35999    0.86372  -0.417  0.67688
## Instrumentpiano:HarmonyIV-I-V:Voicepar5th  -0.81407    0.85358  -0.954  0.34035
## Instrumentstring:HarmonyIV-I-V:Voicepar5th -0.82428    0.85860  -0.960  0.33716
##
## (Intercept)                                ***
## Instrumentpiano                            *
## Instrumentstring                           ***
## HarmonyI-V-IV
## HarmonyI-V-VI                              *
## HarmonyIV-I-V
## Voicepar3rd
## Voicepar5th
## KnowRob
## KnowAxis                                   **
## Instrumentpiano:HarmonyI-V-IV
## Instrumentstring:HarmonyI-V-IV
## Instrumentpiano:HarmonyI-V-VI
## Instrumentstring:HarmonyI-V-VI
## Instrumentpiano:HarmonyIV-I-V
## Instrumentstring:HarmonyIV-I-V
## Instrumentpiano:Voicepar3rd
## Instrumentstring:Voicepar3rd
## Instrumentpiano:Voicepar5th
## Instrumentstring:Voicepar5th
## HarmonyI-V-IV:Voicepar3rd
## HarmonyI-V-VI:Voicepar3rd                  **
```

```
## HarmonyIV-I-V:Voicepar3rd
## HarmonyI-V-IV:Voicepar5th
## HarmonyI-V-VI:Voicepar5th
## HarmonyIV-I-V:Voicepar5th
## HarmonyI-V-IV:KnowRob
## HarmonyI-V-VI:KnowRob                      ***
## HarmonyIV-I-V:KnowRob
## HarmonyI-V-IV:KnowAxis
## HarmonyI-V-VI:KnowAxis                     *
## HarmonyIV-I-V:KnowAxis
## Instrumentpiano:HarmonyI-V-IV:Voicepar3rd
## Instrumentstring:HarmonyI-V-IV:Voicepar3rd
## Instrumentpiano:HarmonyI-V-VI:Voicepar3rd
## Instrumentstring:HarmonyI-V-VI:Voicepar3rd *
## Instrumentpiano:HarmonyIV-I-V:Voicepar3rd
## Instrumentstring:HarmonyIV-I-V:Voicepar3rd
## Instrumentpiano:HarmonyI-V-IV:Voicepar5th
## Instrumentstring:HarmonyI-V-IV:Voicepar5th
## Instrumentpiano:HarmonyI-V-VI:Voicepar5th
## Instrumentstring:HarmonyI-V-VI:Voicepar5th
## Instrumentpiano:HarmonyIV-I-V:Voicepar5th
## Instrumentstring:HarmonyIV-I-V:Voicepar5th
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.207 on 1860 degrees of freedom
## Multiple R-squared:  0.3078, Adjusted R-squared:  0.2918
## F-statistic: 19.24 on 43 and 1860 DF,  p-value: < 2.2e-16
```