Classical vs. Popular: What Factors Influence the Answer?

Richard Shu*
shiyus@andrew.cmu.edu

6 December 2019

Abstract

We aim to find out some possible factors that could make an impact on rating music stimuli as popular or classical. We start with building linear models for three main design variables Instrument, Harmony and Voice, adding random effects to the model by manual subset selection, including other fixed effects through automatic methods and testing for specific considerations. We discover that certain types of Instrument, Harmony or Voice are closely related to Classical music or Popular music. Popular rating is mainly affected by Instrument while classical rating is influenced by more factors, and when respondents give their evaluations, different taste and knowledge in the design variables also influence their decisions.

1 Introduction

Music has been playing an important role since the beginning of civilization and throughout the history various genres of music have been developed. Classical music and Popular music are two typical genres that have attracted a lot of audience. There are definitions from musicians that distinguish between Classical music and Popular music but how do listeners identify the different types of music?

in this paper, we are interested in finding out what factors drive listeners to categorize a piece of music as Classical or Popular. We will use the dataset collected by musicologist and composer Ivan Jimenez and his student Vincent Rossi in 2012 to develop hierchical models that statistically analyze the relationship between genre ratings and other explanatory variables. In that experiment 70 listeners were presented with 36 musical stimuli played with different combinations of instrument, harmonic motion and voice leading, and ratings from respondents along with other information were included in the dataset. We start with preliminary exploratory data analysis, filter and transform variables when necessary, construct fixed linear models and random effect models and give our conclusion. When developing models, we also put interpretability, musical knowledge and research goals into consideration. In particular, we aim to answer the questions below:

- Among the three design factors, does Instrument exert the strongest influence on musical ratings?
- Does the specific Harmonic Motion I-V-vi have a strong association with classical ratings?

*Master of Statistical Practice, Department of Statistics & Data Science, Carnegie Mellon University

- Does contrary motion among the voice leading levels, have a strong association with classical ratings?
- Does it seem to matter whether the respondent is familiar with one or the other of the Pachelbel rants/comedy bits?
- Are there differences in the way that musicians and non-musicians identify classical music?
- Are there differences in the things that drive classical vs. popular ratings?

Finally, we will recap our findings, discuss the big picture and mention any other interesting observations during our analysis in the discussion section.

2 Methods

The data for this study comes from a designed experiment conducted in 2012 aiming to measure the influence of instrument, harmonic motion and voice leading on listeners' identification of music as "classic" or "popular". Each of the 70 listeners gave their ratings to 36 musical stimuli and we have 2520 observations in total. Description of all 26 variables from the dataset is listed below in Table 1. Readers should refer to Professor Brian Junker at Carnegie Mellon University or 36617 course website for detailed information about the data. We use the R language and environment for statistical computing (R Core Team, 2017).

We start our analysis with EDA plots to look at the relationship between musical ratings and explanatory variables. Necessary data cleaning and transformation are performed and will be discussed in the result section.

Next, we focus solely on the three main experimental factors Instrument, Voice and Harmony, and build a linear model of classical rating against the three variables. We start from this model and compare the repeated measures model(random intercept only model) with the best random effect model through subset selection of all possible random effect combinations(function lmer is used here). After we determine what random effects to be included in the model we add more fixed effect terms in and use automatic methods(function fitLMER.fnc is used here) to reach our final model.

We categorize a new variable called ismusician which dichotomizes the Selfdeclare variable into two levels. We want to check if there exist any significant interactions between this term and other variables and if the results are sensitive to where we dichotomize. We are also curious whether familiarity with Pachelbel rants/comedy bits affects slope of harmony, so we build another model to check for that.

Finally, we repeat the process above and come up with a final model for popular rating. We compare coefficients and variables in the two models to see what factors drive classical vs. popular ratings.

Table 1: Table 1: Variable Description

Index	Variable	Description
1	Classical	categorical variable ranging from 1 to 10, 10=very classical sounding
2	Popular	categorical variable ranging from 1 to 10, 10=very popular sounding
3	Subject	unique subject ID
4	Harmony	Harmonic motion(4 levels)
5	Instrument	Instrument(3 levels)
6	Voice	Voice Leading(3 levels)
7	Selfdeclare	Self rating as a musician(1-6, 1 = not at all)
8	OMSI	score on a test of musical knowledge, numeric
9	X16.minus.17	auxiliary measure of musical ability
10	ConsInstr	Concentration devoted to instrument while listening(0-5, 0=not at all
11	ConsNotes	Concentration devoted to notes while listening(0-5, 0=not at all)
12	Instr.minus.Notes	Difference between prev. two variables
13	PachListen	Familiarity with Pachelbel's Canon(0-5, 0=not at all)
14	ClsListen	Familiarity with classical music(0-5, 0=not at all)
15	KnowRob	Familiarity with Pachelbel Rant(0-5, 0=not at all)
16	KnowAxis	Familiaity with Pachelbel comedy bits(0-5, 0=not at all)
17	X1990s2000s	Familiarity with rock and pop music in this decade(0-5, 0=not at all)
18	X1990s2000s.minus.1960s1970s	The difference of the previous measurement in two decades
19	CollegeMusic	Have you taken music classes in college(0=no, 1=yes)
20	NoClass	How many music classes have you taken
21	APTheory	Did you take AP music theory class in High School(0=no,1=yes)
22	Composing	Have you done any music composing(0-5, 0=not at all)
23	PianoPlay	Do you play piano(0-5, 0=not at all)
24	GuitarPlay	Do you play guitar(0-5, 0=not at all)
25	X1stInstr	Proficiency at first musical instrument(0-5, 0=not at all)
26	X2ndInstr	Proficiency at second musical instrument(0-5, 0=not at all)

3 Results

In order to answer the research questions, we are going to select final models for Classical ratings and Popular ratings.

Before we start modelling, we perform EDA and data cleaning to better understand the dataset. Utilizing the skim function in R, we notice that our dependent variables Popular and Classical each contains 27 missing values and some explanatory variables have ineligible amount of NA values. We decide that any column with more than 200 missing values will be removed(except KnowAxis because we will need it in later analysis) because that is around 10 percent of missingness. After deleting columns we delete observations that still have NA values because we can not perform imputation without good understanding of the data and we end up with 1937 rows, which is a good sample size for modelling.

From Figure 1 and 2, we can see that both classical ratings and popular ratings differ between Instrument groups. Similar boxplots are created for Voice and Harmony(please see Appendix

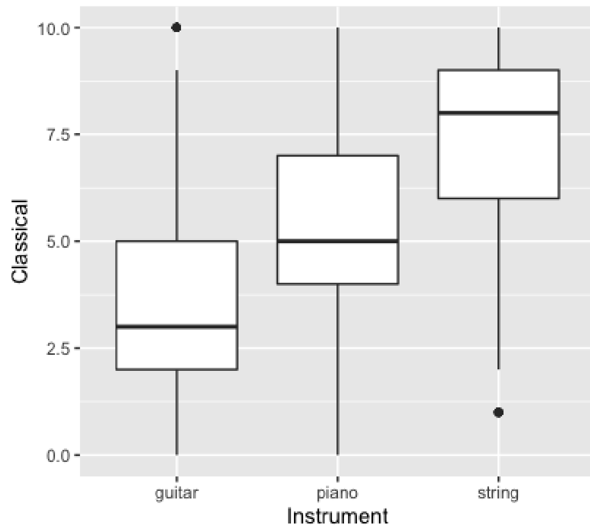


Figure 1: Box plots of classical rating against three instruments

section A) but the difference between groups is not as obvious as what we see in the Instrument variable, which indicates that Instrument might have the greatest influence on ratings. From the skimming result before(Appendix A), we decide to transform three variables OMSI, ConsInstr and Instr.minus.Notes. OMSI is a numeric variable that ranges from 67 to 970 and its scale is quite different from those of other variables, so we standardize it in order to change its scale to be comparative to other variables. ConsInstr and Instr.minus.Notes are supposed to be categorical levels but we observe values such as 4.33, so we round all values to the nearest integer. Most other variables are categorical and we decide not to transform them.

3.1 Building final model for Classical Rating

1. Getting linear model lm1 with all three design variables:

We start constructing the model by determining how the three main experimental variables should be included. We run a linear regression of classical rating against all possible interaction terms between Instrument, Voice and Harmony. This model is put into stepAIC selection and stepBIC selection and we end up with two candidate models. We use an anova test to find the better model and will use it for later parts. The model chosen from AIC is the better model and we have the formula $Y = \beta_0 + \beta_1 I + \beta_2 H + \beta_3 V + \beta_4 H * V$. The normal qq-plot of residuals seem to be satisfactory and from this model we can see that harmonic motion I-V-Vi and contrary voice is positively associated with classical ratings while guitar is negatively associated. We will discuss more about the interpretation later and the summary output of lm1 and residual check can be found in appendix B.

2. Getting random effect model with all three design variables lmer1:

However, we believe that each subject has his or her own taste in music and hearing the same harmonic motion might have different effect on how the classical rating is given, so next we decide to find the best random effect model with the three variables. We use the lmer function to fit a random intercept model first and the exacRLRT test suggests that random effects should be included. Next

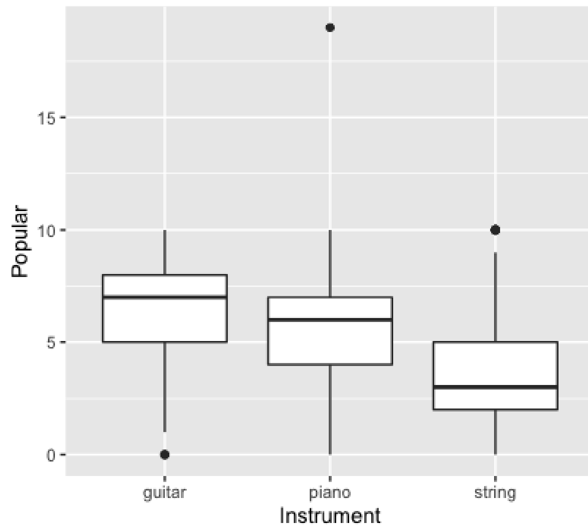


Figure 2: Box plots of popular rating against three instruments

all 14 possible types of random effects with the three variables are tried(note that (I+V—Subject) and (I—Subject)+(V—Subject) are different because we have different assumptions about their correlation). All 15 models are compared in an anova table and we use AIC and BIC values as criterion for choosing the best model. The resulted two models contain random slopes for Instrument and Harmony and the only difference is whether they are correlated. Without any further information we choose to use the model with correlation because that model has more parameters. The model selection procedure and residual check can be found in Appendix B.

Below is the summary of model lmer1 and we will interpret coefficients of the model here. All coefficients of other models can be interpreted in a similar fashion and can be referred back here.

```
   Random effects:
 Groups   Name             Variance Std.Dev. Corr
 Subject  (Intercept)      1.69692  1.3027
          HarmonyI-V-IV    0.09083  0.3014    0.78
          HarmonyI-V-VI    1.82270  1.3501    0.08  0.10
          HarmonyIV-I-V    0.09533  0.3088    0.20  0.19  0.19
          Instrumentpiano  1.60338  1.2662   -0.24 -0.71 -0.35 -0.37
          Instrumentstring 3.41501  1.8480   -0.50 -0.73 -0.60 -0.27  0.59
 Residual                  2.38852  1.5455
Number of obs: 1937, groups:  Subject, 54

Fixed effects:
                          Estimate Std. Error t value
(Intercept)                 3.8322     0.2205  17.377
HarmonyI-V-IV               0.2677     0.1768   1.514
HarmonyI-V-VI               1.2799     0.2519   5.081
HarmonyIV-I-V              -0.1481     0.1768  -0.838
```

```
Instrumentpiano              1.5508     0.1927   8.048
Instrumentstring             3.4526     0.2657  12.992
Voicepar3rd                 -0.2593     0.1717  -1.510
Voicepar5th                 -0.1801     0.1720  -1.047
HarmonyI-V-IV:Voicepar3rd  -0.4158     0.2431  -1.711
HarmonyI-V-VI:Voicepar3rd  -0.6490     0.2434  -2.666
HarmonyIV-I-V:Voicepar3rd   0.5601     0.2431   2.304
HarmonyI-V-IV:Voicepar5th  -0.2485     0.2434  -1.021
HarmonyI-V-VI:Voicepar5th  -0.4733     0.2436  -1.943
HarmonyIV-I-V:Voicepar5th   0.2109     0.2431   0.868
```

- Since all three variables are categorical, we treat the estimated Classical rating 3.83 from a respondent who listens to guitar, contrary voice and I-VI-V as the baseline.
- If we fix the voice at 3rd and keep instrument constant, listening to harmony I-V-IV decreases the classical rating by 0.42-0.27 = 0.15.(Note this is the combined effect of Harmony I-V-IV and the interaction term) But this value may not be significant enough, suggesting there can be no effect.
- If we fix the voice at 3rd and keep instrument constant, listening to harmony I-V-VI increases the classical rating by 1.28-0.65 = 0.63.
- If we fix the voice at 3rd and keep instrument constant, listening to harmony IV-I-V increases the classical rating by 0.56-0.15 = 0.41.
- If we fix the voice at 5th and keep instrument constant, listening to harmony I-V-IV increases the classical rating by 0.17-0.25 = 0.02. This is also not significant enough.
- If we fix the voice at 5th and keep instrument constant, listening to harmony I-V-VI increases the classical rating by 1.28-0.47 = 0.81.
- If we fix the voice at 5th and keep instrument constant, listening to harmony IV-I-V increases the classical rating by 0.21-0.14 = 0.07. This is not significant.
- If we hold voice and harmony constant, listening to piano on average increases the classical rating by 1.55 compared to listening to guitar.
- If we hold voice and harmony constant, listening to string quartet on average increases the classical rating by 3.45 compared to listening to guitar.
- If we fix the harmony at I-V-IV and keep instrument constant, listening to 3rd voice decreases the classical rating by 0.26+0.42 = 0.68.
- If we fix the harmony at I-V-IV and keep instrument constant, listening to 5th voice decreases the classical rating by 0.18+0.25 = 0.43.
- If we fix the harmony at I-V-VI and keep instrument constant, listening to 3rd voice decreases the classical rating by 0.26+0.65 = 0.91.

- If we fix the harmony at I-V-VI and keep instrument constant, listening to 5th voice decreases the classical rating by 0.18+0.48 = 0.66.
- If we fix the harmony at IV-I-V and keep instrument constant, listening to 3rd voice increases the classical rating by 0.56-0.26 = 0.3.
- If we fix the harmony at IV-I-V and keep instrument constant, listening to 5th voice increases the classical rating by 0.21-0.18 = 0.03. Note most of the effects from Voice might not be significant.
- To interpret the random effects, we can imagine that each individual has a distinct taste in music, and the impact of the three experimental variables is the combined effect of the population slope and a random draw from a normal distribution with mean 0 and variance shown in the random effect part. For example, for one subject the effect of piano on his classical rating is $1.5508 + rnorm(0, 1.6, 1)$

3. Adding more fixed effects to the previous model lmer1plus:

As we want to include more fixed effects in our model, we include the fixed effects from the previous model combined with all other variables possible. We do not consider interactions here since we don't think any interaction makes intuitive sense. We use stepAIC and stepBIC to get two different models, and we choose variables that appear in both models as the fixed effects for our lmer model. Our final list of variables will be Harmony, Instrument, Voice, Selfdeclare, Pachlisten, Clslisten, KnowRob, KnowAxis, X1990s2000s.minus.1960s1970s, X1990s2000s, CollegeMusic, Composing, PianoPlay, GuitarPlay and the interaction term between Harmony and Voice. This fixed part is put into the function fitLMER.fnc, which is an automatic method that gives the final mixed model. We select the three random effects as previous models and the result from automatic methods coincide with what we have in the previous model, i.e Instrument and Harmony should have random slopes. We add four more covariates ClsListen, X1990s2000s.minus.1960s1970s, Selfdeclare and Composing. (Please see Appendix C for reference)

4. Testing for some possible interaction terms lmer1final:

In order to answer other research questions we propose, we build two more models and compare with lmer1plus to get our final model.

First, we add terms KnowRob and KnowAxis along with their interactions with Harmony to the previous model and use an anova table to determine whether these variables matter. Both AIC and BIC increase after we include Pachelbel terms so we will stick with our previous model.

Second, we replace the Selfdeclare variable with a dichotomized new variable called ismusician based on Selfdeclare scores and rerun the fixed effects part with all possible interaction terms with ismusician. We try three different cutoff points and fit the variable into the regression. We use stepBIC to select the fixed effect because that gives a relatively simpler model. We add random effects to the selected fix effects. and check for anova result. Setting cutoff at Selfdeclare equals 1 gives us a better model in terms of AIC and BIC and we will present that as the final model.(Please see Appendix D for reference)

5. Answering research questions based on final model lmer1final:

The summary of final model after our analysis in the previous four parts is presented below and we will briefly interpret the results:

- Note that the effects of three main experimental variables do not change much from what we have in part2, so the interpretation will be similar.
- All three levels of Instrument have significant coefficients and their magnitude, compared to Voice and Harmony are the largest. This matches researchers' hypothesis that Instrument exerts the strongest influence among the three design variables.
- Among the four levels, listening to Harmonic motion I-V-VI increases the expected classical rating most and we can say that it has the strongest association with classical ratings.
- KnowAxis and KnowRob do not appear in our final model either as a single variable or as interaction terms, so we don't think that the familiarity with these particular music affects respondents' classical rating on the stimuli music.
- There are differences in the way musicians and non-musicians identify classical music and generally respondents who see themselves as musicians give 0.63 points lower on classical ratings. There is also an interaction effect between difference in frequency of listening to pop/rock music in two decades and the musician variable.
- We want to keep all design variables in our model even if some are not significant. Two observations from the regression output are 1. for the random part, we want all variable variances to be greater than the residual variance but we fail here. Further research might be needed to eventually justify the usage of hierchichal model. 2. for variables with a random effect, we can approximate it as a normal distribution with mean at the fixed estimate and variance from the random model. In this sense we can see that string quartet is somewhat significant from zero while piano is not.
- For other explanatory variables, if respondents listen to pop/rock more in the 90s then in the 60s, the classical rating will likely to increase. The classical rating will generally increase if the respondents listen to Classical music more often.

```
Random effects:
 Groups   Name             Variance Std.Dev. Corr
 Subject  (Intercept)      1.01732  1.0086
          HarmonyI-V-IV    0.08721  0.2953    0.81
          HarmonyI-V-VI    1.81997  1.3491    0.00  0.09
          HarmonyIV-I-V    0.09296  0.3049    0.07  0.16  0.19
          Instrumentpiano  1.60021  1.2650   -0.24 -0.72 -0.35 -0.37
          Instrumentstring 3.41294  1.8474   -0.57 -0.74 -0.60 -0.27  0.59
 Residual                  2.38919  1.5457
Number of obs: 1937, groups:  Subject, 54

Fixed effects:
                                           Estimate Std. Error t value
(Intercept)                                  2.5731     0.8955   2.873
HarmonyI-V-IV                                0.2677     0.1767   1.515
HarmonyI-V-VI                                1.2811     0.2518   5.087
HarmonyIV-I-V                               -0.1481     0.1767  -0.838
```

Instrumentpiano	1.5505	0.1925	8.053
Instrumentstring	3.4523	0.2657	12.994
Voicepar3rd	-0.2593	0.1717	-1.510
Voicepar5th	-0.1801	0.1720	-1.047
X1990s2000s.minus.1960s1970s-3	-0.3592	1.2598	-0.285
X1990s2000s.minus.1960s1970s-2	-2.7849	1.2598	-2.211
X1990s2000s.minus.1960s1970s0	3.1488	1.1429	2.755
X1990s2000s.minus.1960s1970s1	2.7726	1.1799	2.350
X1990s2000s.minus.1960s1970s2	1.2735	0.9532	1.336
X1990s2000s.minus.1960s1970s3	1.5380	0.9019	1.705
X1990s2000s.minus.1960s1970s4	0.6163	1.0972	0.562
X1990s2000s.minus.1960s1970s5	1.1255	1.0222	1.101
ClsListen1	-0.1174	0.3836	-0.306
ClsListen3	1.0386	0.4028	2.578
ClsListen4	0.9352	0.9020	1.037
ClsListen5	2.0785	0.5257	3.954
ismusician1	-0.6256	0.4393	-1.424
X1990s2000s.minus.1960s1970s0:ismusician1	-2.1416	0.9274	-2.309
X1990s2000s.minus.1960s1970s1:ismusician1	-2.1839	1.0417	-2.096
HarmonyI-V-IV:Voicepar3rd	-0.4158	0.2431	-1.711
HarmonyI-V-VI:Voicepar3rd	-0.6502	0.2435	-2.670
HarmonyIV-I-V:Voicepar3rd	0.5601	0.2431	2.304
HarmonyI-V-IV:Voicepar5th	-0.2485	0.2435	-1.021
HarmonyI-V-VI:Voicepar5th	-0.4745	0.2436	-1.948
HarmonyIV-I-V:Voicepar5th	0.2109	0.2431	0.868

3.2 Building final model for Popular Rating

We use similar steps as described in the previous section and we will just talk about the results here. Readers should feel free to refer to Appendix E for technical details.

The final linear model selected from the step-wise function and anova suggests that popular is only related to instrument but to fulfill the purpose of this experiment we still include other two variables in the model. Later in the subset selection we find that Instrument and Harmony should have random effects. Again we want the two random slopes to be correlated since that makes sense musically.

We next add other covariates into the fixed model. With candidates including Selfdeclare,OMSI,ConsInstr,PachI and PianoPlay. Surprisingly the automatic method suggests that only Instrument and its random effect are associated with Popular rating. We last try to include ismusician into the fixed model and we choose dichotomization at Selfdeclare=2 and stepAIC for variable selection because that model keeps all design variables.

The final output is shown below and I will briefly interpret the results and answer research questions:

- Most coefficients other than those of Instrument seem to be not significant both statistically and numerically. This is reasonable because we force those variables in this model. In a way this also reflects that Instrument plays the most important role in affecting popular ratings.

- The numerical values seem to be totally contrary to what we see in the classical model. Guitar and harmony I-IV-V are positively associated with popular ratings while contrary voice is negatively associated with popular ratings. We will talk more about this interesting phenomenon in the discussion section. Being a musician increases the expected rating.

```
Random effects:
 Groups   Name             Variance Std.Dev. Corr
 Subject  (Intercept)      1.5394   1.2407
          Instrumentpiano  1.4481   1.2034   -0.15
          Instrumentstring 2.5280   1.5900   -0.29  0.65
          HarmonyI-V-IV    0.1634   0.4042    0.47 -0.13 -0.23
          HarmonyI-V-VI    0.9491   0.9742   -0.13 -0.23 -0.29 -0.27
          HarmonyIV-I-V    0.3533   0.5944   -0.30 -0.35 -0.37 -0.61 -0.14
 Residual                  2.5244   1.5888
Number of obs: 1937, groups:  Subject, 54

Fixed effects:
                           Estimate Std. Error t value
(Intercept)                 6.70141    0.25275  26.514
HarmonyI-V-IV              -0.02431    0.15046  -0.162
HarmonyI-V-VI              -0.06508    0.21731  -0.299
HarmonyIV-I-V              -0.22222    0.16903  -1.315
Instrumentpiano            -1.31968    0.24170  -5.460
Instrumentstring           -3.31838    0.30359 -10.930
Voicepar3rd                 0.19987    0.08845   2.260
Voicepar5th                 0.20896    0.08845   2.363
ismusician                  0.06901    0.38784   0.178
HarmonyI-V-IV:ismusician    0.02867    0.23617   0.121
HarmonyI-V-VI:ismusician   -0.71283    0.34072  -2.092
HarmonyIV-I-V:ismusician   -0.10071    0.26507  -0.380
Instrumentpiano:ismusician  0.50765    0.37913   1.339
Instrumentstring:ismusician 0.89035    0.47561   1.872
```

In conclusion, Instrument almost solely drives the rating for the stimuli to be Popular but Voice and musician status could also be influential. On the other hand, there are many variables that determine the classical rating. Besides musician status and the three design variables, frequency of listening to pop music and classical music both matters and there is an interactive effect between harmony and voice. In both models we believe that there exists a subject-specific impact from harmony and instrument.

4 Discussion

We will begin our discussion by summarizing findings in the previous section. For both Classical and Popular ratings, Instrument exerts the strongest influence among the three design factors, and for popular ratings, instrument is almost the only factor that makes an impact. Certain types of Harmony, Voice and Instrument are suggested by the model as typical "Classical" or "Popular" component: Harmonic Motion I-V-VI, string quartet and contrary voice leading increase the expected classical rating significantly while Harmonic motion I-VI-V and guitar increase the expected popular rating. This is intuitively valid since many Classical music pieces are played by string instruments while guitar, a relatively new invention, are used intensively in the rise of pop/rock music. We do not have professional knowledge about voice and harmony but we believe musicians like Mr.Jimenez can make sense out of that. For now the explanation will be that popular music and classical music are two so distinct genres that it is reasonable they have unique properties.

In terms of identifying classical music, musicians and non-musicians give different ratings for the same piece of music and their musician status also correlates with how their music-listening patterns affect the outcome. We can suppose that even if a musician and a non-musician have the same X1990s2000s.minus.1960s1970s value, their difference in background musical knowledge drives them to give different ratings.

When we compare models with ismusician variable at different dichotomization point, we notice that different set of variables remain in the final model and the significance changes as we change our dichotomization point. This is an interesting pattern and maybe further steps can consider more about this variable. However, I think using only Selfdeclare to categorize whether a person is musician or not is rather arbitrary because there are many other effects that also reflect the musical ability of the respondent. If we can define ismusician as a combined effect of several variables such as NoClass, Composing and APTheory, we not only shrink the set of parameters we have and could possibly get simpler models, but also define the variable better. A side note here is that OMSI¿500 is a definition for musician but many respondents with score lower than 500 still categorize themselves as musicians.

During the data cleaning process, we drop out variables with more than 200 NAs and then delete rows with missing values. Since we do not know why some answers are missing, we can not tell whether those missing values are missing at random or not, and it might not be wise to delete them directly. Possible imputation such as regress the missing variable on other variables and replace empty entries with the regression estimation. Future researchers should try such methods to see if the model can be improved. Another issue that we are not confident about is that we keep all three major experimental variables in the model regardless of what automatic methods suggest. For the sake of this analysis, we can defend that we want to use the variables collected, but there is possibility that Harmony and Voice have no impact on Popular ratings, and as long as people hear guitar sounds, they vote for popular music. More analyses should be performed to determine if we should include all design variables.

Our last concern is about how the data was collected. Experimenters interviewed only undergraduates from Pittsburgh University and there could be a cluster effect. There could be students not willing to answer the questions and voluntary bias should also be something we worry about.

Admittedly there are drawbacks in our model, but we still believe that our model captures some important information about ratings. The main takeway from this study is that Popular rating is connected to guitar and Harmony I-VI-V and Classical rating is associated with string quartet, Harmony I-V-VI and contrary voice leading. Whether the respondent is a musician will influence the outcome rating and generally there are more factors affecting Classical rating than Popular rating.

References

R Core Team (2017), *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sheather, S.J. (2009), *A Modern Approach to Regression with R*. New York: Springer Science + Business Media LLC.

Appendix

A:Data Cleaning and EDA

Below are the EDA plots I have mentioned and the code used for variable transformation.

```
require(gridExtra)
plot1 = ggplot(rating,aes(x = Harmony, y = Classical)) + geom_boxplot()
plot2 = ggplot(rating,aes(x = Instrument, y = Classical)) + geom_boxplot()
plot3 = ggplot(rating,aes(x = Voice, y = Classical)) + geom_boxplot()
plot4 = ggplot(rating,aes(x = Harmony, y = Popular)) + geom_boxplot()
plot5 = ggplot(rating,aes(x = Instrument, y = Popular)) + geom_boxplot()
plot6 = ggplot(rating,aes(x = Voice, y = Popular)) + geom_boxplot()
grid.arrange(plot1,plot2,plot3,plot4,plot5,plot6, ncol = 2)
OMSI = (OMSI-mean(OMSI))/sd(OMSI)
ConsInstr = round(ConsInstr,digit = 0)
Instr.minus.Notes = round(Instr.minus.Notes, digit = 0)
```

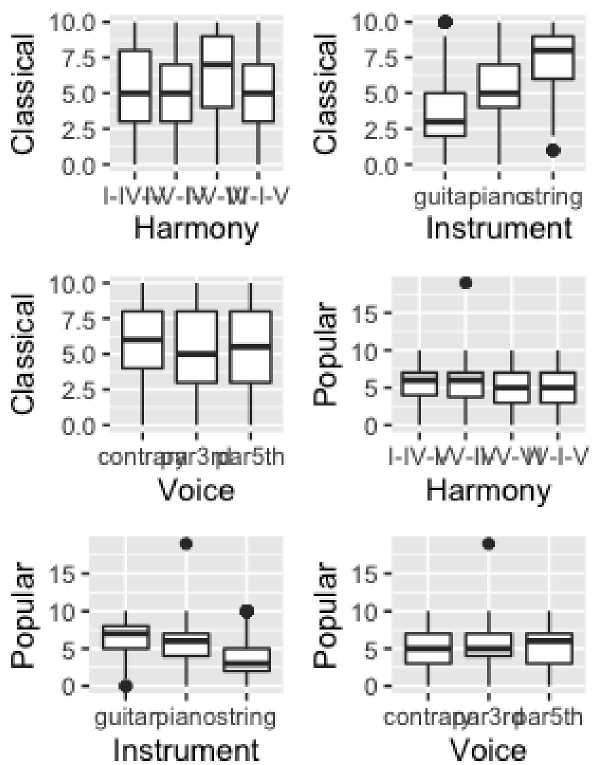
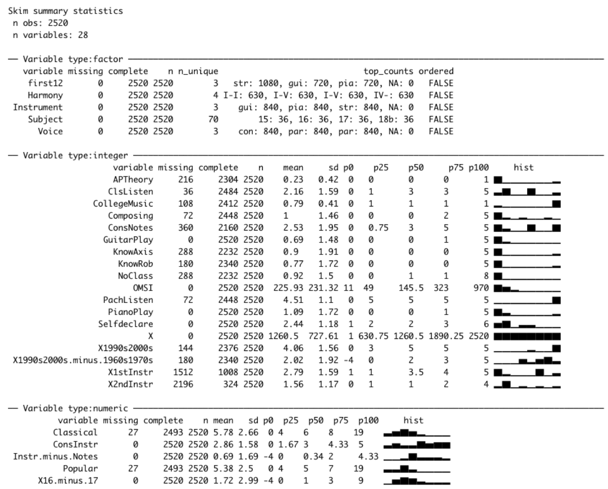


Figure 3: Box plots of Classical/Popular against design variables



Skim summary statistics
n obs: 2520
n variables: 28

— Variable type:factor

variable	missing	complete	n	n_unique	top_counts	ordered
first12	0	2520	2520	3	str: 1080, gui: 720, pia: 720, NA: 0	FALSE
Harmony	0	2520	2520	4	I-I: 630, I-V: 630, I-V: 630, IV-: 630	FALSE
Instrument	0	2520	2520	3	gui: 840, pia: 840, str: 840, NA: 0	FALSE
Subject	0	2520	2520	70	15: 36, 16: 36, 17: 36, 18b: 36	FALSE
Voice	0	2520	2520	3	con: 840, par: 840, par: 840, NA: 0	FALSE

— Variable type:integer

variable	missing	complete	n	mean	sd	p0	p25	p50	p75	p100	hist
APTheory	216	2304	2520	0.23	0.42	0	0	0	0	1	
ClsListen	36	2484	2520	2.16	1.59	0	1	3	3	5	
CollegeMusic	108	2412	2520	0.79	0.41	0	1	1	1	1	
Composing	72	2448	2520	1	1.46	0	0	0	2	5	
ConsNotes	360	2160	2520	2.53	1.95	0	0.75	3	5	5	
GuitarPlay	0	2520	2520	0.69	1.48	0	0	0	1	5	
KnowAxis	288	2232	2520	0.9	1.91	0	0	0	0	5	
KnowRob	180	2340	2520	0.77	1.72	0	0	0	0	5	
NoClass	288	2232	2520	0.92	1.5	0	0	1	1	8	
OMSI	0	2520	2520	225.93	231.32	11	49	145.5	323	970	
PachListen	72	2448	2520	4.51	1.1	0	5	5	5	5	
PianoPlay	0	2520	2520	1.09	1.72	0	0	0	1	5	
Selfdeclare	0	2520	2520	2.44	1.18	1	2	2	3	6	
X	0	2520	2520	1260.5	727.61	1	630.75	1260.5	1890.25	2520	
X1990s2000s	144	2376	2520	4.06	1.56	0	3	5	5	5	
X1990s2000s.minus.1960s1970s	180	2340	2520	2.02	1.92	-4	0	2	3	5	
X1stInstr	1512	1008	2520	2.79	1.59	1	1	3.5	4	5	
X2ndInstr	2196	324	2520	1.56	1.17	0	1	1	2	4	

— Variable type:numeric

variable	missing	complete	n	mean	sd	p0	p25	p50	p75	p100	hist
Classical	27	2493	2520	5.78	2.66	0	4	6	8	19	
ConsInstr	0	2520	2520	2.86	1.58	0	1.67	3	4.33	5	
Instr.minus.Notes	0	2520	2520	0.69	1.69	-4	0	0.34	2	4.33	
Popular	27	2493	2520	5.38	2.5	0	4	5	7	19	
X16.minus.17	0	2520	2520	1.72	2.99	-4	0	1	3	9	

Figure 4: Data summary from skim function

B:Modelling Classical rating against three design variables

Fitting the fixed model of main effects with residual check. As residuals are within the confidence interval and shows a nearly normal distribution, we are satisfied.

```
fit <- lm(Classical~(Instrument+Harmony+Voice)^3)
fit_aic <- stepAIC(fit, direction = "both", k = 2)
summary(fit_aic)

fit_bic <- stepAIC(fit, direction = "both", k = log(nrow(rating)))
summary(fit_bic)

anova(fit_bic, fit_aic)
plot(fit_aic)
```

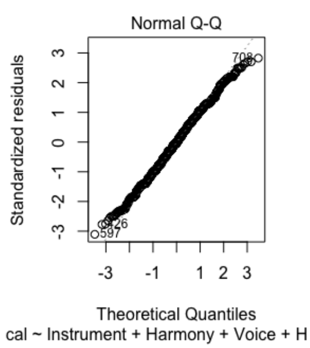


Figure 5: residual plots of the initial linear model lm1(mentioned in section 3.1.1)

Pick the best random effect model:

```
lmer1 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(1|Subject), control = lmerCont
summary(lmer1)
exactRLRT(lmer1)
lmer2 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Instrument|Subject),control = l
lmer3 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony|Subject), control = lme
lmer4 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Voice|Subject),control = lmerCo
lmer5 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Instrument|Subject) + (Voice|Su
lmer6 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Instrument|Subject)+(Harmony|Su
lmer7 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony|Subject)+(Voice|Subject
lmer8 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony|Subject)+(Voice|Subject
lmer9 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony+Voice|Subject),control
lmer10 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Voice+Instrument|Subject),cont
lmer11 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony+Instrument|Subject),co
lmer12 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony+Instrument+Voice|Subje
lmer13 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony|Subject)+(Instrument+V
lmer14 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony+Instrument|Subject)+(V
lmer15 <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+(Harmony+Voice|Subject)+(Instru
anova(lmer1,lmer2,lmer3,lmer4,lmer5,lmer6,lmer7,lmer8,lmer9,lmer10,lmer11,lmer12,lmer13,lmer14
summary(lmer11)
```

Residual checks of the lmer model.

```
xvar_lmer_q2 = predict(lmer11)
yvar_lmer_q2 = resid(lmer11)
binnedplot(xvar_lmer_q2, yvar_lmer_q2)


resid.cond <- r.cond(lmer11)
fit.cond <- yhat.cond(lmer11)
qqnorm(resid.cond,main="Conditional Residuals (epsilon)")
qqline(resid.cond)
```

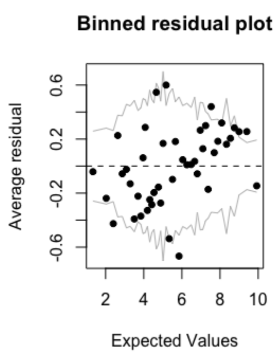


Figure 6: binned residual plot of lmer1(mentioned in section 3.1.2)

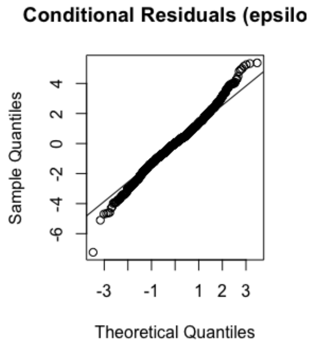


Figure 7: residual check of conditional residuals

C:Adding other fixed effects

I first transform variables that should be factors as factors. I use variables that appear in both stepAIC and stepBIC results to run the lmer model. I use automatic methods to generate a model and force design variables in the model to get the final model of this section. Residual plots are shown in the end. As residuals are within the confidence interval and shows a nearly normal distribution, we are satisfied.

```
rating$ClsListen = as.factor(rating$ClsListen)
rating$CollegeMusic = as.factor(rating$CollegeMusic)
rating$Composing = as.factor(rating$Composing)
rating$GuitarPlay = as.factor(rating$GuitarPlay)
rating$KnowRob = as.factor(rating$KnowRob)
rating$KnowAxis = as.factor(rating$KnowAxis)
rating$PachListen = as.factor(rating$PachListen)
rating$PianoPlay = as.factor(rating$PianoPlay)
rating$Selfdeclare = as.factor(rating$Selfdeclare)
rating$X1990s2000s = as.factor(rating$X1990s2000s)
rating$X1990s2000s.minus.1960s1970s = as.factor(rating$X1990s2000s.minus.1960s1970s)

lm_3 <- lm(Classical ~ .-Popular-Subject+Harmony:Voice, data = rating)
lm_3_fixed_aic <- stepAIC(lm_3, direction = "both", k = 2)
summary(lm_3_fixed_aic)
lm_3_fixed_bic <- stepAIC(lm_3, direction = "both", k = log(nrow(rating)))
summary(lm_3_fixed_bic)
lmer_3_1 <- lmer(Classical~Harmony+Instrument+Voice+Selfdeclare+PachListen+ClsListen+KnowRob+Kn


lmer_q3 <- fitLMER.fnc(lmer_3_1, ran.effects = c("(Harmony|Subject)","(Instrument|Subject)","(V
summary(lmer_q3)
lmer_q3_final <- lmer(Classical~Harmony+Instrument+Voice+Harmony:Voice+ClsListen+X1990s2000s.mi
summary(lmer_q3_final)
```

```
xvar_lmer_q3 = predict(lmer_q3_final)
yvar_lmer_q3 = resid(lmer_q3_final)
binnedplot(xvar_lmer_q3, yvar_lmer_q3)

resid.cond <- r.cond(lmer_q3_final)
fit.cond <- yhat.cond(lmer_q3_final)
qqnorm(resid.cond,main="Conditional Residuals (epsilon)")
qqline(resid.cond)
```

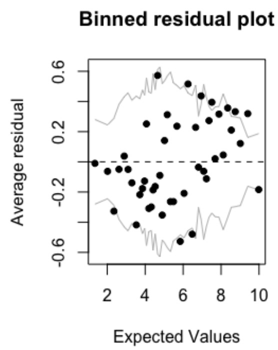


Figure 8: binned residual plot of lmer1plus(mentioned in section 3.1.3)

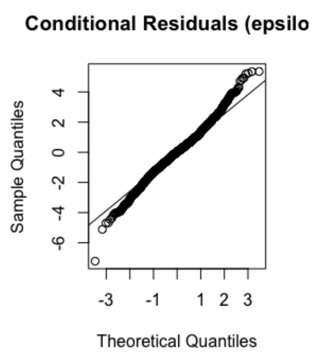


Figure 9: residual check of conditional residuals

D:Final model for classical rating

Some specific effects asked by research questions are tested here to see if they should actually be included here.

```
lmer_q3_pachel <- lmer(Classical~Instrument+Voice+Harmony:Voice+ClsListen+X1990s2000s.minus.196
```

```
anova(lmer_q3_pachel, lmer_q3_final)

rating <- rating %>% mutate(ismusician = ifelse(Selfdeclare %in% c(1,2), 0, 1))

rating <- rating %>% mutate(ismusician1 = ifelse(Selfdeclare %in% 1, 0, 1))

rating <- rating %>% mutate(ismusician2 = ifelse(Selfdeclare %in% c(1,2,3), 0, 1))
```

```
lm_4_1 <- lm(Classical~(Harmony+Voice+Instrument+Composing+X1990s2000s.minus.1960s1970s+ClsList
lm_4_1_aic <- stepAIC(lm_4_1, direction = "both", k = 2)
summary(lm_4_1_aic)
lm_4_1_bic <- stepAIC(lm_4_1, direction = "both", k = log(nrow(rating)))
summary(lm_4_1_bic)
lmer_4_1 <- lmer( Classical ~ Harmony + Instrument + Voice + Composing + X1990s2000s.minus.196
summary(lmer_4_1)
anova(lmer_q3_final, lmer_4_1)


lm_4_2 <- lm(Classical~(Harmony+Voice+Instrument+Composing+X1990s2000s.minus.1960s1970s+ClsList
lm_4_2_aic <- stepAIC(lm_4_2, direction = "both", k = 2)
summary(lm_4_2_aic)
lm_4_2_bic <- stepAIC(lm_4_2, direction = "both", k = log(nrow(rating)))
summary(lm_4_2_bic)
lmer_4_2 <- lmer(Classical ~ Harmony + Instrument + Voice + X1990s2000s.minus.1960s1970s + Cls
    data = rating, control = lmerControl(optimizer = "bobyqa"),REML = FALSE)
summary(lmer_4_2)
anova(lmer_4_2,lmer_q3_final)

lm_4_3 <- lm(Classical~(Harmony+Voice+Instrument+Composing+X1990s2000s.minus.1960s1970s+ClsList
lm_4_3_aic <- stepAIC(lm_4_3, direction = "both", k = 2)
summary(lm_4_3_aic)
lm_4_3_bic <- stepAIC(lm_4_3, direction = "both", k = log(nrow(rating)))
summary(lm_4_3_bic)
lmer_4_3 <- lmer(Classical ~ Harmony + Instrument + Voice + Composing + X1990s2000s.minus.1960s
summary(lmer_4_3)
anova(lmer_4_3,lmer_q3_final)
```

The two anova tables below suggest that including KnowRob and KnowAxis do not improve the model but including ismusician does.

```
    Data: rating
Models:
lmer_q3_final: Classical ~ Harmony + Instrument + Voice + Harmony:Voice + ClsListen +
```

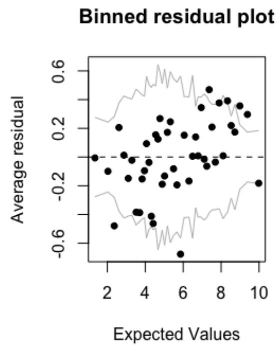


Figure 10: binned residual plot of lmer1final(mentioned in section 3.1.4)

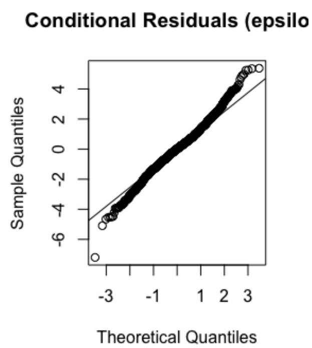


Figure 11: residual check of conditional residuals

```
lmer_q3_final:     X1990s2000s.minus.1960s1970s + Selfdeclare + Composing +
lmer_q3_final:     (Instrument + Harmony | Subject)
lmer_q3_pachel: Classical ~ Instrument + Voice + Harmony:Voice + ClsListen +
lmer_q3_pachel:     X1990s2000s.minus.1960s1970s + Selfdeclare + Composing +
lmer_q3_pachel:     KnowAxis * Harmony + KnowAxis * Harmony + (Instrument + Harmony |
lmer_q3_pachel:     Subject)
               Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer_q3_final  57 7710.9 8028.3 -3798.4   7596.9
lmer_q3_pachel 65 7720.4 8082.4 -3795.2   7590.4 6.4466      8     0.5973



Data: rating
Models:
lmer_4_2: Classical ~ Harmony + Instrument + Voice + X1990s2000s.minus.1960s1970s +
lmer_4_2:     ClsListen + ismusician1 + X1990s2000s.minus.1960s1970s:ismusician1 +
lmer_4_2:     Harmony:Voice + (Harmony + Instrument | Subject)
```

```
lmer_q3_final: Classical ~ Harmony + Instrument + Voice + Harmony:Voice + ClsListen +
lmer_q3_final:     X1990s2000s.minus.1960s1970s + Selfdeclare + Composing +
lmer_q3_final:     (Instrument + Harmony | Subject)
              Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer_4_2      51 7706.4 7990.4 -3802.2   7604.4
lmer_q3_final 57 7710.9 8028.3 -3798.4   7596.9 7.5068      6     0.2765
```

E:Finding final model for popular rating

We basically repeat steps in the previous parts. Will only show code here and will not present output.

```
fit_pop <- lm(Popular~(Instrument+Harmony+Voice)^3,data = rating)
fit_pop_aic <- stepAIC(fit_pop, direction = "both", k = 2)
summary(fit_pop_aic)

fit_pop_bic <- stepAIC(fit_pop, direction = "both", k = log(nrow(rating)))
summary(fit_pop_bic)

anova(fit_pop_aic, fit_pop_bic)


lmer1_pop <- lmer(Popular~Harmony+Instrument+Voice+(1|Subject),data = rating, control = lmerCo
summary(lmer1_pop)

exactRLRT(lmer1_pop)

lmer2_pop <- lmer(Popular~Harmony+Instrument+Voice+(Instrument|Subject),data = rating,control =
lmer3_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony|Subject),data = rating,control = l
lmer4_pop <- lmer(Popular~Harmony+Instrument+Voice+(Voice|Subject),data = rating,control = lme
lmer5_pop <- lmer(Popular~Harmony+Instrument+Voice+(Instrument|Subject) + (Voice|Subject),data
lmer6_pop <- lmer(Popular~Harmony+Instrument+Voice+(Instrument|Subject)+(Harmony|Subject),data
lmer7_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony|Subject)+(Voice|Subject),data = ra
lmer8_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony|Subject)+(Voice|Subject)+(Instrume
lmer9_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony+Voice|Subject), data = rating,cont
lmer10_pop <- lmer(Popular~Harmony+Instrument+Voice+(Voice+Instrument|Subject), data = rating,
lmer11_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony+Instrument|Subject),data = rating
lmer12_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony+Instrument+Voice|Subject),data =
lmer13_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony|Subject)+(Instrument+Voice|Subjec
lmer14_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony+Instrument|Subject)+(Voice|Subjec
lmer15_pop <- lmer(Popular~Harmony+Instrument+Voice+(Harmony+Voice|Subject)+(Instrument|Subjec
anova(lmer1_pop,lmer2_pop,lmer3_pop,lmer4_pop,lmer5_pop,lmer6_pop,lmer7_pop,lmer10_pop,lmer11_

lmer.pop <- fitLMER.fnc(lmer1_pop, ran.effects=c("(Harmony|Subject)", "(Instrument|Subject)","
summary(lmer.pop)
summary(lmer11_pop)
```

```
#Might use lmer6_pop here
xvar_lmer_q5a = predict(lmer6_pop)
yvar_lmer_q5a = resid(lmer6_pop)
binnedplot(xvar_lmer_q5a, yvar_lmer_q5a)

resid.cond <- r.cond(lmer6_pop)
fit.cond <- yhat.cond(lmer6_pop)
qqnorm(resid.cond,main="Conditional Residuals (epsilon)")
qqline(resid.cond)

lm_5 <- lm(Popular ~ .-Classical-Subject, data = rating)
lm_5_fixed_aic <- stepAIC(lm_5, direction = "both", k = 2)
summary(lm_5_fixed_aic)#Harmony + Instrument + Voice + Selfdeclare + OMSI + X16.minus.17 + Cons
lm_5_fixed_bic <- stepAIC(lm_5, direction = "both", k = log(nrow(rating)))
summary(lm_5_fixed_bic)#Instrument + Selfdeclare + OMSI + ConsInstr + PachListen + ClsListen +
lmer_5_1 <- lmer(Popular~Harmony+Instrument+Voice+Selfdeclare+OMSI+ConsInstr+PachListen+ClsLis


lmer_q5 <- fitLMER.fnc(lmer_5_1, ran.effects = c("(Harmony|Subject)","(Instrument|Subject)","(
summary(lmer_q5)

xvar_lmer_q5b = predict(lmer_q5)
yvar_lmer_q5b = resid(lmer_q5)
binnedplot(xvar_lmer_q5b, yvar_lmer_q5b)

resid.cond <- r.cond(lmer_q5)
fit.cond <- yhat.cond(lmer_q5)
qqnorm(resid.cond,main="Conditional Residuals (epsilon)")
qqline(resid.cond)

lm_5_1 <- lm(Popular~(Harmony+Voice+Instrument)*ismusician, data = rating)
lm_5_1_aic <- stepAIC(lm_5_1, direction = "both", k = 2)
summary(lm_5_1_aic)
lm_5_1_bic <- stepAIC(lm_5_1, direction = "both", k = log(nrow(rating)))
summary(lm_5_1_bic)
lmer_5_1<-lmer(Popular ~ Harmony + Instrument +Voice+ ismusician + Harmony:ismusician +Instrum
summary(lmer_5_1)

lm_5_2 <- lm(Popular~(Harmony+Voice+Instrument)*ismusician1, data = rating)
lm_5_2_aic <- stepAIC(lm_5_2, direction = "both", k = 2)
summary(lm_5_2_aic)
lm_5_2_bic <- stepAIC(lm_5_2, direction = "both", k = log(nrow(rating)))
summary(lm_5_2_bic)
lmer_5_2<-lmer(Popular ~ Harmony + Instrument + +Voice+ismusician1 + Harmony:ismusician1+(Inst
summary(lmer_5_2)
```

```
lm_5_3 <- lm(Popular~(Harmony+Voice+Instrument)*ismusician2, data = rating)
lm_5_3_aic <- stepAIC(lm_5_3, direction = "both", k = 2)
summary(lm_5_3_aic)
lm_5_3_bic <- stepAIC(lm_5_3, direction = "both", k = log(nrow(rating)))
summary(lm_5_3_bic)
lmer_5_3 <- lmer(Popular ~ Harmony + Instrument +Voice+ ismusician2 + Harmony:ismusician2+Inst
summary(lmer_5_3)
```