How do we Perceive Music?

Julia Stelman
jstelman@andrew.cmu.edu

December 9, 2019

1 Abstract

In this paper, I look at what factors relate to how listeners perceive how classical or popular a song is. The effects of Instrument, Harmonic Motion, and Voice Leading are looked at through the lens of mixed effects models, correlation statistics, and boxplots. Using mixed model regression analysis, I examine the connections between ratings and combinations of these factors and traits specific to each research subject. My final models shows that the strongest linear predictor of both classical and popular ratings is instrument, but that instrument has a different association with popular ratings than with classical ratings. Additionally, we see that listeners give different rankings if they are musicians, or if they are familiar with Rob Paravonian's Pachelbel Rant on Youtube. Finally, I discuss possible explanations for my findings.

2 Introduction

In 2012, Ivan Jimenez, a composer and musicologist visiting the University of Pittsburgh[1], and student Vincent Rossi, collected 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". A track's resemblance to a genre (Classical or Popular), as perceived by each research subject, was collected as a 1-10 rating reported by the subject. Researchers were interested in seeing how genre resemblance, as perceived by subjects, relates not only to musical traits of the track, but further more, how that is influenced, if at all, by characteristics of the individual research subjects. In this paper, I will be investigating the following questions through my analysis.

1. What experimental factor or combination of factors (**Instrument, Harmony,** and **Voice**) has the strongest influence on how Classical or Popular a subject perceives a given track?
   - What insight does the analysis elicit into the following questions?
     - Does **Instrument** exert the strongest influence with respect to tracks' perceived resemblance to each genre?
     - Among the levels of **Harmonic Motion**:
       * Does *I-V-VI* have the strongest association with perceived resemblance to Classical music (perceived Classical-ness)?

[1] Now at the Sibelius Institute, University of the Arts, Helsinki Finland.

    * Does a subject's familiarity with either of the Pachelbel rants/comedy bits seem to interact with the influence of **harmony** on his or her ratings of perceived Classical-ness?
  - Among the levels of **Voice Leading**, does *contrary motion* have the strongest association with perceived Classical-ness?

2. Are there differences in the ways that musicians and non-musicians perceive classical music?

3. Are there differences in the variables that drive ratings of perceived Classical-ness vs perceived Popular-ness?

3 Methods

The researchers 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

The subjects also answered several survey-style questions and completed a test of musical knowledge. The survey questions were posed as a mix of *yes/no* questions and *scale of A to B* questions. A few of the questions posed in the latter manner were ones I felt ought to have been asked in the form of *yes/no* questions. When it comes to self-reported *scale of A to B*-style answers on surveys, there's always a need for caution. It's likely that each person has their own idea of what differentiates a 1 from a 2 or a 3. For reliability, I chose to take advantage of the idea that most everyone has the same standards for what warrants an answer of A. By turning zeros and non-zeros into *true*s and *false*s respectively, I transformed eight of the *scale of A to B*-style variables into *yes/no*-style variables. For further details, refer to Subsection **??** on page **??** of the Appendix.

Other steps taken to care for missing and unclean data are listed below:

- Two original predictor variables were completely dropped to a high proportion of missing values.
- All observations with missing response variables (i.e. Classical or Popular) were removed.
- There where two subjects for which at least 10 of the values recorded for the response variables were

- Three unaffiliated observations had impossible values for the response variables, assumed to be typos. I used by best judgement to replace them with the correct values.

- For all other variables with missing data, imputation was performed on a case by case basis.

For exact details, refer to Subsection **??** on page **??** of the Appendix for exact details. In the course of answering the three research questions, I used a combination of logic, research, and statistical analysis using R software and packages (R Core Team). The variables analyzed in this study are summarized in Figure 1 below. Of the 24 different variables described, two are response variables (Classical and Popular), three are main experimental factors (Instrument, Harmony, and Voice), and the remaining 19 represent some form of all but three of the other original variables from the original data set.

3.1 What experimental factor or combination of factors (Instrument, Harmony, and Voice) has the strongest influence on how Classical or Popular a subject perceives a given track?

For early EDA, a mix of correlation plots, diagnostic residual plots, boxplots, conventional linear models, and variable selection plots (using BIC) were used. Later computational variable selection methods from R's `LMERConvenienceFunctions` library's `fitLMER.fnc()` series were used in combination with personal judgements based on understanding of the subject matter to perform backwards elimination of fixed effects (experimental factors and individual covariates) and forwards selection of random effects (three main experimental factors).

3.2 Are there differences in the ways that musicians and non-musicians perceive classical music?

Is.musician was a variable fabricated by transforming the *scale of 1-6* variable `Selfdeclare` by dichotomizing along the minimum. The effect this variable had as a predictor in the Classical ratings model was investigated through performing anova F tests in R.

3.3 Are there differences in the variables that drive ratings of perceived Classical-ness vs perceived Popular-ness?

For this, I used the same methods to model popular that I used to model classical. I then compared the most prevalent factors in the classical model to those in the popular model to see what differences there were in the variables correlated with the how subjects perceive each type of music.

4 Results

Once data cleaning was finished, I began an extensive process of exploratory data analysis. Full details on this are available in the attached Appendix, and I will indicate what parts of it to refer to for more detail about specific procedures as they come up in this paper.

First, I modeled the influence of the *Instrument*, *Harmony*, and *Voice* variables on *Classical*. I began with a conventional linear model including all possible main and interaction effects, and used R's `arm` library's `stepAIC()` function to pick away the ones that weren't needed. (To see details, refer to section A.2.1 of the Appendix.)

A brief description of all variables in the data set follows:

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

Figure 1: A brief description of all variables in the clean data.

I used R's `lme4` library's `lmer()` function to put a subject-based random intercept in the model, and with help from R's `LMERConvenienceFunctions` library's `fitLMER.fnc()` series, I performed backwards elimination of fixed effects and forwards selection of random effects for the three main experimental factors (*Instrument*, *Harmony*, and *Voice*) in the model. (To see details, refer to section A.2.2 of the Appendix. To see conditional residual plots for this model, see section A.2.3.)

The the model that resulted is denoted in Figure 2:

Roughly the same process was performed for Popular ratings. (Gory details available in sections A.3.1 and A.3.2 of the Appendix.) Figure 3 details the result. There was nothing restricting it using the same structure as the Classical ratings model, but it just happened that it did.

I will expand on both of these more in subsection 5.1 of Results.

The analysis is not be complete without consideration of all available information. There exist at least 19 subject-specific variables in the data set that have not been explored in either of the models so far discussed so far. To begin exploring these variables, they were each plotted against `Classical`. Those whose plots seemed to indicate some sort of a relationship with `Classical`, as well as those in which the experimenters took special interest, were placed into models on which several types of automated variable selection were performed. (For more detail and plots, refer to section A.2.4 of the Appendix.) The same was done later for `Popular` (section A.3.3 of the Appendix).

Ultimately, it was a combination of heuristic-based variable selection (mostly AIC and BIC), and informed personal judgement-based variable selection that led to the final models for Classical Ratings detailed in Figure 4, and Popular Ratings detailed in Figure 5, each in hierarchical format. (More information in sections A.2.5 and A.3.4 of the Appendix.)

Interestingly, in the presence of everything, it turned out that there were some personal biases represented in the models after all. They are figured in as random effects in the Classical and Popular ratings models. The existence of these makes the interpretation of the coefficients two-fold. In addition to interpreting the average effect, we have a metric for how much the effect varies from one subject to another. The effects in the models can be interpreted as follows:

Interpretation:

For classical:

For Popluar:

4.1 What experimental factor (combination) has the strongest influence on how Classical or Popular a subject perceives a given track?

As far as single experimental effects, for both Classical and Popular ratings, the latter especially, it seems that, between the three main experimental factors, `instrument` had the biggest influence on subject's perception of the music genre. The coefficients of each genre's final model are most significant for `Instrument`.

The boxplots in Figure 6 provide a give some visual insight into the relationships between each main experimental effect and `Classical`. There is a very clear difference between the interquartile ranges and medians of each level of `Instrument`, which isn't something that can be said for either of the other two variables. For `Harmony`, *I-V-VI* tends to associate with higher Classical ratings more than other levels of `Harmony`, and for `Voice`, *contrary* might associate with lower Classical ratings less often than other levels of `Voice`. But aside from that, there's really no competition.

For Popular ratings, the difference is even more pronounced. A look at the boxplots in Figure 7 provides some visual insight. Like there was with `Classical`, there is a very clear difference between

$$Classical_i = \alpha_{0i} + \alpha_{1i} + \alpha_{2i} + \vec{\alpha_3}^T \vec{1_{Vi}} + \epsilon_i, \epsilon_i \text{-} N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j} \text{-} N(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \vec{\beta_1}^T \vec{1_{Ii}} + \vec{\eta_{1j[i]}}^T \vec{1_{Ii}}, \vec{\eta_{1j}} \text{-} N(\vec{0}, \vec{\tau_1^2})$$

$$\alpha_{2j[i]} = \vec{\beta_2}^T \vec{1_{Hi}} + \vec{\eta_{2j[i]}}^T \vec{1_{Hi}}, \vec{\eta_{2j}} \text{-} N(\vec{0}, \vec{\tau_2^2})$$

	factorLevel	Estimate	Std. Error
beta0	(Intercept)	3.9506	0.2106
beta1[1]	Instrumentpiano	1.3616	0.1723
beta1[2]	Instrumentstring	3.1107	0.2360
beta2[1]	HarmonyI-V-IV	-0.0317	0.0915
beta2[2]	HarmonyI-V-VI	0.7717	0.1758
beta2[3]	HarmonyIV-I-V	0.0383	0.0882
alpha3[1]	Voicepar5th	0.0373	0.0760
alpha3[2]	Voicecontrary	0.3950	0.0760

	factorLevel	Estimate
sigma_squared	Residual	2.3984
tau0_squared	(Intercept)	2.5673
tau1_squared[1]	Instrumentpiano	1.6657
tau1_squared[2]	Instrumentstring	3.4958
tau2_squared[1]	HarmonyI-V-IV	0.0476
tau2_squared[2]	HarmonyI-V-VI	1.6242
tau2_squared[3]	HarmonyIV-I-V	0.0063

Figure 2: Personal Biases Model for Classical: Only main experimental factors considered.
Note: $\vec{1_{\square i}}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

$$Popular_i = \alpha_{0i} + \alpha_{1i} + \alpha_{2i} + \vec{\alpha_3}^T \vec{1_{Vi}} + \epsilon_i, \epsilon_i \text{-} N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j} \text{-} N(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \vec{\beta_1}^T \vec{1_{Ii}} + \vec{\eta_{1j[i]}}^T \vec{1_{Ii}}, \vec{\eta_{1j}} \text{-} N(\vec{0}, \vec{\tau_1^2})$$

$$\alpha_{2j[i]} = \vec{\beta_2}^T \vec{1_{Hi}} + \vec{\eta_{2j[i]}}^T \vec{1_{Hi}}, \vec{\eta_{2j}} \text{-} N(\vec{0}, \vec{\tau_2^2})$$

	factorLevel	Estimate	Std. Error
beta0	(Intercept)	6.1542	0.2525
beta1[1]	Instrumentpiano	-0.8882	0.1617
beta1[2]	Instrumentstring	-2.4092	0.2405
beta2[1]	HarmonyI-V-IV	-0.0329	0.0944
beta2[2]	HarmonyI-V-VI	-0.2036	0.1368
beta2[3]	HarmonyIV-I-V	-0.1720	0.1053
alpha3[1]	Voicepar3rd	0.1348	0.0736
alpha3[2]	Voicepar5th	0.1618	0.0736

	factorLevel	Estimate
sigma_squared	Residual	2.2499
tau0_squared	(Intercept)	3.9574
tau1_squared[1]	Instrumentpiano	1.4450
tau1_squared[2]	Instrumentstring	3.6698
tau2_squared[1]	HarmonyI-V-IV	0.1185
tau2_squared[2]	HarmonyI-V-VI	0.8037
tau2_squared[3]	HarmonyIV-I-V	0.2709

Figure 3: Personal Biases Model for Popular: Only main experimental factors considered.
Note: $\vec{1_{\square i}}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

$$Classical_i = \alpha_{0i} + \alpha_{1i} + \alpha_{2i} + \vec{\alpha_3}^T \vec{1_{Vi}} + \alpha_4 M_i + \alpha_5 K_i + \vec{\alpha_6}^T \vec{1_{HMi}} + \vec{\alpha_7}^T \vec{1_{HKi}} + \epsilon_i, \epsilon_i\text{-}N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j}\text{-}N(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \vec{\beta_1}^T \vec{1_{Ii}} + \vec{\eta_{1j[i]}}^T \vec{1_{Ii}}, \vec{\eta_{1j}}\text{-}N(\vec{0}, \vec{\tau_1^2})$$

$$\alpha_{2j[i]} = \vec{\beta_2}^T \vec{1_{Hi}} + \vec{\eta_{2j[i]}}^T \vec{1_{Hi}}, \vec{\eta_{2j}}\text{-}N(\vec{0}, \vec{\tau_2^2})$$

	factorLevel	Estimate	Std. Error
beta0	(Intercept)	5.1301	0.3657
beta1[1]	Instrumentpiano	1.3605	0.1724
beta1[2]	Instrumentstring	3.1103	0.2361
beta2[1]	Voicepar3rd	-0.3948	0.0760
beta2[2]	Voicepar5th	-0.3574	0.0760
beta2[3]	HarmonyI-V-IV	-0.1784	0.1930
alpha3[1]	HarmonyI-V-VI	-0.4721	0.3048
alpha3[2]	HarmonyIV-I-V	-0.1481	0.1866
alpha4	is.musician	-0.9851	0.4007
alpha5	KnowRob	-0.1194	0.4204
alpha6[1]	HarmonyI-V-IV:is.musician	0.1927	0.2217
alpha6[2]	HarmonyI-V-VI:is.musician	1.3681	0.3433
alpha6[3]	HarmonyIV-I-V:is.musician	0.2363	0.2151
alpha7[1]	HarmonyI-V-IV:KnowRob	-0.0113	0.2320
alpha7[2]	HarmonyI-V-VI:KnowRob	0.9350	0.3599
alpha7[3]	HarmonyIV-I-V:KnowRob	0.0193	0.2251

	factorLevel	Estimate
sigma_squared	Residual	2.3978
tau0_squared	(Intercept)	2.3172
tau1_squared[1]	Instrumentpiano	1.6682
tau1_squared[2]	Instrumentstring	3.4981
tau2_squared[1]	HarmonyI-V-IV	0.0629
tau2_squared[2]	HarmonyI-V-VI	1.1784
tau2_squared[3]	HarmonyIV-I-V	0.0127

Figure 4: Final Model for Classical.
Note: $\vec{1_{\square i}}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

$$Popular_i = \alpha_{0i} + \alpha_{1i} + \alpha_{2i} + \vec{\alpha_3}^T \vec{1}_{Vi} + \alpha_4 X_i + \alpha_5 K_i + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$$
$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j} \sim N(0, \tau_0^2)$$
$$\alpha_{1j[i]} = \vec{\beta_1}^T \vec{1}_{Ii} + \vec{\eta_{1j[i]}}^T \vec{1}_{Ii}, \vec{\eta_{1j}} \sim N(\vec{0}, \vec{\tau_1^2})$$
$$\alpha_{2j[i]} = \vec{\beta_2}^T \vec{1}_{Hi} + \vec{\eta_{2j[i]}}^T \vec{1}_{Hi}, \vec{\eta_{2j}} \sim N(\vec{0}, \vec{\tau_2^2})$$

	factorLevel	Estimate	Std. Error
beta0	(Intercept)	5.6467	0.2870
beta1[1]	Instrumentpiano	-0.8885	0.1617
beta1[2]	Instrumentstring	-2.4095	0.2404
beta2[1]	HarmonyI-V-IV	-0.0330	0.0945
beta2[2]	HarmonyI-V-VI	-0.2037	0.1368
beta2[3]	HarmonyIV-I-V	-0.1719	0.1053
alpha3[1]	Voicepar3rd	0.1351	0.0736
alpha3[2]	Voicepar5th	0.1618	0.0736
alpha4	X16.minus.17	0.1342	0.0624
alpha5	KnowRob	1.3832	0.4649

	factorLevel	Estimate
sigma_squared	Residual	2.2500
tau0_squared	(Intercept)	3.8258
tau1_squared[1]	Instrumentpiano	1.4444
tau1_squared[2]	Instrumentstring	3.6687
tau2_squared[1]	HarmonyI-V-IV	0.1189
tau2_squared[2]	HarmonyI-V-VI	0.8038
tau2_squared[3]	HarmonyIV-I-V	0.2710

Figure 5: Final Model for Popular.
Note: $\vec{1_{\square i}}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

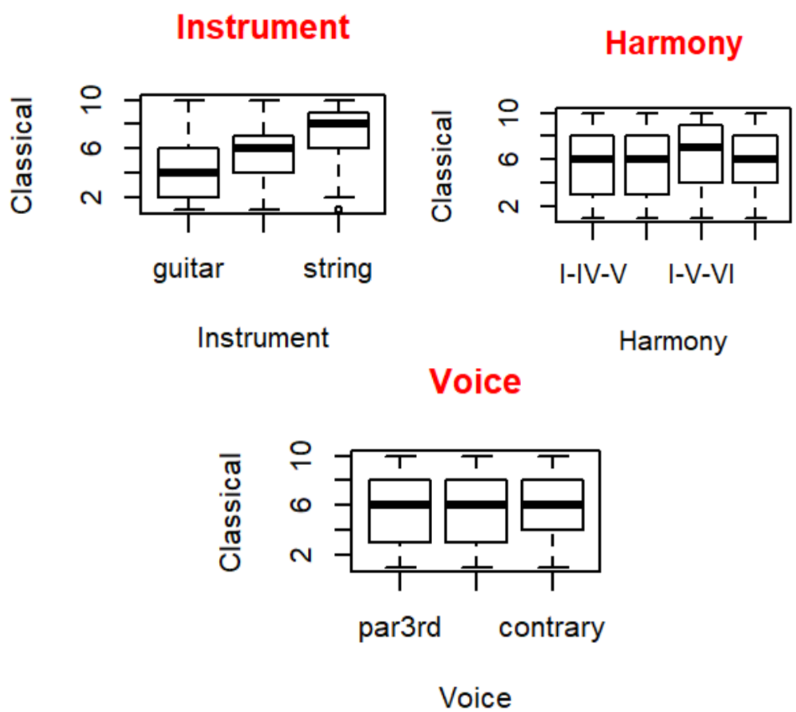


Figure 6: Boxplots of the `Classical` vs each main experimental factor

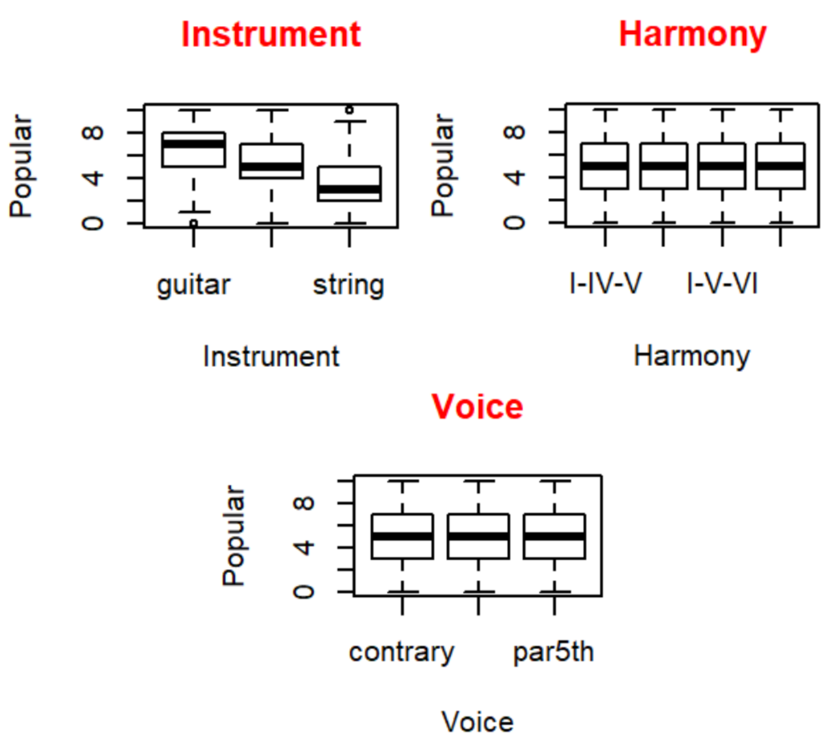


Figure 7: Boxplots of the `Classical` vs each main experimental factor

the interquartile ranges and medians of each level of `Instrument`. For `Popular`, there seems to be the opposite trend between levels of `Instrument` as the ones we saw with it and `Classical`. The interquartile ranges and medians Popular ratings don't seem to differ at all for different levels of `Harmony` and `Voice`.

Speaking on combinations of main experimental effects, the chosen combinations for each genre's final model is three for three. However, `Voice` failed to be selected at least once per genre in automated variable selection methods for both models, and it was actually *never* selected computationally for the Popular ratings model. It's not hard to see why by the large size of it's coefficients' p-values. Furthermore, if we look at the actual coefficients' values, standard errors, and p-values for `Harmony` in both models, we'll see it's heavily reliant on interactions with other variables to keep its spot in both final models. For that reason, it's probably more honest to consult the main experimental factors only-models to answer this question.

Referring back to the Classical ratings model in Figure 2, between `Instrument`, `Harmony`, and `Voice`, the fitted Classical ratings model favored the inclusion of all three as fixed effects and two as random effects. The (modified) fitted Popular ratings model in Figure 3 only includes `Harmony` and `Voice` because I forced it to. As is shown in Figure (the one with just 3MEF for popular), none of the levels of `Voice` nor `Harmony` are significant. This is interesting because the variable selection process favored the omission of `Harmony`, but, under the condition that `Harmony` had to be in the model, it preferred the presence of the random effect to the presence of just the fixed effect. This means that the relationship `Harmony` has with subject's perception of the how popular a track sounds has more to do with the subject's own biases toward certain harmonic motions than a consistent pattern in which subjects across the board find that certain harmonic motions are more resembling of popular music than others.

All in all, compared to those of voice leading or harmonic motion, the differences between the instrument playing the music had more meaningful and cohesive effects on the differences between how classical or popular the subjects perceived the track. Additionally, there were still significant `Harmony`- and `Voice`-dependent differences in how classical the subjects perceived tracks, but to a lesser level, the weights of which were less proportionally distributed across levels in comparison to Instrument. There were no significant harmony- or voice-dependent differences in how popular the subjects perceived tracks. Only `Instrument` was statistically significant, and the differences between consecutive, (ordinally sorted) per-level effects were similarly large and significant across all of its levels. Sections , , and will expand on this more.

4.1.1 Does Instrument exert the strongest influence on perceived resemblance to each genre?

If we take a look at the correlation plots in Figure 8, it seems as though the correlations are strongly negative for Classical ratings in the same instances where they are strongly positive for popular ratings, and vice-versa. The electric guitar is strongly correlated positively with perceived resemblance to popular music, but negatively with perceived resemblance to classical music. Meanwhile the opposite is the case for the string quartet. And piano lies very close to the middle, yet still very weakly echoes the correlation pattern we saw with electric guitar. The boxplots from Figures 7 and 6 show this too.

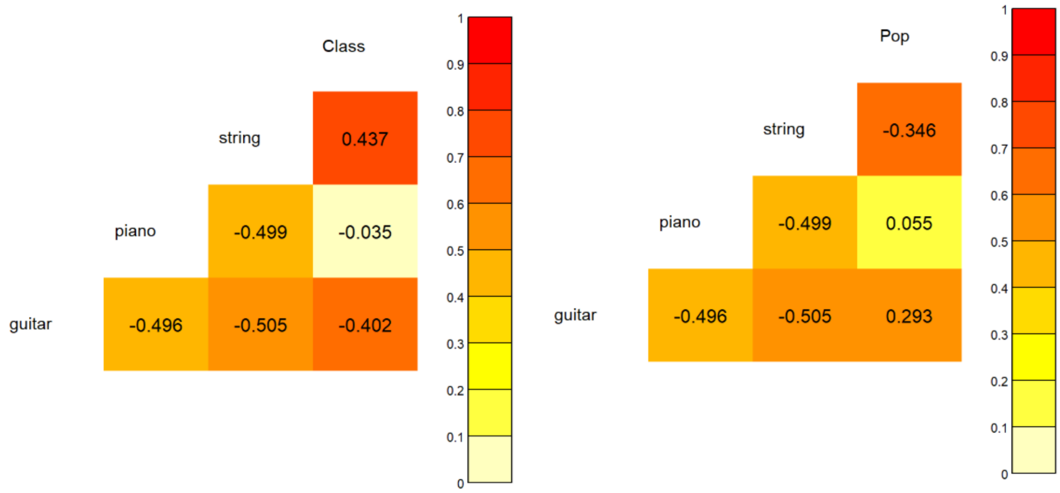


Figure 8: Correlation Plots of the levels of `Instrument` with each genre. The cells are color-coded by absolute value of correlation.

4.1.2 Among the levels of Harmonic Motion, does *I-V-VI* have the strongest association with perceived Classical-ness?

Within Harmonic Motion, the level that stands out for its effect on Classical rating the most is the *I-V-VI* level. Another look at that boxplot in Figure 6 demonstrates this too. The correlation matrix in Figure **??** puts this into numbers. It seems that, based on the shading and number is this figure, the differences between average classical ratings of songs splitting any two motions that are not *I-V-VI* don't seem to be very large, while a difference between average classical ratings of songs with a Harmonic Motion of *I-V-VI* and songs of any one of the other three is valid.

Going back to the final model in Figure 4, out of the three fixed effects on Classical ratings that contain `Harmony`, in both cases in which a statistically significant effect for a level of `Harmony` exists, *I-V-VI* is the only level with a statistically significant effect size. The others' effects on Classical ratings not are significantly different from the baseline's (*I-IV-V*).

4.1.3 Does a subject's familiarity with either of the Pachelbel rants/comedy bits seem to interact with the influence of harmony on his or her ratings of perceived Classical-ness?

Looking at the model in Figure 4, we see that interactions between `KnowRob` and `Harmony` are statistically significant only in the *HarmonyI-V-VI* case, same as we saw in the levels of main effects for `Harmony`. In fact, we see that this is one of the two interactions terms that has claimed responsibility for *HarmonyI-V-VI*'s effect to the point that the main effect *HarmonyI-V-VI* isn't significant anymore. However, it should be noted that, other than as an interaction with `Harmony`, familiarity with either comedy bit has no significant effect on how classical a subject perceives a track. This suggests that the rant had an influence on the people who watched it.

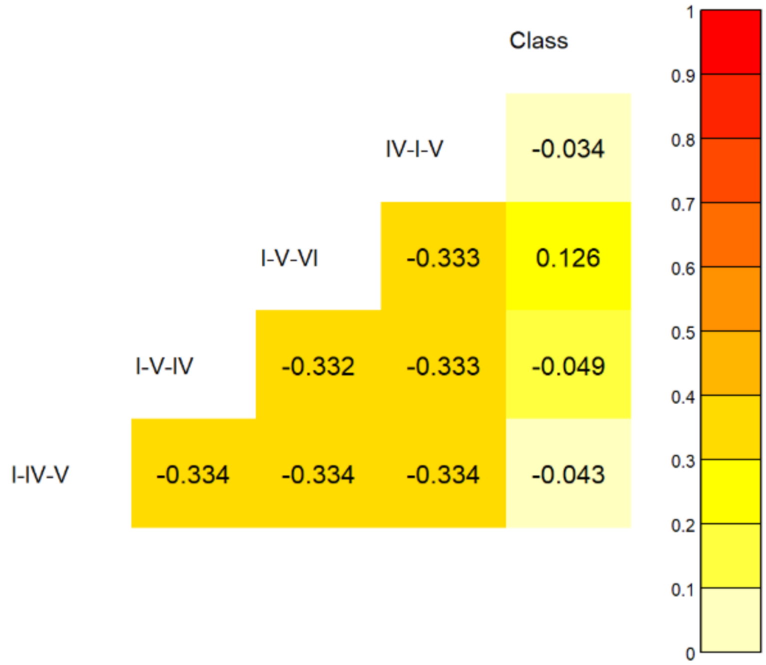


Figure 9: Correlation Plots of the levels of `Harmony` with Classical ratings. The cells are color-coded by absolute value of correlation.

4.1.4 Among the levels of Voice Leading, does *contrary motion* have the strongest association with perceived Classical-ness?

Within Voice Leading, it's contrary voices that have the most significant effects on Classical rating. The bottom boxplot in Figure 6 shows some evidence of this. The correlation matrix in Figure 10 shows this as well, but in a different way. The patterns we are observing here are much like what we saw with `Harmony`, though the overall magnitudes are smaller. The correlations between classical rating and each of *Voicepar3rd* and *Voicepar3rd* are very small and both negative, while the correlation between classical rating *Voicecontrary* is large in comparison, and positive. Returning, once again, to the model in Figure 2, we see that the highest, and sole significant effect size of a level of Voice is that of *Voicecontrary*. *Voicepar3rd* and *Voicepar3rd* are hardly different from each other.

4.2 Are there differences in the ways that musicians and non-musicians perceive classical music?

When I performed an analysis of variance (ANOVA) test to see if there were significant differences between the model with `is.musician` and without it, I found that the model with `is.musician` and the interaction of `is.musician` and `Harmony` was strongly significantly better than the one lacking the interaction. Yet, there was not statistically significant evidence of improvement in the jump from from a model with no effect for `is.musician` to a model with `is.musician`, but no `Harmony:is.musician` interaction.

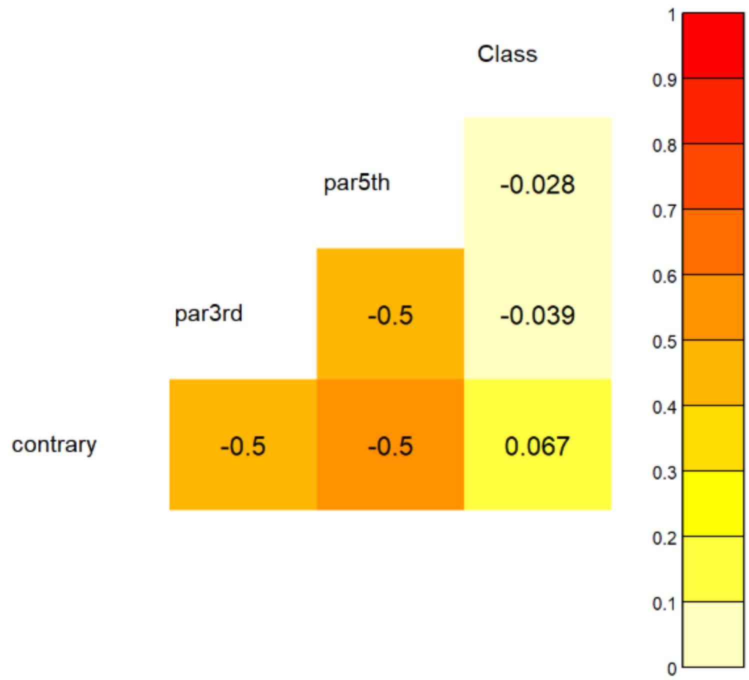


Figure 10: Correlation Plots of the levels of `Voice` with Classical ratings. The cells are color-coded by absolute value of correlation.

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
wo_musician	34	9899.475	10097.40	-4915.738	9831.475	NA	NA	NA
wo_HM	35	9900.579	10104.32	-4915.289	9830.579	0.8961472	1	0.3438169
final.model	38	9892.222	10113.43	-4908.111	9816.222	14.3565139	3	0.0024579

Figure 11: Anova Table for test of significance of `is.musician` effect.

4.3 Are there differences in the variables that drive ratings of perceived Classical-ness vs perceived Popular-ness?

While the models of Classical rating and Popular ratings contain the same combinations of main experimental factors, granted this is partly due to force, the individual covariates are different for different genres.

While subjects' perceptions of how classical a track is seems to depend to whether or not they are a musician, subjects' perceptions of how classical a track is seems to depend to whether or not they are a musician, subjects' perceptions of how popular a track is seems to depend to an auxiliary measure of their ability to distinguish classical vs popular music. Meanwhile, whether or not subjects are familiar with the Pachelbel rant is an attribute of both of the genres' models. For each genre, the coefficients can be interpreted as follows:

Classical:

On average, the expected perception of "Classical-ness," for guitar music with a harmonic motion of I-IV-V and Parallel 3rds for leading voice, of non-musicians who didn't see the Pachelbel rant and rank is 5.13, but there's some variation among subjects, and the variance associated with this variation is 2.317, all else held constant.

On average, the overall effects on subjects' perception of "Classical-ness" associated with sound from piano or strings instead of guitar are 1.36 and 3.11 respectively, but there's some variation among subjects, and the variances associated with this variation are 1.67 and 3.50 respectively, all else held constant.

On average, the overall effects on subjects' perception of "Classical-ness" associated with a harmonic motion of I-V-IV, I-V-VI, or IV-I-V instead of I-IV-V are -0.17, -0.47, and -0.15 respectively, but there's some variation among subjects, and the variances associated with this variation are 0.06, 1.18 and 0.01 respectively, all else held constant.

On average, the overall effects on subjects' perception of "Classical-ness" associated with a leading voice of par5th or contrary instead of par3rd are -0.39 and -0.36 respectively, all else held constant.

On average, a musician's perception of how Classical a song sounds is expected to be 1.0 unit less that a non-musician's, all else held constant.

On average, the overall effect on subjects' perception of "Classical-ness" associated with subjects who've seen Rob's rant is lower than that associated with subjects who've never seen it by 0.12 units, all else held constant.

On average, the overall effect on musicians' perception of "Classical-ness" is higher than non-musicians' by an additional 0.19 units 1.37, or 0.24 when harmonic motion is I-V-IV, I-V-VI, or IV-I-V instead of I-IV-V, respectively, all else held constant.

On average, the overall effect on subjects' perception of "Classical-ness" associated with subjects who've seen Rob's rant is higher than that associated with subjects who've never seen it by an additional -0.01 units 0.93, or 0.02 when harmonic motion is I-V-IV, I-V-VI, or IV-I-V instead of I-IV-V, respectively, all else held constant.

Popular:

On average, the expected perception of "Popular-ness," for guitar music with a harmonic motion of I-IV-V and Parallel 3rds for leading voice, of subjects who didn't see Rob's rant is 5.65, but there's some variation among subjects, and the variance associated with this variation is 3.83, all else held constant.

On average, the overall effects on subjects' perception of "Popular-ness" associated with sound from piano or strings instead of guitar are -0.89 and -2.41 respectively, but there's some variation among subjects, and the variances associated with this variation are 1.44 and 3.67 respectively, all else held constant.

On average, the overall effects on subjects' perception of "Popular-ness" associated with a harmonic motion of I-V-IV, I-V-VI, or IV-I-V instead of I-IV-V are -0.03, -0.20, and -0.17 respectively, but there's some variation among subjects, and the variances associated with this variation are 0.12, 0.80 and 0.27 respectively, all else held constant.

On average, the overall effects on subjects' perception of "Popular-ness" associated with a leading voice of par5th or contrary instead of par3rd are 0.14 and 0.16 respectively, all else held constant.

On average, the overall effect on subjects' perception of "Popular-ness" associated with a one unit increase in an auxiliary measure of their ability to distinguish classical vs popular music, is 0.13, all else held constant.

On average, the overall effect on subjects' perception of "Classical-ness" associated with subjects who've seen Rob's rant is higher than that associated with subjects who've never seen it by 1.38 units, all else held constant.

5 Discussion

5.1 What experimental factor (combination) has the strongest influence on how Classical or Popular a subject perceives a given track?

After careful analysis of the data, I have discovered some interesting patterns. It seems as though the effects of music played on pianos and strings instruments are more often perceived as classical compared to music played on the guitar. However, the opposite seems to be the case for popular music. It's also apparent that differences within each of the three main experimental factors seem to be significantly associated with differences in how "Classical" a song is perceived as by subjects. Yet this really was not true of the differences in how "Popular" a song is perceived as by subjects. As a matter of fact, it seems as though, of the three, differences in the instrument used to play the music is the only main effect that was significantly associated with differences in perceived "Popular-ness." Meanwhile, harmony and voice were associated with differences in perceived "Classical-ness" in some prominent way, be it as main effects or interactions.

All in all, more needs to be done before anything can be concluded, but these are some initial findings.

5.1.1 Does Instrument exert the strongest influence on perceived resemblance to each genre?

The results were crystal clear. Instrument was the strongest deciding factor both for classical and popular ratings. But it won by a larger margin in the popular ratings model. The instrument probably most associated with popular music is the guitar, especially in its electric state. Although classical guitar does exist, I've never heard of it begin played on electric guitar. On the flip side, the string quartet is a rare find in pop music, but common in classical. Piano can go either way. (Think Coldplay or Mozart.) It's a spectrum, which is something like what we observed in the model, as an instrument gets more associated with pop music, it is less associated with classical.

5.1.2 Among the levels of Harmonic Motion, does *I-V-VI* have the strongest association with perceived Classical-ness?

Harmonic Motion was an interesting player in the classical ratings model. Essentially, the only level of it that had any effect at all was *I-V-VI* level, which represented the Harmonic Motion of Pachelbel's Canon in D.

5.1.3 Does a subject's familiarity with either of the Pachelbel rants/comedy bits seem to interact with the influence of harmony on his or her ratings of perceived Classical-ness?

The fact that the influence of `KnowRob` alone has no significant effect on classical rating, yet when paired with `Harmony:I-V-VI` becomes significant (according to the ANOVA test) is very powerful. First off, when it comes to observational studies like this, there is always the chance of lurking variables. It could be the case that someone who has seen the rant saw it because they are particularly interested in music, and, as a side effect, they watch videos of musical youtubers more than the average person. This would also make them more likely to just know more about music, and perhaps, consider themself a musician to some degree. But the fact that both `is.musician` and `KnowRob` are significant interactions of `Harmony` means that they don't cover very much of the same ground, and therefore, the information we are gaining from having one in a model that already contains the other is new information, that probably represents a real underlying relationship.

That makes us more free to say that, in light of this, and bringing in the fact that the strongest level of `Harmony` associated with `KnowRob` is the one from Pachelbel's Canon, the people who saw the video might even have learned something from watching it that the people who didn't see the video didn't learn. Of course, correlation does not equal causation, so more information would need to be gathered in a controlled setting to draw any conclusion for real.

5.1.4 Among the levels of Voice Leading, does *contrary motion* have the strongest association with perceived Classical-ness?

The correlation plot and the coefficients agree that contrary motion is more influential that either 3rd parallels or 5th parallels as far as Leading Voice goes. However, the correlation and effect sizes were not very big either way, even if they were significant.

5.2 Are there differences in the ways that musicians and non-musicians perceive classical music?

When it's up to each person to declare whether or not they are a musician, there is always the issue of personal opinion. Nowadays, especially, when people can make music on a laptop, being a musician doesn't always imply that music literacy, theory lessons, or even instruments, were part of the process. This means it can be hard to tell how useful the information really is. However, based on the results of the ANOVA test, it looks, on average, the information tells us something useful. Assuming that this difference is due to the idea that a musician is likely to have a more trained ear for classical music than a non-musician, and not the other way around, I interpret these results in the following way:

Whereas non-musicians and musicians alike can tell the difference between instruments and voices, and can choose to weigh these differences consistently while deciding how classical to rate a song they hear, harmonic motion does not follow suit. While someone with no music-making experience may be able to note that a difference exists between the way that two different harmonic motions sound when put to song, they might not be able to identify what the difference springs from. It's probable that musician would be more likely to attribute that difference to harmonics, but that's not something we can infer from the results. We can only see that there is a stronger relationship between the harmonics and the classical ratings given by musicians than the ones given by non musicians. From that, the most reasonable inference I can make is that the strength of the relationship comes from consistency. In a situation in which song A and song C are in *I-IV-V* while song B and song D are in *I-V-VI*, the musician more than the non-musician, on average, would mentally connect the distinction he made between song A and song B as being the same as the distinction he makes later between song C and song D.

5.3 Are there differences in the variables that drive ratings of perceived Classical-ness vs perceived Popular-ness?

The effects of Instrument on Classical vs Popular ratings have inverse effects. As "popular" is kind of a non-genre, you can't expect a lot of consistency when a in the rules surrounding it. Compared to "pop", even "classical" seems like a cohesive genre (though this can be debated too). There aren't any standards restricting the notes or harmonics used in popular music. Meanwhile, classical ratings were influenced by things such as Harmonics and even Leading Voice, even if not to as much of an extent as by instrument. Aside from danceability, if that can even be measured, I'm not sure there is an audio feature that's good for classifying music as pop. Of the features available in this data, `Instrument` is the only main experimental factor with levels I could imagine associating with different degrees of pertinence to pop music.

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/.

Project 3 Code Appendix

Julia Stelman

12/8/2019

Appendix

A.1. Data Cleaning and Variable Transformation

I first had to take care of unclean and missing data.

A.1.1 Variable Removal

Two variables were removed immediately for high missing value rate.

```
dat <- dat %>% select(-X2ndInstr, -X1stInstr)
dat <- dat[,c(sort(names(dat)))]
```

A.1.2 Response Data-cleaning

If an observation was missing a response variable, (i.e. `Classical` or `Popular`) It was obliterated. There were some instances of slopiness in the response variables, or as I call them, illegal values (values not in the integer range of [1,10]). Decisions were made based on patterns in the data about how to legalize the illegal values.

```
# if there are missing response variable, just delete it.
dat <- dat %>% filter(!is.na(Classical))
dat <- dat %>% filter(!is.na(Popular))
# make variables alphabetical for consistency purposes
dat <- dat[,c(sort(names(dat)))]

# now that we've done "NA"s let's look at zero instances as missingness candidates
(dat[which(dat$Classical == 0),"Subject"]) %>% as.character() %>% as.factor() %>% table()
# .
# 40 71
#  1  7

#### let's see if we can infer s40's intended behavior regarding classical based on popular and
 viceversa
(dat[which(dat$Subject == 40),"Classical"] + dat[which(dat$Subject == 40),"Popular"]) %>% table
()
# .
#  9 10 11 12
#  1 32  2  1
#### ^ looks like most of the time they add up to around 10
dat[which(dat$Classical == 0 & dat$Subject == 40),c("Classical","Popular")]
#     Classical Popular
# 883         0      10
##### let's use intuition to impute a 1 here
dat[which(dat$Classical == 0 & dat$Subject == 40),"Classical"] <- 1

#### let's see if we can infer s71's intended behavior regarding classical based on popular and
 viceversa
(dat[which(dat$Subject == 71),"Classical"] + dat[which(dat$Subject == 71),"Popular"]) %>% table
()
# .
#  4  5  7  8  9 10 11 12 13
#  1  1  4  3  7 15  2  2  1
#### ^ looks this is not going to be so easy, let's keep exploring
dat[which(dat$Subject == 71),c("Classical","Popular")] %>% table()
#          Popular
# Classical 0 1 2 3 4 5 7 8 9 10
#        0  0 0 0 0 1 1 0 1 2  2
#        1  0 0 0 0 0 0 0 1 1  0
#        2  0 0 0 0 0 0 1 2 0  0
#        3  0 0 0 0 0 0 1 0 0  0
#        4  0 0 0 0 0 0 1 0 0  0
#        5  0 0 0 0 0 2 1 0 0  0
#        7  4 0 1 1 1 1 0 0 0  0
#        8  2 1 0 0 0 1 0 0 0  0
#        9  1 0 0 0 0 0 0 0 0  0
#        10 6 0 0 0 0 0 0 0 0  0
##### ^ looks like this subject gave responses of 0 for both genre ratings, and there doesn't se
em to be an obvious reason for when and where they did this.
##### I'm thinking they just didn't understand that the minimum of the scale was 1.
##### So I'm going to go ahead and correct that by making the 0s into 1s.
dat[which(dat$Classical == 0 & dat$Subject == 71),"Classical"] <- 1
dat[which(dat$Popular == 0 & dat$Subject == 71),"Popular"] <- 1
```

```
# now for popular
(dat[which(dat$Popular == 0),"Subject"]) %>% as.character() %>% as.factor() %>% table()
# .
# 18b  51  53
#   1   1  10

#### let's see if we can infer s40's intended behavior regarding popular based on classical and
 viceversa
(dat[which(dat$Subject == "18b"),"Classical"] + dat[which(dat$Subject == "18b"),"Popular"]) %>%
 table()
# .
#  6  7  8  9 10 11
#  1  1  3 12 17  2
#### ^ looks this is not going to be so easy, let's keep exploring
dat[which(dat$Subject == "18b"),c("Classical","Popular")] %>% table()
#          Popular
# Classical 0 1 2 3 4 5 6 7 8 9
#         1 0 0 0 0 0 0 0 0 1 2
#         2 0 0 0 0 0 0 0 2 3 0
#         3 0 0 0 0 0 0 3 3 1 0
#         4 0 0 0 1 2 1 3 1 0 0
#         5 0 0 0 1 3 1 0 0 0 0
#         6 1 0 0 2 2 0 0 0 0 0
#         7 0 0 0 1 0 0 0 0 0 0
#         8 0 0 1 0 0 0 0 0 0 0
#         9 0 1 0 0 0 0 0 0 0 0
##### ^ looks like this subject gave responses on a scale of 0-9 instead 1-10 for both Classical
and popular. However, I don't know enough to be comfortable changing non-impossible answers.
##### Therefore, I'm going to go ahead and say, I think they meant that 0 to be the minimum
##### So I'll just shift it up to the in-scale minimun and make it a 1.
dat[which(dat$Popular == 0 & dat$Subject == "18b"),"Popular"] <- 1

#### let's see if we can infer s51's intended behavior regarding popular based on classical and
 viceversa
(dat[which(dat$Subject == 51),"Classical"] + dat[which(dat$Subject == 51),"Popular"]) %>% table
()
# .
#  9 10 11 12
#  2 16 14  4
#### ^ looks this is not going to be so easy, let's keep exploring
dat[which(dat$Popular == 0 & dat$Subject == 51),c("Classical","Popular")]
#      Classical Popular
# 1286        10       0
##### ^ looks like another runaway minimum
##### So I'll just shift it up to the in-scale minimun and make it a 1.
dat[which(dat$Popular == 0 & dat$Subject == 51),"Popular"] <- 1

#### let's see if we can infer s53's intended behavior regarding Popular based on popular and vi
ceversa
(dat[which(dat$Subject == 53),"Classical"] + dat[which(dat$Subject == 53),"Popular"]) %>% table
()
# .
```

```
#   5   6   7   8   9 10
#   1   1   4   3 10 17
#### ^ looks like most of the time they add up to around 7-10, but never more than 10
(dat[which(dat$Popular == 0 & dat$Subject == 53),c("Classical","Popular")]) %>% table()
#          Popular
# Classical  0
#         10 10
##### So this looks like an easy fix. Runaway minimum. 0s will become 1s.
dat[which(dat$Popular == 0 & dat$Subject == 53),"Popular"] <- 1

## Great! Now NAs and 0s are taken care of! Onto other problems with response variables.
# other illegal classical ratings
dat[-which(dat$Classical %in% 1:10),c("Subject","Classical")]
#      Subject Classical
# 1951      73      19.0
# 2196      80       9.5
# 2358      91       4.6
# 2366      91       3.5
# 2374      91       4.2
### If you look at a keyboard, it's clear that 19 was almost certainly a typo made by someone tr
ying to type "10"
dat[which(dat$Classical == 19 & dat$Subject == 73),"Classical"] <- 10

### As for the rest, let's just say some subjects were indecisive, and round the values to the n
earest legal number.
dat[-which(dat$Classical %in% 1:10 & dat$Classical <= 10),"Classical"] <-
  round(dat[-which(dat$Classical %in% 1:10 & dat$Classical <= 10),"Classical"])

## Great! Now for popular!
dat[-which(dat$Popular %in% 1:10),c("Subject","Popular")]
#      Subject Popular
# 1148      47    19.0
# 2194      80     3.5
# 2358      91     4.6
# 2366      91     6.8
# 2374      91     4.2
### another "10" disguised as a "19"
dat[which(dat$Popular == 19 & dat$Subject == 47),"Popular"] <- 10

### As for the rest, we're going to help some more indecisive subjects make a legal choice.
dat[-which(dat$Popular %in% 1:10 & dat$Popular <= 10),"Popular"] <-
  round(dat[-which(dat$Popular %in% 1:10 & dat$Popular <= 10),"Popular"])
```

A.1.3 Missing Music Education Data: Imputation

Then, I used a case-by-case imputation method for missing cases of the following four variables:

- APTheory
- CollegeMusic
- Composing (only observations where all four of these variables were missing)
- NoClass

The method I used was designed to take into account all the information we can borrow from any present variables that cover some of the same ground as missing a variable. For example, if `NoClass` (number of music classes) = 1, `APTheory` = 0, and `CollegeMusic` is missing, my method takes the liberty of setting `CollegeMusic` = 1, reasoning that, if someone took one music class, and it wasn't AP Music Theory in high school, they must have taken some music class in college.

```
# if all four missing, fill with 0
dat[!(!is.na(dat$CollegeMusic) | !is.na(dat$NoClass) | !is.na(dat$APTheory) | !is.na(dat$Composi
ng)),17:20] <- dat[!(!is.na(dat$CollegeMusic) | !is.na(dat$NoClass) | !is.na(dat$APTheory) | !i
s.na(dat$Composing)),17:20] %>%
  mutate(CollegeMusic = 0, NoClass = 0, APTheory = 0, Composing = 0)
# if all three missing, fill with 0
dat[!(!is.na(dat$CollegeMusic) | !is.na(dat$NoClass) | !is.na(dat$APTheory)),17:19] <- dat[!(!i
s.na(dat$CollegeMusic) | !is.na(dat$NoClass) | !is.na(dat$APTheory)),17:19] %>%
  mutate(CollegeMusic = 0, NoClass = 0, APTheory = 0)
# if there is No music classes is 1 and college classes is 1, and AP is missing, fill with 0
dat[(is.na(dat$CollegeMusic) | is.na(dat$NoClass) | is.na(dat$APTheory)),17:19] <-
  dat[(is.na(dat$CollegeMusic) | is.na(dat$NoClass) | is.na(dat$APTheory)),17:19] %>%
  mutate(APTheory = 0)
# if both missing: 0's (subj 25)
dat[(is.na(dat$CollegeMusic) & is.na(dat$NoClass)),17:18] <- dat[(is.na(dat$CollegeMusic) & is.n
a(dat$NoClass)),17:18] %>%
  mutate(CollegeMusic = 0, NoClass = 0)
# if college music is 1, then say they have taken 1 class, if it's 0, say 0
dat[which(dat$CollegeMusic == 1 & is.na(dat$NoClass)),18] <- 1
dat[which(dat$CollegeMusic == 0 & is.na(dat$NoClass)),18] <- 0
```

A.1.4 Other Missing Data: Median Imputation

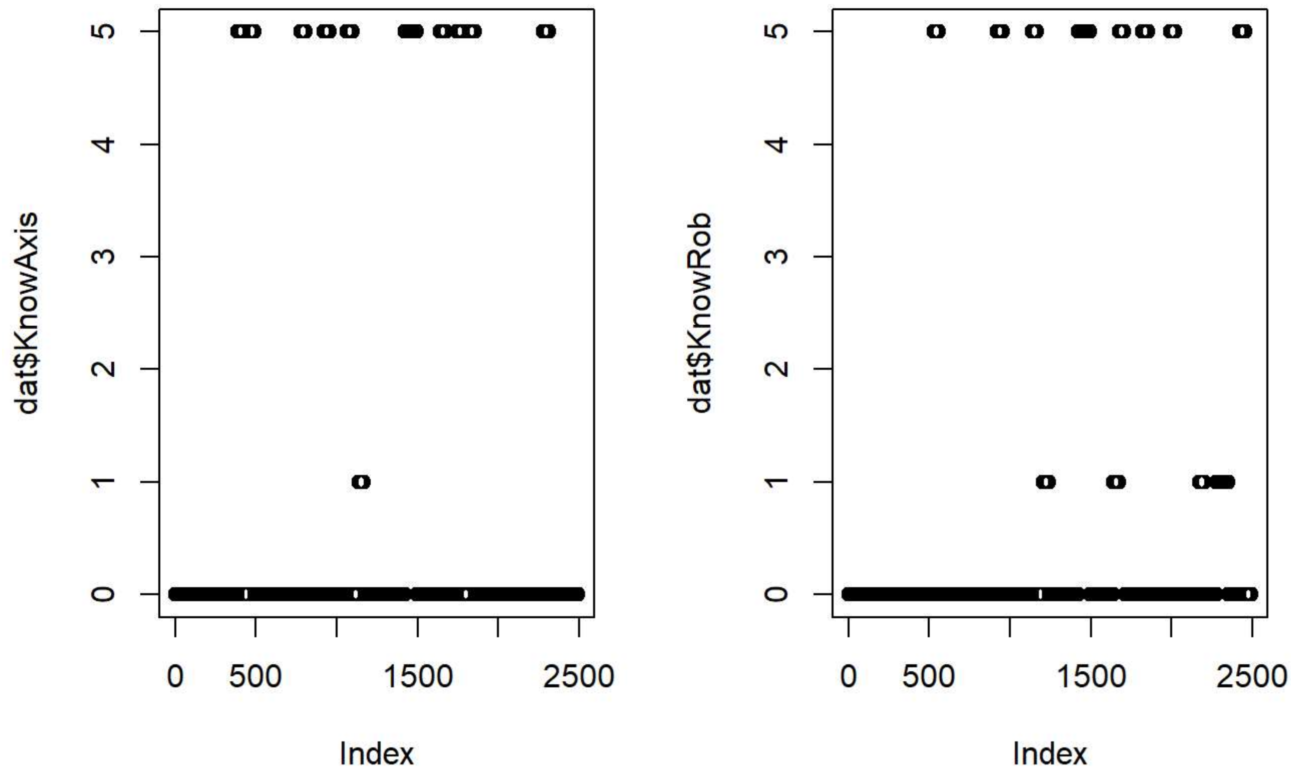
For the remaining numeric variables (including the rest of observations missing `Composing`), I use simple median imputation. There were 8 of these:

- ClsListen
- ConsNotes
- KnowAxis
- KnowRob
- PachListen
- X1990s2000s
- X1990s2000s.minus.1960s1970s
- Composing (what remained)

I chose median imputation simply because I couldn't think of a fitting case-based (yet practical) imputation strategy to use instead, and as most of these are ordinal discrete variables, the median made more sense than the mean.

A.1.5 Transforming Data: Dichotimizing Rank-scale Answers

`KnowAxis` and `KnowRob`, plotted below, are two examples of the "scale of 0-5"-style questions that I strongly believe should have been asked in the form of a "yes or no"-style question.



Judging by the plot, it seems as though there was no need for the scale, as all subjects, save one, chose either the maximun or the minimun. Any self-reported *a* to *b*-rank scale-style answers (usually *0* to *5*, but *1* to *6* for `Selfdeclare`) are likely going to be unreliable, as each person has their own idea of what differentiates a 2 from a 3 or a 4. I have more trust in the idea that most everyone would have the same standards for an answer of 0 (1 for `Selfdeclare`), since the definitions of 0 (1 for `Selfdeclare`) were explicitly defined by the experimenters, and therefore are the same for all subjects. Thus, I am making 8 new variables by dichotimizing rank-scale variables between 0 and 1 (1 and 2 for `Selfdeclare`), resulting in *yes/no*-style answers. I've left most of these variables' original versions in the data just for caution, however, KnowAxis and KnowRob have been completely replaced, as it's clear from the plots that the numbers won't provide useful insight. The transformed boolean variables' names are listed below:

- ClsListenDum
- ComposingDum
- GuitarPlayDum
- KnowAxis
- KnowRob
- PachListenDum
- PianoPlayDum
- is.musician (from Selfdeclare)

Technically, I don't used the original versions of these variables in any analysis, Though `Selfdeclare` does appear in two plots.

```
#I will change these and others like it to factors

attach(dat)
########## Edited! ###############
dat$PachListenDum <- as.integer(PachListen>0)
dat$ClsListenDum <- as.integer(ClsListen>0)
dat$ComposingDum <- as.integer(Composing>0)
# don't even bother keeping originals of next two
dat$KnowAxis <- as.integer(KnowAxis>0)
dat$KnowRob <- as.integer(KnowRob>0)
dat$PianoPlayDum <- as.integer(PianoPlay>0)
dat$GuitarPlayDum <- as.integer(GuitarPlay>0)
########## Inserted! ###############
dat$is.musician <- as.integer(Selfdeclare>1)
######### End insert
####### End edit
detach(dat)
```

A.1.6 Summary of the Final Data Cleaning

The finalized dataset is summarized below.

```
##     APTheory       Classical        ClsListen      CollegeMusic   
##  Min.   :0.000   Min.   : 1.000   Min.   :0.000   Min.   :0.0000  
##  1st Qu.:0.000   1st Qu.: 4.000   1st Qu.:1.000   1st Qu.:1.0000  
##  Median :0.000   Median : 6.000   Median :3.000   Median :1.0000  
##  Mean   :0.211   Mean   : 5.783   Mean   :2.162   Mean   :0.7978  
##  3rd Qu.:0.000   3rd Qu.: 8.000   3rd Qu.:3.000   3rd Qu.:1.0000  
##  Max.   :1.000   Max.   :10.000   Max.   :5.000   Max.   :1.0000  
##                                                                   
##    Composing        ConsInstr       ConsNotes       GuitarPlay    
##  Min.   :0.0000   Min.   :0.000   Min.   :0.000   Min.   :0.0000  
##  1st Qu.:0.0000   1st Qu.:1.670   1st Qu.:1.000   1st Qu.:0.0000  
##  Median :0.0000   Median :3.000   Median :3.000   Median :0.0000  
##  Mean   :0.9635   Mean   :2.869   Mean   :2.604   Mean   :0.6807  
##  3rd Qu.:2.0000   3rd Qu.:4.330   3rd Qu.:4.000   3rd Qu.:1.0000  
##  Max.   :5.0000   Max.   :5.000   Max.   :5.000   Max.   :5.0000  
##                                                                   
##    Harmony    Instr.minus.Notes  Instrument     KnowAxis     
##  I-IV-V:625   Min.   :-4.000    guitar:833   Min.   :0.0000  
##  I-V-IV:622   1st Qu.: 0.000    piano :820   1st Qu.:0.0000  
##  I-V-VI:622   Median : 0.670    string:840   Median :0.0000  
##  IV-I-V:624   Mean   : 0.698                 Mean   :0.1733  
##               3rd Qu.: 2.000                 3rd Qu.:0.0000  
##               Max.   : 4.330                 Max.   :1.0000  
##                                                              
##     KnowRob          NoClass            OMSI         PachListen   
##  Min.   :0.0000   Min.   :0.0000   Min.   : 11.0   Min.   :0.000  
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.: 49.0   1st Qu.:5.000  
##  Median :0.0000   Median :1.0000   Median :145.0   Median :5.000  
##  Mean   :0.1998   Mean   :0.9314   Mean   :225.5   Mean   :4.523  
##  3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:323.0   3rd Qu.:5.000  
##  Max.   :1.0000   Max.   :8.0000   Max.   :970.0   Max.   :5.000  
##                                                                   
##    PianoPlay        Popular        Selfdeclare       Subject    
##  Min.   :0.000   Min.   : 0.000   Min.   :1.000   15     :  36  
##  1st Qu.:0.000   1st Qu.: 3.000   1st Qu.:2.000   16     :  36  
##  Median :0.000   Median : 5.000   Median :2.000   17     :  36  
##  Mean   :1.095   Mean   : 5.046   Mean   :2.444   18b    :  36  
##  3rd Qu.:1.000   3rd Qu.: 7.000   3rd Qu.:3.000   19     :  36  
##  Max.   :5.000   Max.   :10.000   Max.   :6.000   20     :  36  
##                                                   (Other):2277  
##       Voice      X16.minus.17    X1990s2000s   
##  par3rd  :830   Min.   :-4.00   Min.   :0.000  
##  par5th  :832   1st Qu.: 0.00   1st Qu.:3.000  
##  contrary:831   Median : 1.00   Median :5.000  
##                 Mean   : 1.69   Mean   :4.105  
##                 3rd Qu.: 3.00   3rd Qu.:5.000  
##                 Max.   : 9.00   Max.   :5.000  
##                                                
##  X1990s2000s.minus.1960s1970s PachListenDum     ClsListenDum   
##  Min.   :-4.000               Min.   :0.0000   Min.   :0.0000  
##  1st Qu.: 1.000               1st Qu.:1.0000   1st Qu.:1.0000  
##  Median : 2.000               Median :1.0000   Median :1.0000  
##  Mean   : 2.017               Mean   :0.9856   Mean   :0.8412  
```

```
##  3rd Qu.: 3.000               3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   : 5.000               Max.   :1.0000   Max.   :1.0000
##
##   ComposingDum     PianoPlayDum   GuitarPlayDum     is.musician
##  Min.   :0.0000   Min.   :0.000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:1.0000
##  Median :0.0000   Median :0.000   Median :0.0000   Median :1.0000
##  Mean   :0.3791   Mean   :0.416   Mean   :0.2571   Mean   :0.7738
##  3rd Qu.:1.0000   3rd Qu.:1.000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.000   Max.   :1.0000   Max.   :1.0000
##
```

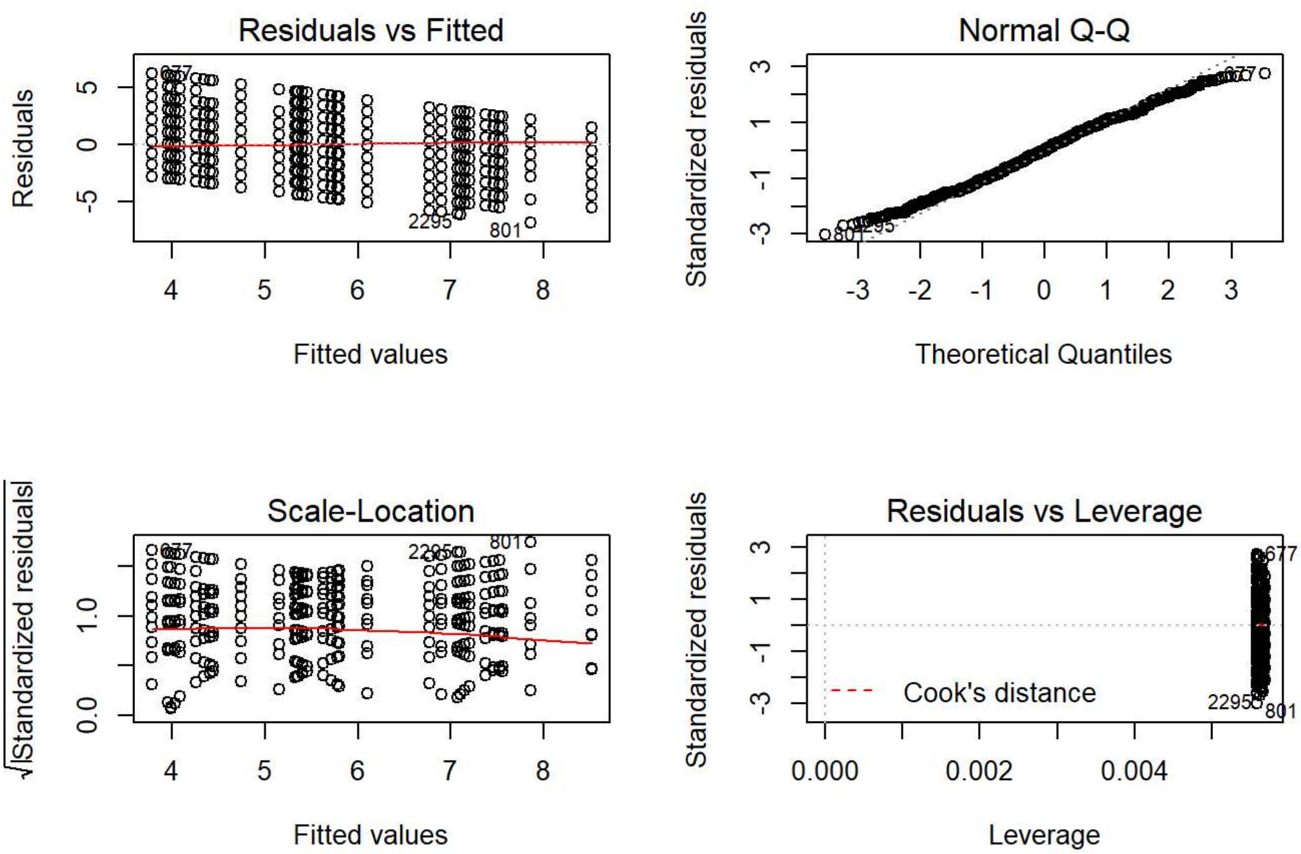
A.2 Modeling Classical

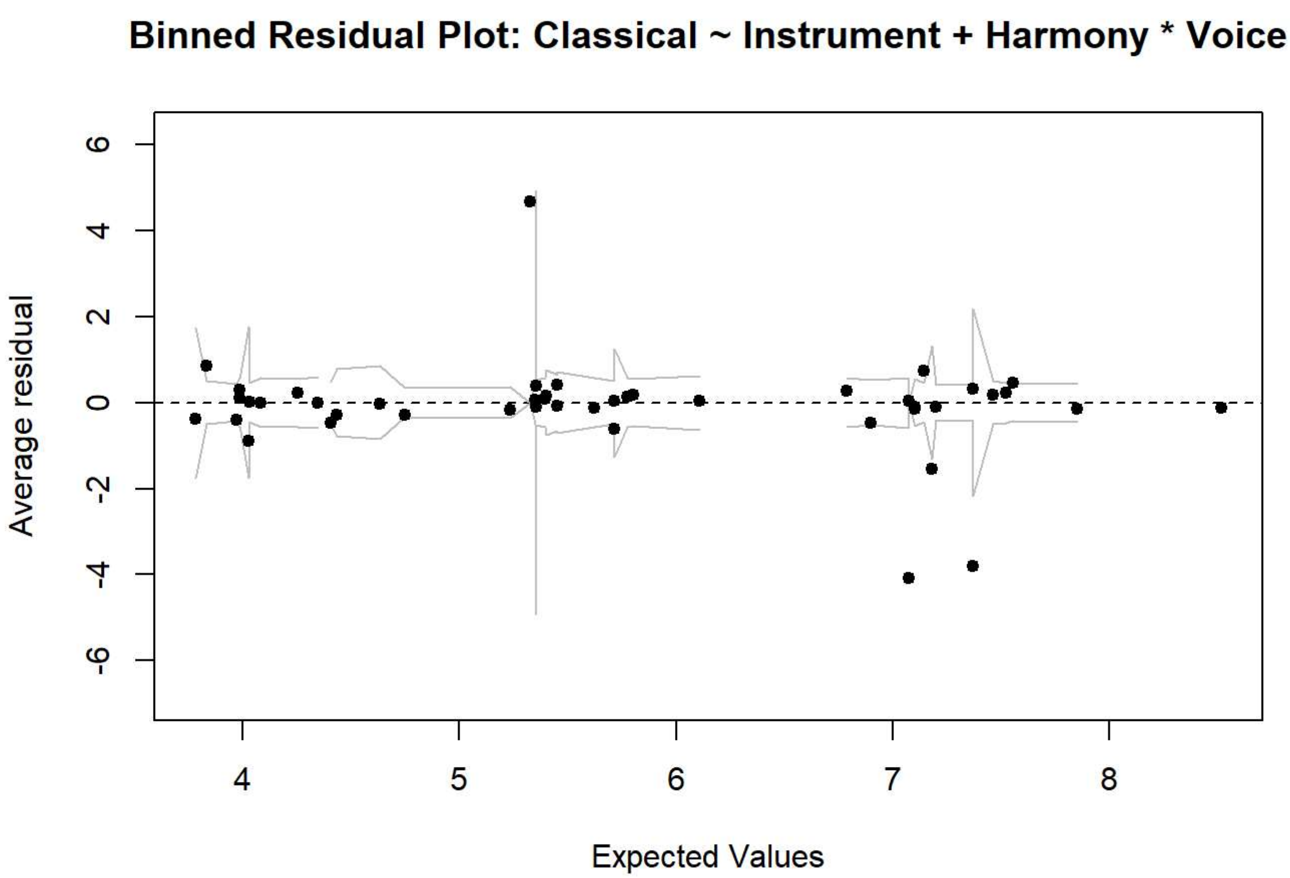
A.2.1 EDA: Conventional Linear Model

Using conventional linear models, I modeled the influence of the `Instrument`, `Harmony`, & `Voice` variables on `Classical`. I began with all possible main and interaction effects, and used R's **arm** library's *stepAIC()* function to pick away the ones that weren't needed. Ultimately, no main effects, and all interactions besides voice and harmony, were dropped. See the summary and code below.

```
lmFull <- lm(Classical ~ Instrument * Harmony * Voice , data = dat)
lm3 <- stepAIC(lmFull,trace = F)
summary(lm3)
# Call:
# lm(formula = Classical ~ Instrument + Harmony + Voice + Harmony:Voice,
#     data = dat)
#
# Residuals:
#    Min     1Q Median     3Q    Max
# -6.855 -1.740  0.013  1.653  6.217
#
# Coefficients:
#                              Estimate Std. Error t value Pr(>|t|)
# (Intercept)                   3.98700    0.17015  23.433   <2e-16 ***
# Instrumentpiano               1.36681    0.11191  12.213   <2e-16 ***
# Instrumentstring              3.11485    0.11124  28.001   <2e-16 ***
# HarmonyI-V-IV                -0.20399    0.22254  -0.917   0.3594
# HarmonyI-V-VI                 0.45175    0.22335   2.023   0.0432 *
# HarmonyIV-I-V                 0.35952    0.22335   1.610   0.1076
# Voicepar5th                   0.04384    0.22308   0.197   0.8442
# Voicecontrary                 0.26970    0.22254   1.212   0.2257
# HarmonyI-V-IV:Voicepar5th     0.15874    0.31548   0.503   0.6149
# HarmonyI-V-VI:Voicepar5th     0.25736    0.31568   0.815   0.4150
# HarmonyIV-I-V:Voicepar5th    -0.43207    0.31568  -1.369   0.1712
# HarmonyI-V-IV:Voicecontrary  0.35692    0.31530   1.132   0.2577
# HarmonyI-V-VI:Voicecontrary  0.69259    0.31568   2.194   0.0283 *
# HarmonyIV-I-V:Voicecontrary -0.53177    0.31530  -1.687   0.0918 .
# ---
# Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Residual standard error: 2.275 on 2479 degrees of freedom
# Multiple R-squared:  0.2607,  Adjusted R-squared:  0.2569
# F-statistic: 67.26 on 13 and 2479 DF,  p-value: < 2.2e-16
```

Some diagnostic plots of that model below:


Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Cook's distance
Leverage


Binned Residual Plot: Classical ~ Instrument + Harmony * Voice
Average residual
Expected Values

A.2.2 EDA: Repeated Measures Model

A.2.2.0 EDA: Repeated Measures Model setup

Adding a random intercept to this model for `Subject`. This is called a *Repeated Measures Model*. Below is the hierarchical model representation.

Note: $\overrightarrow{1}_{\square i}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

$$Classical_i = \alpha_{0i} + \overrightarrow{\alpha_1}^T \overrightarrow{1}_{Hi} + \overrightarrow{\alpha_2}^T \overrightarrow{1}_{Ii} + \overrightarrow{\alpha_3}^T \overrightarrow{1}_{Vi} + \overrightarrow{\alpha_4}^T \overrightarrow{1}_{HVi} + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j} \sim N(0, \tau^2)$$

A.2.2.1 EDA: Confirming Repeated Measures Model is Appropriate!

I used R's **lme4** library's *lmer()* function to model the model stated above. Using R's **RLRsim** library's *exactRLRT()* function, I was able to test the significance of the subject-based random intercept in the model. Based on the tiny p-value, I've concluded that the random intercept is very much an asset to the model. A summary of that model and the results of the significance test are below along with the code.

```
repeated.measures.model <- lmer(Classical ~ Instrument + Harmony * Voice + (1|Subject), data = d
at,
                control = lmerControl(optimizer = "bobyqa"))
display(repeated.measures.model)
# lmer(formula = Classical ~ Instrument + Harmony * Voice + (1 |
#     Subject), data = dat, control = lmerControl(optimizer = "bobyqa"))
#                             coef.est coef.se
# (Intercept)                  3.99     0.21
# Instrumentpiano              1.37     0.09
# Instrumentstring             3.11     0.09
# HarmonyI-V-IV               -0.20     0.18
# HarmonyI-V-VI                0.45     0.18
# HarmonyIV-I-V                0.36     0.18
# Voicepar5th                  0.04     0.18
# Voicecontrary                0.27     0.18
# HarmonyI-V-IV:Voicepar5th    0.15     0.26
# HarmonyI-V-VI:Voicepar5th    0.26     0.26
# HarmonyIV-I-V:Voicepar5th   -0.43     0.26
# HarmonyI-V-IV:Voicecontrary  0.35     0.26
# HarmonyI-V-VI:Voicecontrary  0.69     0.26
# HarmonyIV-I-V:Voicecontrary -0.53     0.26
#
# Error terms:
#  Groups   Name        Std.Dev.
#  Subject  (Intercept) 1.30
#  Residual             1.87
# ---
# number of obs: 2493, groups: Subject, 70
# AIC = 10440.1, DIC = 10347.5
# deviance = 10377.8

## Check to see if (1|Subject) is needed, with a simpler
## call to exactRLRT (note the simpler call to exactRLRT when there is
## just one random effect in HA and no random effect in H0):

exactRLRT(repeated.measures.model)
# RLRT = 778.61, p-value < 2.2e-16
# so it definitely makes sense to include the (1|Subject) random effect.
```

A.2.2.2 EDA: Repeated Measures Model Variable Selection

Using R's **LMERConvenienceFunctions** library's *fitLMER.fnc()* function with `BIC` for the method, I performed backwards elimination on fixed effects of the model above. The interaction was dropped. The resulting model is summarized below.

```
# try the fit function for lmer from the lmerconveniencefunctions library
## backwards elimination of fixed effects:

summary(repeated.measures.model.fit <- fitLMER.fnc(repeated.measures.model, method = "BIC"))
# Linear mixed model fit by REML ['lmerMod']
# Formula: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
#    Data: dat
# Control: lmerControl(optimizer = "bobyqa")
#
# REML criterion at convergence: 10425.7
#
# Scaled residuals:
#     Min      1Q  Median      3Q     Max
# -3.0192 -0.6456 -0.0153  0.6495  3.9190
#
# Random effects:
#  Groups   Name        Variance Std.Dev.
#  Subject  (Intercept) 1.691    1.300
#  Residual             3.514    1.875
# Number of obs: 2493, groups:  Subject, 70
#
# Fixed effects:
#                   Estimate Std. Error t value
# (Intercept)        3.94607    0.18816  20.972
# Instrumentpiano    1.36697    0.09231  14.808
# Instrumentstring   3.11389    0.09170  33.956
# HarmonyI-V-IV     -0.03396    0.10618  -0.320
# HarmonyI-V-VI      0.77082    0.10618   7.260
# HarmonyIV-I-V      0.03647    0.10609   0.344
# Voicepar5th        0.03913    0.09197   0.425
# Voicecontrary      0.40100    0.09200   4.359
#
# Correlation of Fixed Effects:
#             (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr5t
# Instrumntpn -0.243
# Instrmntstr -0.245  0.498
# HrmnyI-V-IV -0.282  0.001    -0.001
# HrmnyI-V-VI -0.281  0.001    -0.001     0.499
# HrmnyIV-I-V -0.280 -0.001    -0.001     0.499  0.499
# Voicepar5th -0.244  0.000     0.001     0.000 -0.004 -0.003
# Voicecntrry -0.245  0.001     0.001     0.002 -0.001 -0.002  0.500

# the model drops the interaction term
display(repeated.measures.model.fit)
# lmer(formula = Classical ~ Instrument + Harmony + Voice + (1 |
#     Subject), data = dat, REML = TRUE, control = lmerControl(optimizer = "bobyqa"))
#                   coef.est coef.se
# (Intercept)        3.95     0.19
# Instrumentpiano    1.37     0.09
# Instrumentstring   3.11     0.09
# HarmonyI-V-IV     -0.03     0.11
# HarmonyI-V-VI      0.77     0.11
# HarmonyIV-I-V      0.04     0.11
```

```
# Voicepar5th       0.04     0.09
# Voicecontrary     0.40     0.09
#
# Error terms:
#  Groups   Name        Std.Dev.
#  Subject  (Intercept) 1.30
#  Residual             1.87
# ---
# number of obs: 2493, groups: Subject, 70
# AIC = 10445.7, DIC = 10380.1
# deviance = 10402.9
```

Since the model was changed, here is the updated greek.

Note: $\overrightarrow{1}_{\square i}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

$$Classical_i = \alpha_{0i} + \overrightarrow{\alpha_1}^T \overrightarrow{1}_{Hi} + \overrightarrow{\alpha_2}^T \overrightarrow{1}_{Ii} + \overrightarrow{\alpha_3}^T \overrightarrow{1}_{Vi} + \epsilon_i, \epsilon_i \text{~} N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j} \text{~} N(0, \tau^2)$$

A.2.3 EDA: Personal Biases Model

A.2.3.1 EDA: Personal Biases Model Variable Selection

Using R's **LMERConvenienceFunctions** library's *fitLMER.fnc()* function, again. This time I started with the final model from 2.2.2, and performed forwards selection of random effects (meaning I tried switching every fixed effect to a random effect, and seeing if there was an improvement). Two random effects were added when all was said and done. It looks like there are per subject effects for `Instrument` and `Harmony`.

```
## I am going to try switching every fixed slope to a random slope.
## first, we grab the vector of all the fixed effects; we're going to
## try a random slope on each one
vars = c("Instrument","Harmony","Voice")
## forward selecton of random effects
random.slopes.fit <- ffRanefLMER.fnc(repeated.measures.model.fit,
                                     ran.effects= list(slopes=vars, by.vars="Subject",
                                            corr=rep(0,length(vars))))
# the backfitting afterwards intented to remove the Harmony (Fixed and Random) effects altogethe
r. However, I did not want that. So I didn't permit backfitting.
# Clean it up.
random.slopes.fit <- update(random.slopes.fit, .~.-
                              (1|Subject)-(Instrument|Subject)-(Harmony|Subject)+
                         (1 + Instrument + Harmony | Subject))
display(random.slopes.fit)
# lmer(formula = Classical ~ Instrument + Harmony + Voice + (1 +
#     Instrument + Harmony | Subject), data = dat, REML = TRUE,
#     control = lmerControl(optimizer = "bobyqa"))
#                   coef.est coef.se
# (Intercept)       3.95     0.21
# Instrumentpiano   1.36     0.17
# Instrumentstring  3.11     0.24
# HarmonyI-V-IV    -0.03     0.09
# HarmonyI-V-VI     0.77     0.18
# HarmonyIV-I-V     0.04     0.09
# Voicepar5th       0.04     0.08
# Voicecontrary     0.40     0.08
#
# Error terms:
#  Groups   Name             Std.Dev. Corr
#  Subject  (Intercept)      1.60
#           Instrumentpiano  1.29     -0.38
#           Instrumentstring 1.87     -0.57  0.66
#           HarmonyI-V-IV    0.22      0.68 -0.65 -0.47
#           HarmonyI-V-VI    1.27     -0.05 -0.26 -0.42  0.22
#           HarmonyIV-I-V    0.08      0.11 -0.34  0.09 -0.05  0.08
#  Residual                  1.55
# ---
# number of obs: 2493, groups: Subject, 70
# AIC = 9922.1, DIC = 9821.6
# deviance = 9841.9

sigma(random.slopes.fit)^2
# 2.398394

anova(random.slopes.fit,
      repeated.measures.model.fit)
# Models:
# repeated.measures.model.fit: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
# random.slopes.fit: Classical ~ Instrument + Harmony + Voice + (1 + Instrument +
# random.slopes.fit:     Harmony | Subject)
#                             Df     AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
# repeated.measures.model.fit 10 10422.9 10481 -5201.4  10402.9
```

```
# random.slopes.fit          30  9901.9 10076 -4920.9   9841.9   561     20  < 2.2e-16
rm(vars)
```

A.2.3.2 Summary of Personal Biases Model

The model is written out in mathematical terms below. The table below shows the values of the various fixed effect parameters that number too many to fit into the model below.

Note: $\overrightarrow{1_{\square i}}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

$$Classical_i = \alpha_{0i} + \alpha_{1i} + \alpha_{2i} + \overrightarrow{\alpha_3}^T \overrightarrow{1_{Vi}} + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j} \sim N(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \overrightarrow{\beta_1}^T \overrightarrow{1_{Ii}} + \overrightarrow{\eta_{1j[i]}}^T \overrightarrow{1_{Ii}}, \overrightarrow{\eta_{1j}} \sim N(\vec{0}, \overrightarrow{\tau_1^2})$$

$$\alpha_{2j[i]} = \overrightarrow{\beta_2}^T \overrightarrow{1_{Hi}} + \overrightarrow{\eta_{2j[i]}}^T \overrightarrow{1_{Hi}}, \overrightarrow{\eta_{2j}} \sim N(\vec{0}, \overrightarrow{\tau_2^2})$$

The values and standard deviations of the coefficient vectors, $\overrightarrow{\alpha_?}$ are inscribed in the table below.

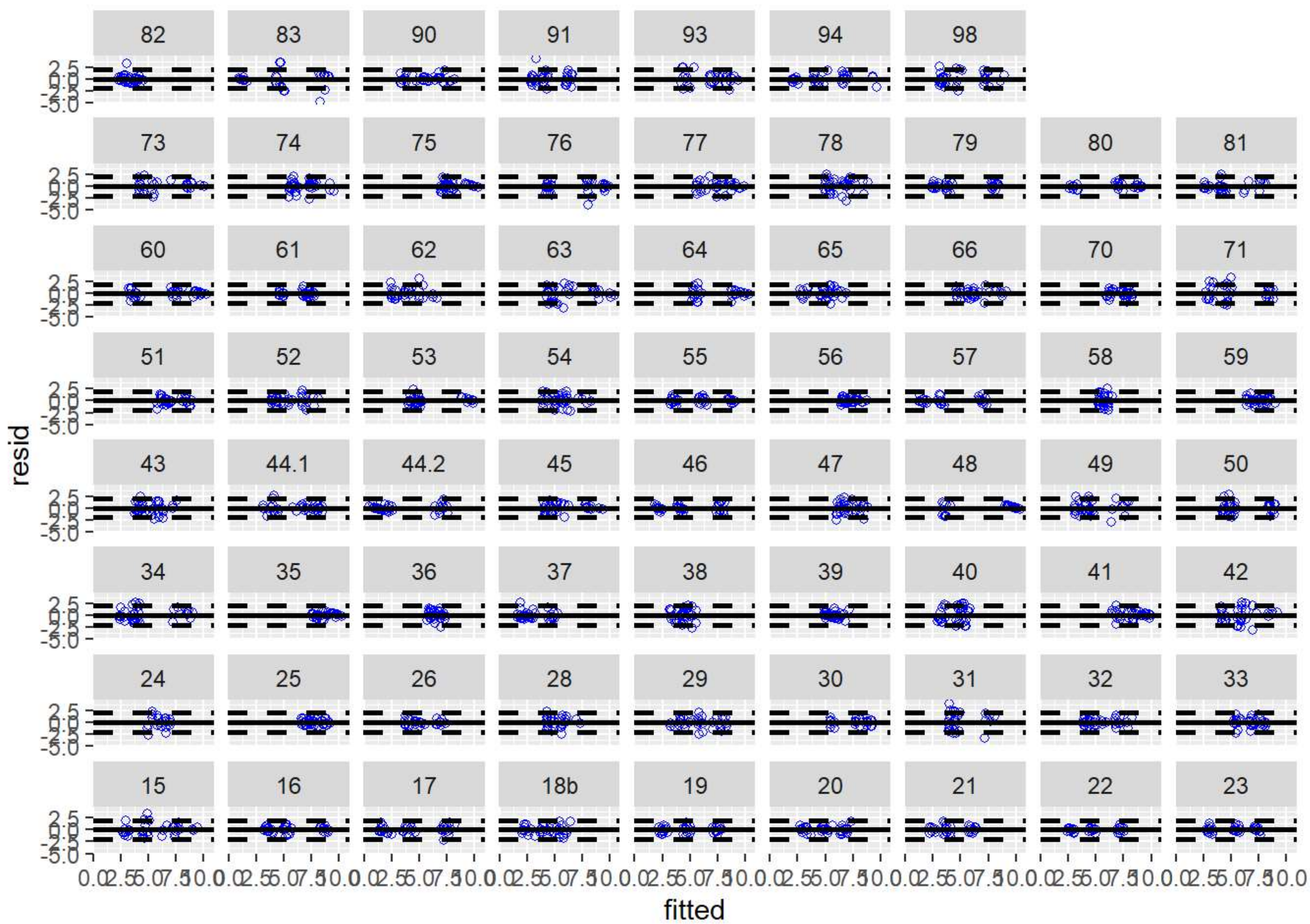
	factorLevel	Estimate	Std. Error
beta0	(Intercept)	3.9506	0.2106
beta1[1]	Instrumentpiano	1.3616	0.1723
beta1[2]	Instrumentstring	3.1107	0.2360
beta2[1]	HarmonyI-V-IV	-0.0317	0.0915
beta2[2]	HarmonyI-V-VI	0.7717	0.1758
beta2[3]	HarmonyIV-I-V	0.0383	0.0882
alpha3[1]	Voicepar5th	0.0373	0.0760
alpha3[2]	Voicecontrary	0.3950	0.0760

	factorLevel	Estimate
sigma_squared	Residual	2.3984
tau0_squared	(Intercept)	2.5673
tau1_squared[1]	Instrumentpiano	1.6657
tau1_squared[2]	Instrumentstring	3.4958
tau2_squared[1]	HarmonyI-V-IV	0.0476
tau2_squared[2]	HarmonyI-V-VI	1.6242
tau2_squared[3]	HarmonyIV-I-V	0.0063

As you can see, the all variances of the random effect for `Harmony` are less than the estimated residual variance. Originally, the model fitter wanted to prune `Harmony` completely at the last minute. However, I'm forcing all models to keep all three main experimental effects as fixed effects. And if, conditional on those terms, the fitter decided a random effect for `Harmony` was better than no random effect for `Harmony`, (though worse than no effects whatsoever for `Harmony`), the random effect stays.

A.2.3.3 Conditional Residual Plots of Personal Biases Model

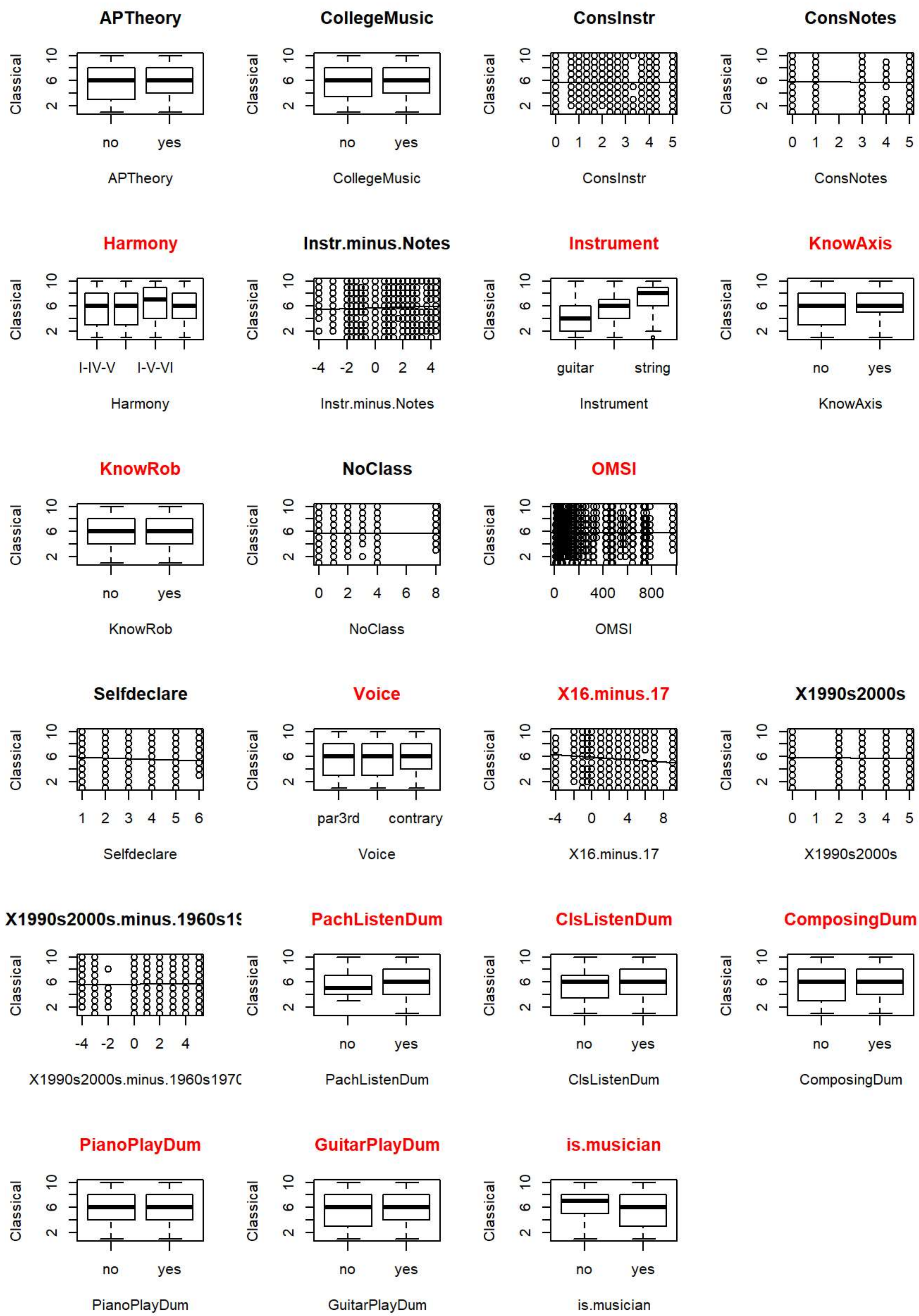
Most of the subjects' conditional residuals stay between the lines. There is clumping on some but not all plots. That means that some of the subjects' perceptions had more correlation with the experimental effects, while the opposite case was true for others.



A.2.4 EDA: Individual Covariates

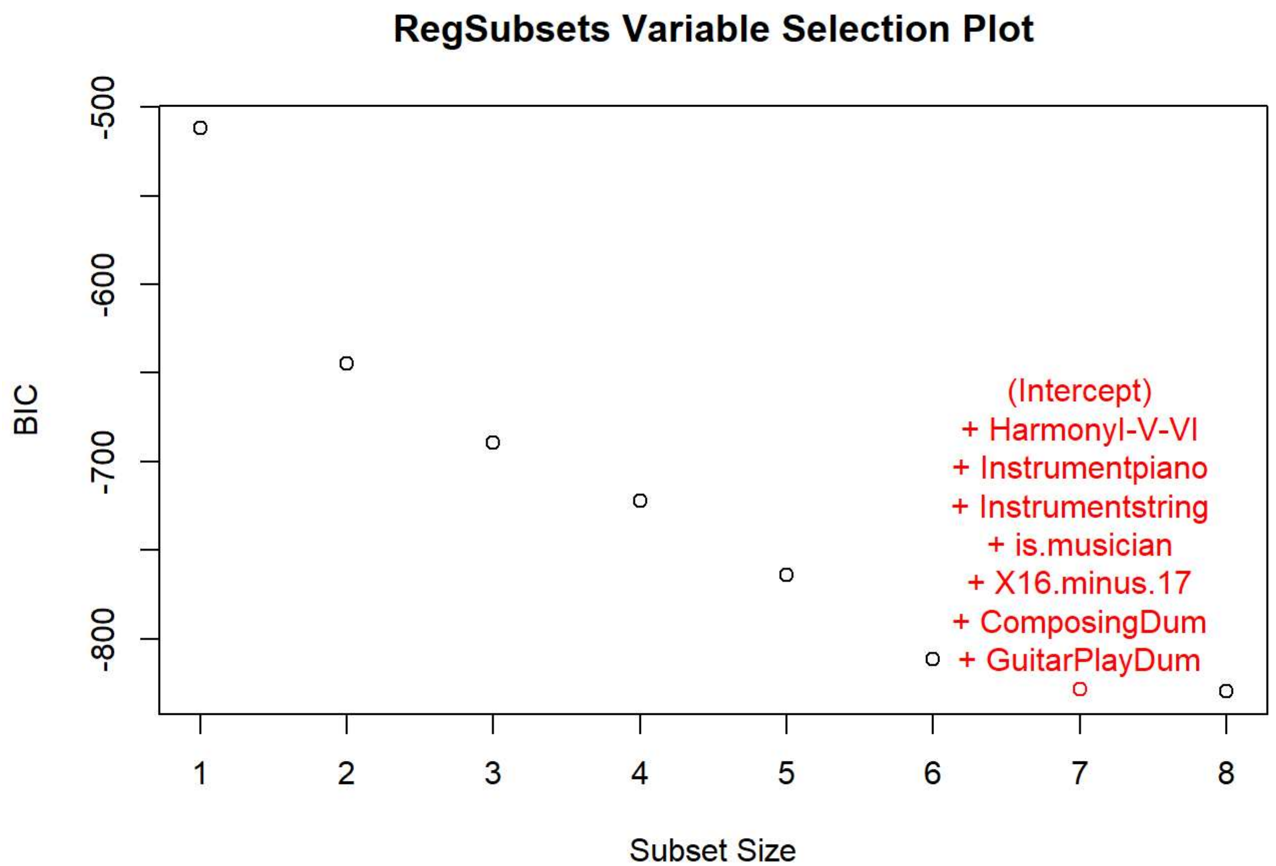
A.2.4.1 EDA: Individual Covariates vs Classical Plots

I started by plotting each individual variable against `Classical`. Then I marked (in red) the ones whose plots seemed to indicate a relationship with `Classical` of some sort, and other ones that I had special interest (also red).


APTheory
CollegeMusic
ConsInstr
ConsNotes
Harmony
Instr.minus.Notes
Instrument
KnowAxis
KnowRob
NoClass
OMSI
Selfdeclare
Voice
X16.minus.17
X1990s2000s
X1990s2000s.minus.1960s19
PachListenDum
ClsListenDum
ComposingDum
PianoPlayDum
GuitarPlayDum
is.musician
Classical
no
yes
I-IV-V
I-V-VI
guitar
string
par3rd
contrary
X1990s2000s.minus.1960s1970

A.2.4.2 EDA: Individual Covariates Conventional Linear Model BIC-Subset Plots

I began with the marked variables (main effects only) and used R's **leaps** library's *regsubsets()* function to find the model with the lowest BIC (up to 8 predictor variables). The plot only indicated three candidates for optimal models. The two least simple versions, and the second most simple model. The second least simple model (spelled out in red), with the addition of `Voice` as a predictor, is the one I've elected as a starting point for forming my final multi-level model.



A.2.4.3 EDA: Individual Covariates Repeated Measures Model Variable Selection

After using a combination of automatic variable selection (mostly the *fitLMER.fnc()* function), heuristic-based variable selection (mostly BIC), and personal judgement-based variable selection, I have decided on the following as a starting point, which I perform automated variable selection on:

```
icrmm0 <- lmer(Classical ~ Instrument * (ComposingDum + GuitarPlayDum) + Voice*ComposingDum +
                    Harmony * (is.musician + KnowAxis + KnowRob + ComposingDum) +
                    X16.minus.17 + (1 |Subject), method = "BIC",
                  data = dat, control = lmerControl(optimizer = "bobyqa"))
icrmm.fit <- bfFixefLMER_t.fnc(icrmm0)
```

```
formula(icrmm.fit)
```

```
## Classical ~ Instrument + ComposingDum + GuitarPlayDum + Voice +
##     Harmony + is.musician + KnowAxis + KnowRob + X16.minus.17 +
##     (1 | Subject) + Instrument:ComposingDum + Instrument:GuitarPlayDum +
##     Harmony:is.musician + Harmony:KnowAxis + Harmony:KnowRob +
##     ComposingDum:Harmony
```

A.2.4.4 EDA: Individual Covariates Personal Biases Model Variable Selection

I used R's **LMERConvenienceFunctions** library's *fitLMER.fnc()* function, again. Starting with the model from 2.4.4, above, I performed forwards selection of random effects and then backwards elimination of fixed effects. This resulted in the additions of `Instrument` and `Harmony` as random effects, followed by the removal of `ComposingDum` and `GuitarPlayDum` along with all interaction terms including them. I also used VIFs, AIC, and BIC comparisons to decide between the two best candidates.

```
# get that bad term out of the slope list
vars = c("Instrument","Harmony","Voice")
icpbm.fit <- fitLMER.fnc(icrmm.fit, ran.effects=
                              list(slopes=vars, by.vars="Subject",
                                   corr=rep(0,length(vars))), set.REML.FALSE = T)
formula(icpbm.fit)
# Classical ~ Instrument + Voice + Harmony + is.musician +
#     KnowRob + (1 | Subject) + (Instrument | Subject) + (Harmony |
#     Subject) + Harmony:is.musician + Harmony:KnowRob

icpbm.fit <- update(icpbm.fit,.~.-(1|Subject)-(Instrument|Subject)-(Harmony|Subject)+
                         (1 + Instrument + Harmony | Subject))
rm(vars)
```

A.2.4.5 EDA: Individual Covariates Personal Biases Model VIFs

Variance Inflation Factors raised a potential concern. A healthy relationship between all variables would result in a VIF table whose last column's values are all below 2. However, the simultaneous presence of both `Harmony` and interactions such as `Harmony:is.musician` jeopardize this (see below). Two new, VIF-rule-conforming models were formed by removing the interaction terms, one at a time, in order of VIF size. An anova test revealed for each interaction term that adding it to the model resulted in a statistically significant improvement (p-value < 0.05) and resulted in lower AIC.

```
##                         GVIF Df GVIF^(1/(2*Df))
## Harmony             92.57847  3        2.126923
## Harmony:is.musician 99.30188  3        2.151921
```

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
no.interaction	32	9904.637	10090.92	-4920.319	9840.637	NA	NA	NA
no.Mus.interaction	35	9900.579	10104.32	-4915.289	9830.579	10.05824	3	0.0180775
icpbm.fit	38	9892.222	10113.43	-4908.111	9816.222	14.35651	3	0.0024579

Since some multicolinearity is justified by the fact that interaction terms are, by definition, multicolinear with their coresponding main effects, I elected to pardon the high VIFs and favor the model with better heuristics.

A.2.5 Summary of Final Model for Classical

Note: I don't know why, but R has more than one different method for AIC calculation. Different functions that internally perform AIC calculations don't always use the same method. As a result, the AIC printed in the last row of the table above and the one shown in the summary below, though derived from the same model, are different.

```
## lmer(formula = Classical ~ Instrument + Voice + Harmony + is.musician +
##     KnowRob + (1 + Instrument + Harmony | Subject) + Harmony:is.musician +
##     Harmony:KnowRob, data = dat, REML = TRUE, control = lmerControl(optimizer = "bobyqa"),
##     method = "BIC")
##                            coef.est coef.se
## (Intercept)                 5.13     0.37
## Instrumentpiano             1.36     0.17
## Instrumentstring            3.11     0.24
## Voicepar3rd                -0.39     0.08
## Voicepar5th                -0.36     0.08
## HarmonyI-V-IV              -0.18     0.19
## HarmonyI-V-VI              -0.47     0.30
## HarmonyIV-I-V              -0.15     0.19
## is.musician                -0.99     0.40
## KnowRob                    -0.12     0.42
## HarmonyI-V-IV:is.musician   0.19     0.22
## HarmonyI-V-VI:is.musician   1.37     0.34
## HarmonyIV-I-V:is.musician   0.24     0.22
## HarmonyI-V-IV:KnowRob      -0.01     0.23
## HarmonyI-V-VI:KnowRob       0.93     0.36
## HarmonyIV-I-V:KnowRob       0.02     0.23
##
## Error terms:
##  Groups   Name             Std.Dev. Corr
##  Subject  (Intercept)      1.52
##           Instrumentpiano  1.29     -0.36
##           Instrumentstring 1.87     -0.55  0.66
##           HarmonyI-V-IV    0.25      0.72 -0.61 -0.47
##           HarmonyI-V-VI    1.09      0.11 -0.40 -0.55  0.15
##           HarmonyIV-I-V    0.11      0.37 -0.39 -0.09  0.03 -0.31
##  Residual                  1.55
## ---
## number of obs: 2493, groups: Subject, 70
## AIC = 9919.1, DIC = 9789.3
## deviance = 9816.2
```

```
## [1] "Sigma^2:"
```

```
## [1] 2.397771
```

A.2.6: The is.musician addition

When I performed an anova test to see if there were significant differences between the model with the the original, I found that the one with `is.musician` and its interaction with Harmony was strongly significantly better than the one without its interaction. But when the interaction isn't an option, the presence of `is.musician` as a main effect isn't statistically significantly. See the anova table below

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
wo_musician	34	9899.475	10097.40	-4915.738	9831.475	NA	NA	NA
wo_HM	35	9900.579	10104.32	-4915.289	9830.579	0.8961472	1	0.3438169
final.model	38	9892.222	10113.43	-4908.111	9816.222	14.3565139	3	0.0024579

A.3 Modeling Popular

A.3.1 EDA: Repeated Measures Model

Repeated Measures Model is Appropriate!

```
#repeated.measures.model
rmm.pop <- lmer(Popular ~ Instrument * Harmony * Instrument +
                                      (1|Subject), data = dat,
                                    control = lmerControl(optimizer = "bobyqa"))
## Check to see if (1|Subject) is needed, with a simpler
## call to exactRLRT (note the simpler call to exactRLRT when there is
## just one random effect in HA and no random effect in H0):

exactRLRT(rmm.pop)
## RLRT = 1333.4, p-value < 2.2e-16
# That's a good sign, we see significance!
```

I performed variable selection on the saturated fixed effect model three times, using three different methods (`t`, `AIC`, and `llrt`), and all three agreed on a repeated measures model with only one main experimental factor as a fixed effect: `Instrument`.

```
## EVERY METHOD COMPLETELY REMOVES HARMONY AND VOICE!

formula(rmm.pop.fit <- fitLMER.fnc(rmm.pop, method = "t"))
# Popular ~ Instrument + (1 | Subject)
formula(rmm.pop.fit <- fitLMER.fnc(rmm.pop, method = "AIC"))
# Popular ~ Instrument + (1 | Subject)
formula(rmm.pop.fit <- fitLMER.fnc(rmm.pop, method = "llrt"))
# Popular ~ Instrument + (1 | Subject)
```

A.3.2 EDA: Personal Biases Model

After re-inserting `Harmony` and `Voice` back into the model, and forward fitting random effects (using `llrt`), it was summarized below with the `display()` command. Given the fixed effects had to be what they were, the model decided to accompany them with two personal biases: `Instrument` and `Harmony`. This looks exactly the same as the `Classical` model, although, I had to be more hands on in the process to arrive here for `Popular` than `Classical`, because of how insignificant `Harmony` and `Voice` were as fixed effects for `Popular`.

```
# pop the other guys back in
rmm.pop.fit <- lmer(Popular ~ Instrument + (1|Subject) + Harmony + Voice,
                data = dat, control = lmerControl(optimizer = "bobyqa"))

vars = c("Instrument","Harmony","Voice")

# llrt method
re.fit.pop <- ffRanefLMER.fnc(rmm.pop.fit, ran.effects =
                            list(slopes=vars, by.vars="Subject",
                                 corr=rep(0,length(vars))))
formual(re.fit.pop)
# Popular ~ Instrument + (1 | Subject) + Harmony +
#     Voice + (Instrument | Subject) + (Harmony | Subject)

# straighten out the main effects
rsm.pop.llrt <- update(re.fit.pop,
                       .~. - ((1 | Subject) + (Instrument | Subject) + (Harmony | Subject)) +
                         (1 + Instrument + Harmony | Subject))
display(rsm.pop.llrt)
# lmer(formula = Popular ~ Instrument + Harmony + Voice + (1 +
#     Instrument + Harmony | Subject), data = dat, REML = TRUE,
#     control = lmerControl(optimizer = "bobyqa"))
#                  coef.est coef.se
# (Intercept)       6.29     0.25
# Instrumentpiano  -0.89     0.16
# Instrumentstring -2.41     0.24
# HarmonyI-V-IV    -0.03     0.09
# HarmonyI-V-VI    -0.20     0.14
# HarmonyIV-I-V    -0.17     0.11
# Voicepar5th       0.03     0.07
# Voicecontrary    -0.13     0.07
#
# Error terms:
#  Groups   Name             Std.Dev. Corr
#  Subject  (Intercept)      1.99
#           Instrumentpiano  1.20     -0.29
#           Instrumentstring 1.92     -0.48  0.75
#           HarmonyI-V-IV    0.34      0.26 -0.29 -0.26
#           HarmonyI-V-VI    0.90     -0.15 -0.18 -0.13 -0.45
#           HarmonyIV-I-V    0.52     -0.25 -0.10  0.04 -0.55 -0.27
#  Residual                  1.50
# ---
# number of obs: 2493, groups: Subject, 70
# AIC = 9832.1, DIC = 9732.5
# deviance = 9752.3

# clean up
rm(rmm.pop,re.fit.pop,rmm.pop.fit)
```

The model is written out in mathematical terms below. The table below shows the values of the various fixed effect parameters that number too many to fit into the model below.

Note: $\overrightarrow{1_{\square i}}$ is a vector with a 1 in the position corresponding to the level of $\square$ characteristic of $observation_i$ if not baseline, and 0s elsewhere.

$$Popular_i = \alpha_{0i} + \alpha_{1i} + \alpha_{2i} + \overrightarrow{\alpha_3}^T \overrightarrow{1_{Vi}} + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_0 + \eta_{0j}, \eta_{0j} \sim N(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \overrightarrow{\beta_1}^T \overrightarrow{1_{Ii}} + \overrightarrow{\eta_{1j[i]}}^T \overrightarrow{1_{Ii}}, \overrightarrow{\eta_{1j}} \sim N(\vec{0}, \overrightarrow{\tau_1^2})$$

$$\alpha_{2j[i]} = \overrightarrow{\beta_2}^T \overrightarrow{1_{Hi}} + \overrightarrow{\eta_{2j[i]}}^T \overrightarrow{1_{Hi}}, \overrightarrow{\eta_{2j}} \sim N(\vec{0}, \overrightarrow{\tau_2^2})$$

The values and standard deviations of the coefficient vectors, $\overrightarrow{\alpha_?}$ are inscribed in the table below.

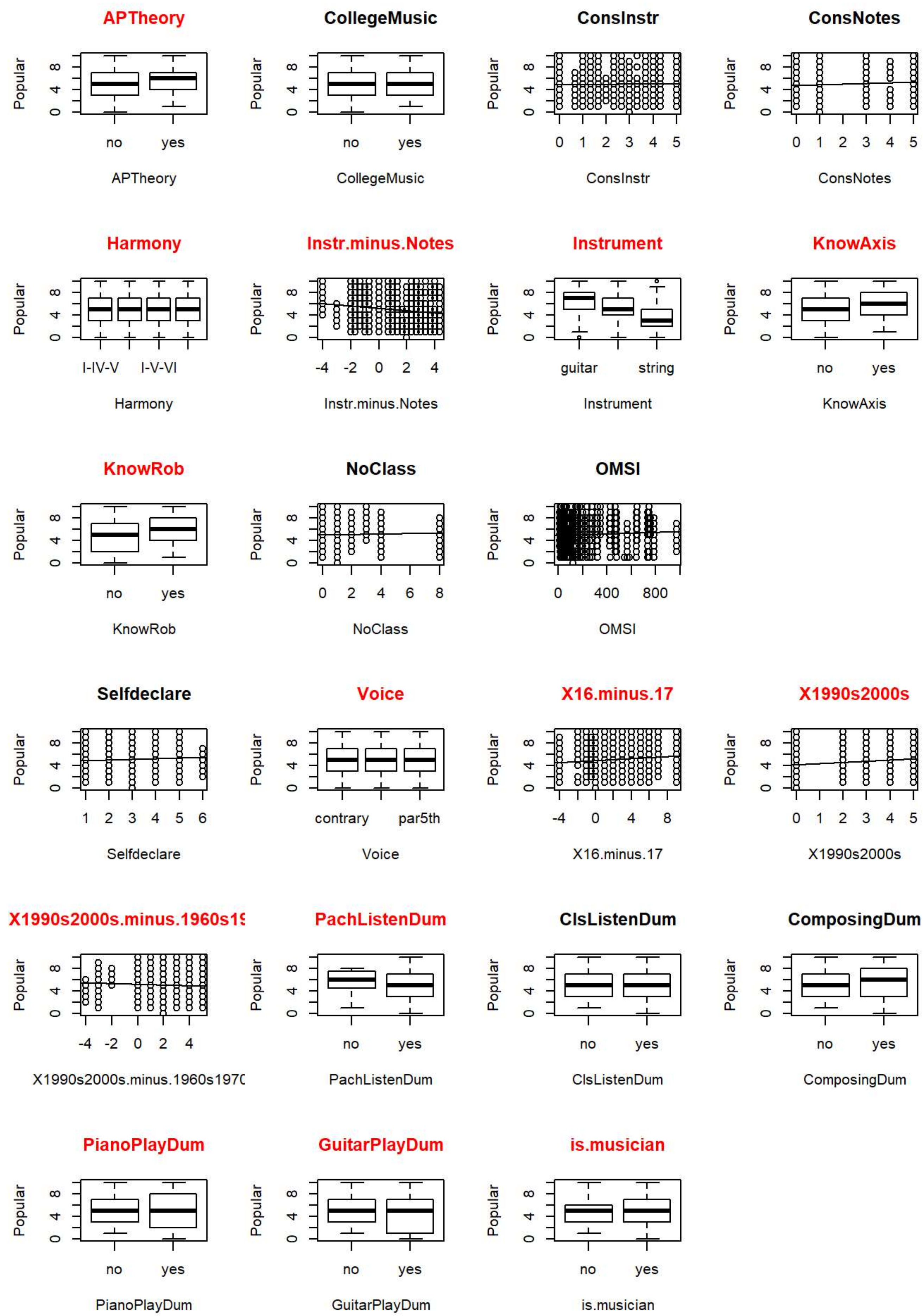
	factorLevel	Estimate	Std. Error
beta0	(Intercept)	6.1542	0.2525
beta1[1]	Instrumentpiano	-0.8882	0.1617
beta1[2]	Instrumentstring	-2.4092	0.2405
beta2[1]	HarmonyI-V-IV	-0.0329	0.0944
beta2[2]	HarmonyI-V-VI	-0.2036	0.1368
beta2[3]	HarmonyIV-I-V	-0.1720	0.1053
alpha3[1]	Voicepar3rd	0.1348	0.0736
alpha3[2]	Voicepar5th	0.1618	0.0736

	factorLevel	Estimate
sigma_squared	Residual	2.2499
tau0_squared	(Intercept)	3.9574
tau1_squared[1]	Instrumentpiano	1.4450
tau1_squared[2]	Instrumentstring	3.6698
tau2_squared[1]	HarmonyI-V-IV	0.1185
tau2_squared[2]	HarmonyI-V-VI	0.8037
tau2_squared[3]	HarmonyIV-I-V	0.2709

A.3.3 EDA: Individual Covariates

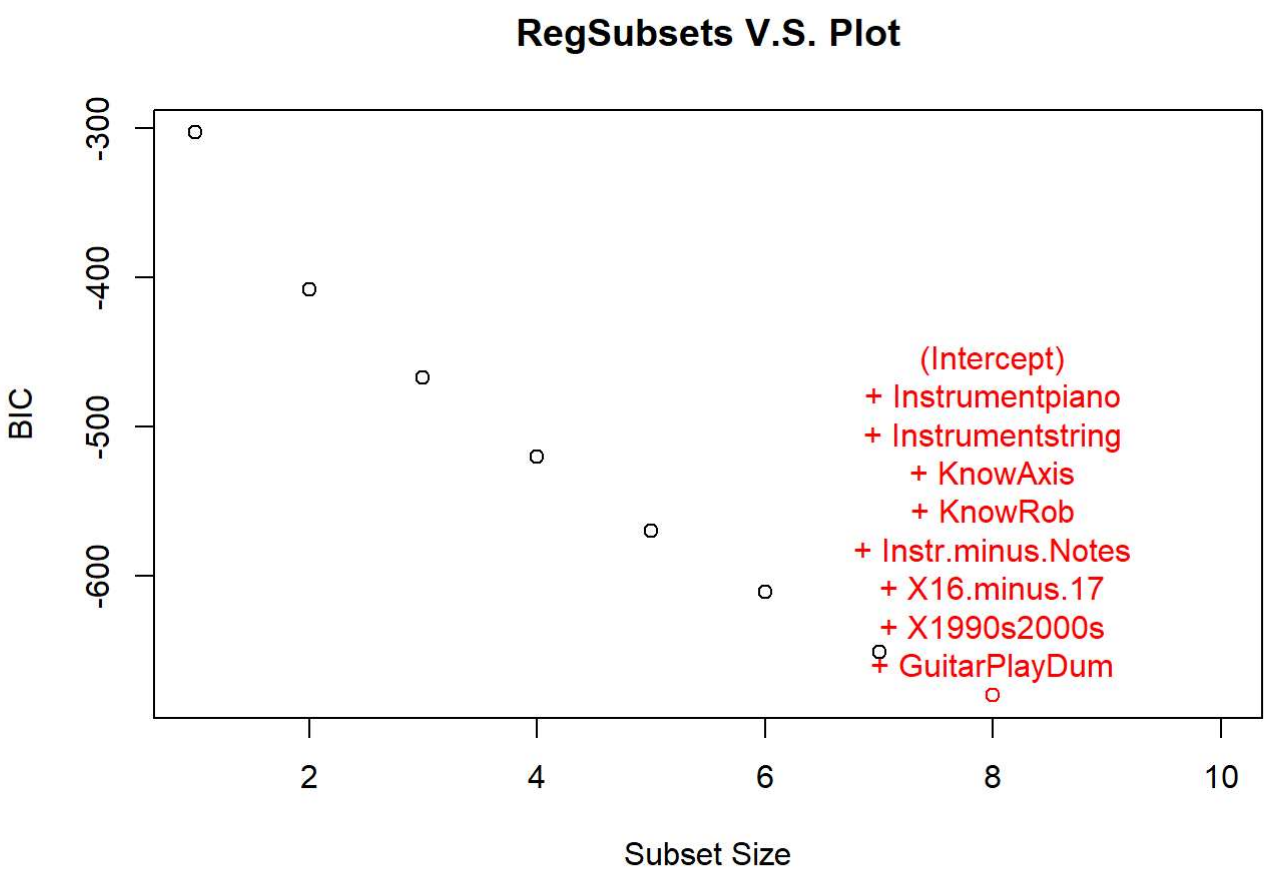
A.3.3.1 EDA: Individual Covariates vs Popular Plots

I started by plotting each individual variable against `Popular`. Then I marked (in red) the ones whose plots seemed to indicate a relationship with `Popular` of some sort, and other ones that I had special interest (also red).


APTheory
CollegeMusic
ConsInstr
ConsNotes
Harmony
Instr.minus.Notes
Instrument
KnowAxis
KnowRob
NoClass
OMSI
Selfdeclare
Voice
X16.minus.17
X1990s2000s
X1990s2000s.minus.1960s19
PachListenDum
ClsListenDum
ComposingDum
PianoPlayDum
GuitarPlayDum
is.musician
Popular
no
yes
I-IV-V
I-V-VI
guitar
string
contrary
par5th
X1990s2000s.minus.1960s197C

A.3.3.2 EDA: Individual Covariates Conventional Linear Model BIC-Subset Plots

I began with the marked variables (main effects only) and used R's **leaps** library's *regsubsets()* function to find the model with the lowest BIC (up to 8 predictor variables). The plot only indicated three candidates for optimal models. The two most simple versions, and the least simple model. The least simple model (spelled out in red), with the additions of `Voice` and `Harmony` as predictors, is the one I've elected as a starting point for forming my final multi-level model.



A.3.3.3 EDA: Individual Covariates Repeated Measures Model Variable Selection

I used R's **LMERConvenienceFunctions** library's *fitLMER.fnc()* function, again. I started by adding the individual covariates plus each two-way interaction between one of them and a main experimental factor. I performed forwards selection of random effects and then backwards elimination of fixed effects. This resulted in the additions of `Instrument` and `Harmony` as random effects, followed by the removal of `KnowAxis`, `Instr.minus.Notes`, `GuitarPlayDum`, and `is.musician` along with all interaction terms including them. A long and painful walk through my process is shown below.

```
popular.2 <- lmer(Popular ~ (Instrument + Harmony + Voice) *
                    (X16.minus.17 + Instr.minus.Notes + X1990s2000s + KnowAxis +
                    KnowRob + GuitarPlayDum + is.musician) + (1|Subject),
                  data = dat, control = lmerControl(optimizer = "bobyqa"))

## BACKWARD FIT FIXED EFFECTS

popular.BIC.2 <- bfFixefLMER_t.fnc(popular.2, method = "BIC")
formula(popular.BIC.2)
# Popular ~ Instrument + Instr.minus.Notes + (1 | Subject) + Instrument:Instr.minus.Notes

popular.AIC.2 <- bfFixefLMER_t.fnc(popular.2, method = "AIC")
formula(popular.AIC.2)
# Popular ~ Instrument + Harmony + X16.minus.17 + Instr.minus.Notes +
#     KnowAxis + KnowRob + is.musician + (1 | Subject) + Instrument:X16.minus.17 +
#     Instrument:Instr.minus.Notes + Instrument:is.musician + Harmony:KnowAxis +
#     Harmony:KnowRob + Harmony:is.musician

## FORWARD FIT RANDOM EFFECTS

vars <- c("Instrument","Harmony")
popular.AIC.2 <- ffRanefLMER.fnc(popular.AIC.2, ran.effects =
                              list(slopes=vars, by.vars="Subject",
                                   corr=rep(0,length(vars))))
display(popular.AIC.2)
# lmer(formula = Popular ~ Instrument + Harmony + X16.minus.17 +
#     Instr.minus.Notes + KnowAxis + KnowRob + is.musician + (1 |
#     Subject) + (Instrument | Subject) + (Harmony | Subject) +
#     Instrument:X16.minus.17 + Instrument:Instr.minus.Notes +
#     Instrument:is.musician + Harmony:KnowAxis + Harmony:KnowRob +
#     Harmony:is.musician, data = dat, REML = TRUE, control = lmerControl(optimizer = "bobyqa"))
#                                    coef.est coef.se
# (Intercept)                         5.24     0.56
# Instrumentpiano                    -0.38     0.36
# Instrumentstring                   -1.77     0.53
# HarmonyI-V-IV                      -0.36     0.20
# HarmonyI-V-VI                       0.28     0.27
# HarmonyIV-I-V                      -0.11     0.22
# X16.minus.17                        0.07     0.08
# Instr.minus.Notes                   0.02     0.14
# KnowAxis                            0.55     0.61
# KnowRob                             1.35     0.58
# is.musician                         0.66     0.60
# Instrumentpiano:X16.minus.17        0.03     0.05
# Instrumentstring:X16.minus.17       0.10     0.08
# Instrumentpiano:Instr.minus.Notes  -0.17     0.09
# Instrumentstring:Instr.minus.Notes -0.32     0.14
# Instrumentpiano:is.musician        -0.57     0.38
# Instrumentstring:is.musician       -0.77     0.56
# HarmonyI-V-IV:KnowAxis              0.14     0.27
# HarmonyI-V-VI:KnowAxis              0.78     0.38
# HarmonyIV-I-V:KnowAxis             -0.23     0.31
# HarmonyI-V-IV:KnowRob               0.01     0.26
```

```
# HarmonyI-V-VI:KnowRob             -0.94     0.36
# HarmonyIV-I-V:KnowRob             -0.18     0.30
# HarmonyI-V-IV:is.musician          0.39     0.23
# HarmonyI-V-VI:is.musician         -0.56     0.31
# HarmonyIV-I-V:is.musician          0.01     0.26
#
# Error terms:
#  Groups    Name             Std.Dev. Corr
#  Subject   (Intercept)      0.00
#  Subject.1 (Intercept)      1.17
#            Instrumentpiano  1.18     -0.57
#            Instrumentstring 1.84     -0.98  0.72
#  Subject.2 (Intercept)      1.58
#            HarmonyI-V-IV    0.30     -0.01
#            HarmonyI-V-VI    0.80     -0.14 -0.45
#            HarmonyIV-I-V    0.52     -0.22 -0.63 -0.34
#  Residual                   1.50
# ---
# number of obs: 2493, groups: Subject, 70
# AIC = 9844.7, DIC = 9679.1
# deviance = 9717.9

## BACKWARD FIT FIXED EFFECTS WITH RANDOM EFFECTS INCLUDED

popular.AIC.2.1 <- bfFixefLMER_t.fnc(popular.AIC.2, method = "AIC", prune.ranefs = F)
display(popular.AIC.2.1)
# lmer(formula = Popular ~ Instrument + KnowRob + (1 | Subject) +
#     (Instrument | Subject), data = data, REML = TRUE, control = lmerControl(optimizer = "bobyq
a"))
#                  coef.est coef.se
# (Intercept)       5.85     0.25
# Instrumentpiano  -0.89     0.16
# Instrumentstring -2.41     0.24
# KnowRob           1.50     0.48
#
# Error terms:
#  Groups    Name             Std.Dev. Corr
#  Subject   (Intercept)      1.21
#  Subject.1 (Intercept)      1.47
#            Instrumentpiano  1.18     -0.43
#            Instrumentstring 1.90     -0.73  0.76
#  Residual                   1.60
# ---
# number of obs: 2493, groups: Subject, 70
# AIC = 9910, DIC = 9876.7
# deviance = 9881.4

## method = LLRT
popular.AIC.2.2 <- bfFixefLMER_t.fnc(popular.AIC.2, method = "llrt")
display(popular.AIC.2.2)
# lmer(formula = Popular ~ Instrument + X16.minus.17 + KnowRob +
#     (1 | Subject) + (Instrument | Subject), data = data, REML = TRUE,
#     control = lmerControl(optimizer = "bobyqa"))
#                  coef.est coef.se
```

```
# (Intercept)       5.65     0.28
# Instrumentpiano  -0.89     0.16
# Instrumentstring -2.41     0.24
# X16.minus.17      0.11     0.06
# KnowRob           1.54     0.47
#
# Error terms:
#  Groups    Name             Std.Dev. Corr
#  Subject   (Intercept)      1.29
#  Subject.1 (Intercept)      1.42
#            Instrumentpiano  1.18     -0.46
#            Instrumentstring 1.90     -0.80  0.76
#  Residual                   1.60
# ---
# number of obs: 2493, groups: Subject, 70
# AIC = 9912.7, DIC = 9869.9
# deviance = 9878.3

# LLRT Better option for second round of backfitting

# add Voice and Harmony back in
final.pop <- lmer(formula = Popular ~ Instrument + Harmony + Voice +
                          X16.minus.17 + KnowRob +
                          (1 + Instrument + Harmony| Subject),
                        data = dat, REML = TRUE,
                        control = lmerControl(optimizer = "bobyqa"))
```

A.3.4 Summary of Final Model for Popular

```
## lmer(formula = Popular ~ Instrument + Harmony + Voice + X16.minus.17 +
##     KnowRob + (1 + Instrument + Harmony | Subject), data = dat,
##     REML = TRUE, control = lmerControl(optimizer = "bobyqa"))
##                  coef.est coef.se
## (Intercept)       5.65     0.29
## Instrumentpiano  -0.89     0.16
## Instrumentstring -2.41     0.24
## HarmonyI-V-IV    -0.03     0.09
## HarmonyI-V-VI    -0.20     0.14
## HarmonyIV-I-V    -0.17     0.11
## Voicepar3rd       0.14     0.07
## Voicepar5th       0.16     0.07
## X16.minus.17      0.13     0.06
## KnowRob           1.38     0.46
##
## Error terms:
##  Groups   Name             Std.Dev. Corr
##  Subject  (Intercept)      1.96
##           Instrumentpiano  1.20     -0.30
##           Instrumentstring 1.92     -0.56  0.75
##           HarmonyI-V-IV    0.34      0.21 -0.29 -0.26
##           HarmonyI-V-VI    0.90     -0.07 -0.18 -0.13 -0.45
##           HarmonyIV-I-V    0.52     -0.28 -0.10  0.04 -0.55 -0.27
##  Residual                  1.50
## ---
## number of obs: 2493, groups: Subject, 70
## AIC = 9828.3, DIC = 9717.5
## deviance = 9740.9
```