Do You Hear Classical Or Popular Music?

Ying Zhu
yingzhu2@andrew.cmu.edu

8 Dec 2019

Abstract

We mainly address the question of how experimental factors and individual covariates have influence on listeners identification of music as "classical" or "popular". We examine the data containing 36 musical stimuli presented to 70 listeners presented by Ivan Jimenez and Vincent Rossi in a designed experiment. Multilevel models are carefully selected through ANOVA tests, forward fitting and stepwise AIC/BIC. We find that for classical ratings, instrument exerts the strongest influence among the three design factors; I-V-vi and contrary motion have the strongest association among all levels. Among individual covariates, *ClsListen* drives classical ratings, while *X16.minus.17*, *Composing* and *KnowRob* drives popular ratings. *Harmony I-V-VI* and *GuitarPlay* indicate the difference that musicians and non-musicians identify classical music. In the future, more rigorous way to clean the data and to classify musicians and non-musicians may yield more reliable results.

1 Introduction

Music makes people feel relaxed. However, it might not be an easy thing to distinguish classical music from popular one. Listeners identification of music as "classical" or "popular" can be influenced by some music features such as instrument, harmonic motion, and voice leading. The main purpose for this study is to find what experimental factor, or combination of factors, has the strongest inuence on ratings?

There are several research questions to address in order to fully investigate the data:

- Does Instrument exert the strongest inuence among the three design factors (Instrument, Harmonic Motion,Voice Leading), as the researchers suspect?
- Among the levels of Harmonic Motion does I-V-vi have a strong association (the strongest?) with classical ratings? Does it seem to matter whether the respondent is familiar with one or the other (or both) of the Pachelbel rants/comedy bits?
- Among the levels of Voice Leading, does contrary motion have a strong (the strongest?) association with classical ratings?

In addition to answering the main question posed above, we will address the following questions:

- Are there differences in the way that musicians and non-musicians identify classical music?
- Are there differences in the things that drive classical, vs. popular, ratings?

2 Methods

The data was collected by Ivan Jimenez, a composer and musicologist visiting the University of Pittsburgh, and student Vincent Rossi in a designed experiment in 2012.

In summary, data are available on the following variables:

```
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)
X1stInstr                        How proficient are you at your first
                                     musical instrument (0-5, 0=not at all)
X2ndInstr                        Same, for second musical instrument
first12                          In the experiment, which instrument was
                                 presented to the subject in the first
                                 12 stimuli? (IGNORE FOR THIS ASSIGNMENT.)
```

Figure 1: Variable description

According to summary of the raw dataset, we have 15 variables containing NAs, which are *ConsNotes*, *PachListen*, *ClsListen*, *KnowRob*, *KnowAxis*, *X1990s2000s*, *X1990s2000s.minus.1960s1970s*, *CollegeMusic*, *NoClass*, *APTheory*, *Composing*, *X1stInstr*, *X2ndInstr*, *Classical*, *Popular*. The dependent variables *Classical* and *Popular* have 27 NAs, which is an extremely small proportion out of the whole dataset, so we simply delete the rows with NA in the variables *Classical* and *Popular*. We also find that more than half of the entries in the variables *X1stInstr* and *X2ndInstr* are NAs. We do not think any kind of imputation will make sense due to the little information contained in these two variables and imputation would cause large bias. So we discard these two columns. There are also some variables with miscoded data. For example, *Classical* and *Popular* have values 1 to 10, so we simply delete the observations with value larger than 10 or smaller than 1. Here we treat the variables whose values contain only two levels as factor variables. The only two factor variables are *APTheory* and *CollegeMusic*, and all the others are numeric. The only continuous variable is *OMSI* and all the other numeric variables are discrete. We impute the discrete variables with the mode value of non-missing values.

With the cleaned data, we do some exploratory data analysis. we use correlation plot and boxplots to explore the relationship between each variable pair. Also, qqplot is used to examine the normality of continuous variables and log-transformation is performed if helps. To achieve the best models for ratings, we start with a conventional linear model and then step into multilevel models. Fixed effects are forward selected by hand through ANOVA tests and backward stepwise AIC/BIC. We decide whether the random effects should be included by comparing AIC, BIC, likelihood ratio test or forwardfitting automatically.

3 Results

3.1 What experimental factor, or combination of factors, has the strongest inuence on ratings?

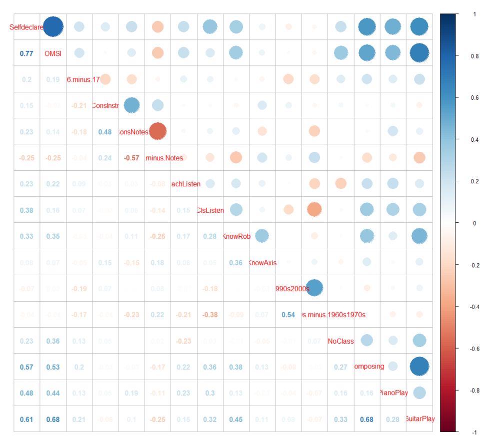


Figure 2: Correlation plot between each numerical variable pair

Before fitting the models, it is necessary to assess whether the original dataset meets modeling assumptions. We perform some exploratory data analysis. We inspect the only continuous variable *OMSI* using QQ plot and histogram. The result shows that the data is heavily skewed, so we perform log-transformation on *OMSI*. After transformation, the distribution of QQ plot and histogram look approximately normal. (For more details about data transformation, please refer to the Appendix on page 4-5.) In addition, we make a correlation plot for numerical variables to explore the relationship between each variable pair. From Figure 2, we find no numerical variable pairs have correlation coefficient greater than 0.8, which means none of the variables is strongly correlated with others. To analyze the relationship between the categorical variables and numerical variables, we make boxplots with *APTheory* or *CollegeMusic* on X axis and other variables on Y axis. From Figure 4

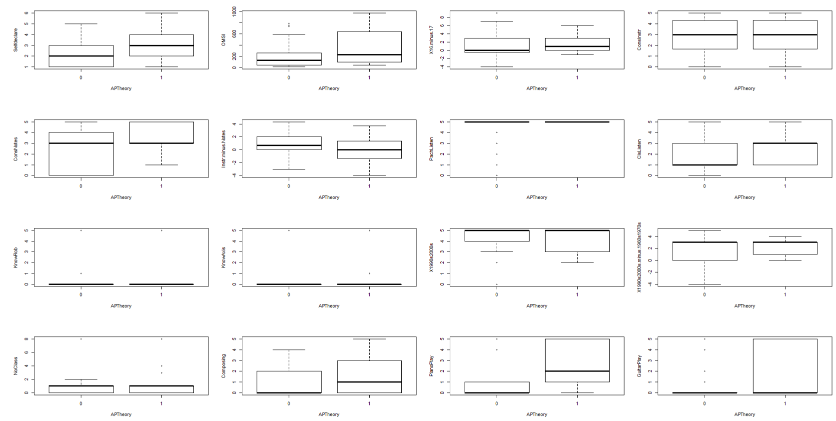


Figure 3: Boxplots between APTheory and numerical variables

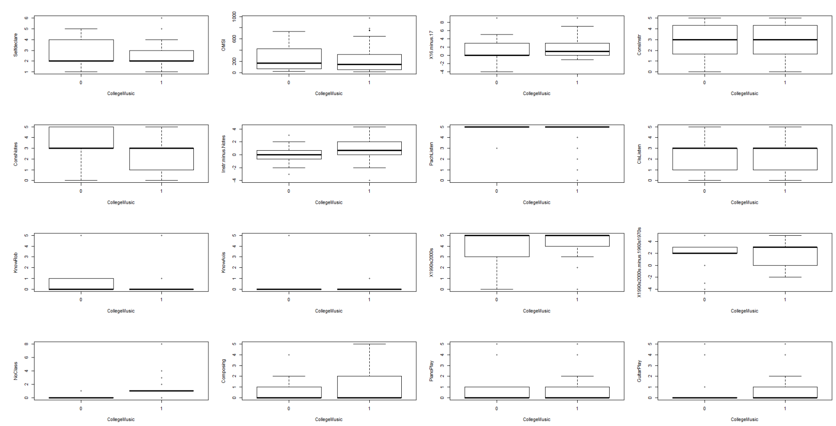


Figure 4: Boxplots between CollegeMusic and numerical variables

and 3, we find *CollegeMusic* and *NoClass* may have strong correlation. However, according to their definition, one can not replace the other, so we keep both of them in the dataset.

we start with fitting conventional linear regression model on Classical and Popular from Instrument, Harmony and Voice to examine the influence of the three main experimental factors (Instrument, Harmony&Voice) on ratings. Since there are only three factors here, it is worth considering interactions of all orders. Therefore, we first include interactions of all orders of the three main effects: Instrument, Harmony and Voice in the conventional linear model. Then we perform stepwise AIC to select interaction terms. Finally, we find that the interaction term *Harmony:Voice* should be included. (For more details about stepAIC, please refer to the Appendix on page 6.)

It can be inferred from summary of the fitted conventional linear model that only the terms *HarmonyI-V-VI*, *Instrumentpiano*, *Instrumentstring* and *HarmonyI-V-VI:Voicepar3rd* are significant under 95% confidence interval. (For more details about model summary, please refer to the Appendix on page 6-7.)

By looking into the diagnostic plots Figure 5, we can tell that the assumption of the normality holds since the points in QQ-plot follow the line pretty well. When we look into the Residuals vs

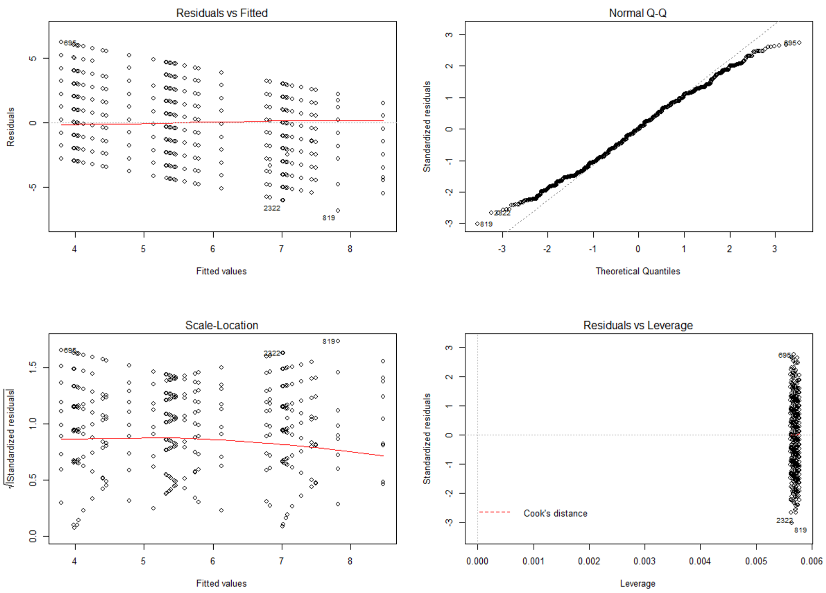


Figure 5: Diagnostic plot

Leverage plot, we find there is no high leverage points. However, the Scale Location plot shows a distinct pattern, which indicates the assumption that the random disturbances have equal variance may not be satisfied, which means the data has obvious heteroscedasticity. Therefore, we further step into fitting multilevel models.

Considering that perhaps person A is more inclined to rate everything as classical, and person B is more inclined to rate everything as popular. This can be accounted for by the random intercept. Then, to avoid "personal biases" in ratings, we fit a standard repeated measures model by adding random intercepts to our conventional linear model, and compare the two models using AIC and BIC. The standard repeated measures model has lower AIC and BIC compared to the previous conventional linear model. Thus, we include the random intercept in the model. (For details about the result of AIC and BIC comparison, please refer to the Appendix on page 8.)

After that, we look at the conditional residuals to assess the current model. As shown in Figure 6, the residual plot does not look good for some subjects. For example, Subject 24 and 39 display distinct decreasing patterns as the fitted value increases. Subject 42 and 48 have conditional residuals far from zero. Therefore, we should further improve the model by adjusting fixed effects and considering whether there should be other random effects.

Perhaps personal biases vary with the type of instrument, type of harmony, and/or type of voice leading. For example, perhaps people vary in the degree to which they are inclined to call music played by a string quartet classical. Therefore, we further consider adding random effects of the three main effects to the previous model by forward-fitting random effects. We assume that the random effect terms of these three main effects have correlations with each other. The result shows that random slope of *Instrument* and *Harmony* need to be included.

Other than the three experimental factors, there may be other factors matter a lot to the ratings, so we improve the model by deciding which individual covariates should be added to the model as fixed effects. We first use ANOVA test to forward select fixed effects by hand. We add five fixed

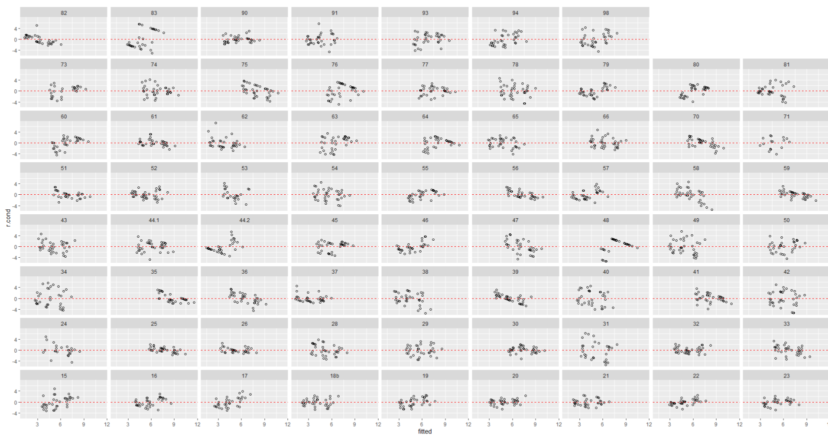


Figure 6: Conditional residual plot after adding random intercept

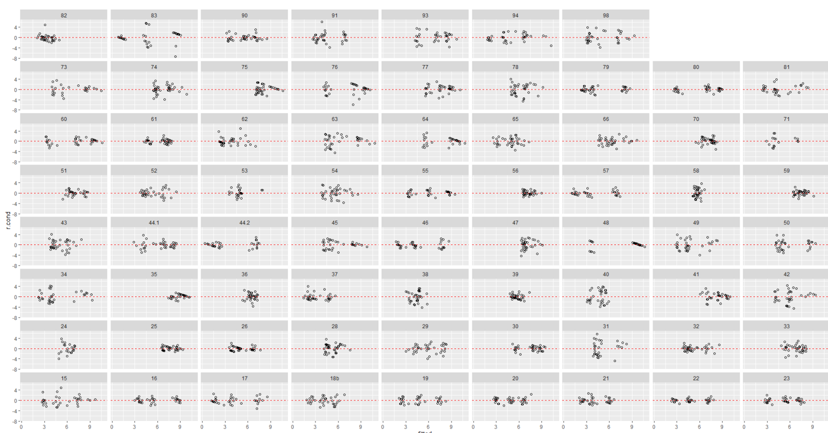


Figure 7: Conditional residual plot for the final model

effects to the previous model: *ClsListen*, *X16.minus.17*, *APTheory*. (For details about the steps of ANOVA tests, please refer to the Appendix on page 13-15.) Then we automatically add fixed effects by applying stepAIC method to the conventional linear model having only the variables of fixed effects in the previous model. The result shows that it adds *Selfdeclare*, *log OMSI*, *X16.minus.17*, *Instr.minus.Notes*, *ClsListen*, *X1990s2000s*, *X1990s2000s.minus.1960s1970s*, *NoClass*, *APTheory*, *Composing* and *PianoPlay*. (For details about the process of automatic selection, please refer to the Appendix on page 17-18.) We find that the two models selected by hand and automatically are nested models, so we apply ANOVA test to decide which model to choose. The test result shows there is no significant difference between the two models. (For details about the result of ANOVA test, please refer to the Appendix on page 18-19.) So finally, we choose the model adding three fixed effects: *ClsListen*, *X16.minus.17*, *APTheory*.

Once the fixed effects are settled, we go back and check to see whether there should be any change in the random effects by forward-fitting the random effect structure. It shows no additional random effects need to be added. Last, we compare current the model with the model without adding fixed effects. The current model has lower AIC, so we can say the model has been improved

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Harmony |
    Subject) + (Instrument - 1 | Subject) + Harmony:Voice + ClsListen +
    X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s +      APTheory.fc - 1
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9683.9   9898.7  -4805.0   9609.9     2416

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.8291 -0.5827  0.0238  0.5720  4.0561

Random effects:
 Groups    Name              Variance  Std.Dev.  Corr
 Subject   (Intercept)       1.363e-09 3.692e-05
 Subject.1 (Intercept)       1.680e+00 1.296e+00
           HarmonyI-V-IV     4.294e-02 2.072e-01  0.50
           HarmonyI-V-VI     1.607e+00 1.268e+00 -0.59  0.33
           HarmonyIV-I-V     3.070e-03 5.540e-02  0.64  0.69 -0.31
 Subject.2 Instrumentguitar 1.298e+00 1.139e+00
           Instrumentpiano  1.201e+00 1.096e+00  0.34
           Instrumentstring 4.959e-01 7.042e-01 -0.96 -0.07
 Residual                    2.323e+00 1.524e+00
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                             Estimate Std. Error t value
Instrumentguitar              4.33162    0.51788   8.364
Instrumentpiano               5.66989    0.51664  10.974
Instrumentstring              7.38468    0.50760  14.548
Voicepar3rd                  -0.23379    0.15014  -1.557
Voicepar5th                  -0.20092    0.15120  -1.329
HarmonyI-V-IV                 0.14533    0.15289   0.951
HarmonyI-V-VI                 1.19158    0.21410   5.566
HarmonyIV-I-V                -0.12756    0.15038  -0.848
ClsListen                     0.25666    0.09324   2.753
X16.minus.17                 -0.07261    0.04590  -1.582
X1990s2000s                  -0.23896    0.10476  -2.281
X1990s2000s.minus.1960s1970s  0.16542    0.09084   1.821
APTheory.fc1                  0.56990    0.33232   1.715
Voicepar3rd:HarmonyI-V-IV    -0.34514    0.21319  -1.619
Voicepar5th:HarmonyI-V-IV    -0.20823    0.21381  -0.974
Voicepar3rd:HarmonyI-V-VI    -0.76420    0.21326  -3.583
Voicepar5th:HarmonyI-V-VI    -0.47531    0.21389  -2.222
Voicepar3rd:HarmonyIV-I-V     0.50026    0.21300   2.349
Voicepar5th:HarmonyIV-I-V     0.06373    0.21300   0.299
convergence code: 0
boundary (singular) fit: see ?isSingular
```

Figure 8: Summary of final model for Classical rating

after adding individual covariates. (For more details about the model comparison, please refer to the Appendix on page 19.) Thus, for classical ratings, our final model includes *Harmony*, *Instrument*, *Voice*, *ClsListen*, *X16.minus.17*, *APTheory*, interaction of *Harmony* and *Voice* with random effect from intercept, *Harmony* and *Instrument*. We plot the conditional residual plot again. It can be inferred from Figure 7 that the conditional residual plot looks good for all subjects. There is no distinct patterns and residuals are close to zero.

We repeat the process and fit a multilevel model for Popular ratings following the same steps. (For details about the process of reaching the final model, please refer to the Appendix on page 24-36.) As a result, for popular ratings, our final model includes *Harmony*, *Instrument*, *Voice*, *Composing*, *X16.minus.17*, *KnowRob*, *X1990s2000s*, *GuitarPlay*, with random effect from intercept, *Harmony* and *Instrument*.

To determine whether *Instrument* exerts the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading), as the researchers suspect, we look at the summary tables of all levels of estimations. (For details about how we get the summary tables of all levels of estimations, please refer to the Appendix on page 22-26 & 40-43.)

- For classical ratings, it can be inferred from Figure 8 that all the levels of *Instrument* have t value greater than 1.96, which means all the levels have significant influence on classical

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Harmony |
    Subject) + (Instrument - 1 | Subject) + Composing + X16.minus.17 +      KnowRob + X1990s2000s + GuitarPlay - 1
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9726.2   9906.1  -4832.1   9664.2     2422

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.0570 -0.5923  0.0293  0.5882  3.3093

Random effects:
 Groups    Name              Variance Std.Dev. Corr
 Subject   (Intercept)       0.08041  0.2836
 Subject.1 (Intercept)       0.86320  0.9291
           HarmonyI-V-IV     0.08319  0.2884    0.58
           HarmonyI-V-VI     0.86398  0.9295   -0.24 -0.43
           HarmonyIV-I-V     0.19911  0.4462   -0.37 -0.66 -0.39
 Subject.2 Instrumentguitar 0.60877  0.7802
           Instrumentpiano  1.16463  1.0792    0.24
           Instrumentstring 1.50517  1.2269   -0.61  0.43
 Residual                    2.39489  1.5475
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                 Estimate Std. Error t value
Instrumentguitar  5.51842    0.41658  13.247
Instrumentpiano   4.56795    0.42644  10.712
Instrumentstring  2.97058    0.43244   6.869
HarmonyI-V-IV    -0.04846    0.09501  -0.510
HarmonyI-V-VI    -0.30122    0.14213  -2.119
HarmonyIV-I-V    -0.23238    0.10320  -2.252
Voicepar3rd       0.14361    0.07664   1.874
Voicepar5th       0.16737    0.07660   2.185
Composing         0.26561    0.11745   2.261
X16.minus.17      0.11376    0.04403   2.583
KnowRob           0.18401    0.08499   2.165
X1990s2000s       0.16525    0.08378   1.972
GuitarPlay       -0.24344    0.12157  -2.002
convergence code: 1
boundary (singular) fit: see ?isSingular
```

Figure 9: Summary of final model for Popular rating

ratings. In order to compare the effect size of the three design factors, we look at Figure 8, Figure 10 and Figure 11. When we look into the estimated coefficients of fixed effects, *Instrument* has the largest effect size among the three experimental factors. The occurrence of instrument Electric Guitar will increase the predicted Classical rating by 4.33 unit. The occurrence of instrument Piano will increase the predicted Classical rating by 5.70 unit. The occurrence of instrument String Quartet will increase the predicted Classical rating by 7.38 unit. In addition, when we look into the random effect table, we find the standard deviation of the all levels of *Instrument* are small enough so that coefficients of all levels are still positive within 2 standard deviations from fixed effect estimation, which means for more than 95% respondents, all levels of *Instrument* exert positive influence. In conclusion, *Instrument* exerts the strongest influence on Classical ratings among the three design factors.

- For Popular ratings, it can be inferred from Figure 9 that all the levels of *Instrument* have t value greater than 1.96, which means all the levels have significant influence on popular ratings. Similarly, we compare Figure 9, Figure 12 and Figure 13 to decide whether *Instrument* has the largest effect size. When we look into the estimated coefficients of fixed effects, levels of *Instrument* has coefficients 5.52, 4.57, 2.97, respectively. However, all the levels of *Harmony* has coefficients larger than 5. Therefore, we can not say that *Instrument* exerts the strongest influence on Popular ratings among the three design factors.

To determine whether *I-V-vi* has a strong association (the strongest?) with classical ratings among the levels of *Harmonic Motion*, we look at the fixed effects table for each level of Harmonic Motion for Classical rating. We find that all levels of *Harmonic Motion* have t value greater than 1.96, which means all levels have strong association with Classical rating. When we look into the estimated coefficients, *I-V-vi* has the largest effect size among the levels of *Harmonic Motion*. The occurrence of *I-V-vi* will increase the predicted Classical rating by 4.99 unit. In addition, referring to the random effect table, *I-V-vi* has the smallest variability within group, which means the majority

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Classical ~ Harmony + Voice + Instrument + (1 | Subject) + (Harmony -
    1 | Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +      X16.minus.17 + APTheory.fc - 1
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9684.8   9888.0  -4807.4   9614.8     2418

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.8300 -0.5793  0.0238  0.5765  4.0467

Random effects:
 Groups    Name            Variance  Std.Dev.  Corr
 Subject   (Intercept)     1.544e-08 0.0001243
 Subject.1 HarmonyI-IV-V   1.731e+00 1.3155627
           HarmonyI-V-IV   1.983e+00 1.4081325 0.99
           HarmonyI-V-VI   1.525e+00 1.2350658 0.51  0.58
           HarmonyIV-I-V   1.738e+00 1.3183990 1.00  0.99  0.49
 Subject.2 (Intercept)     1.334e+00 1.1549491
           Instrumentpiano 1.654e+00 1.2861960 -0.66
           Instrumentstring 3.336e+00 1.8265713 -1.00  0.68
 Residual                  2.322e+00 1.5238935
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                          Estimate Std. Error t value
HarmonyI-IV-V              3.79984    0.31535  12.050
HarmonyI-V-IV              3.94497    0.32123  12.281
HarmonyI-V-VI              4.99198    0.31073  16.065
HarmonyIV-I-V              3.67182    0.31534  11.644
Voicepar3rd               -0.23403    0.15012  -1.559
Voicepar5th               -0.20128    0.15118  -1.331
Instrumentpiano            1.33843    0.17152   7.803
Instrumentstring           3.05373    0.23135  13.200
ClsListen                  0.21309    0.08902   2.394
X16.minus.17              -0.07380    0.04659  -1.584
APTheory.fc1               0.52331    0.34252   1.528
HarmonyI-V-IV:Voicepar3rd -0.34459    0.21316  -1.617
HarmonyI-V-VI:Voicepar3rd -0.76397    0.21324  -3.583
HarmonyIV-I-V:Voicepar3rd  0.50067    0.21297   2.351
HarmonyI-V-IV:Voicepar5th -0.20793    0.21378  -0.973
HarmonyI-V-VI:Voicepar5th -0.47521    0.21387  -2.222
HarmonyIV-I-V:Voicepar5th  0.06482    0.21297   0.304
```

Figure 10: Fixed effects estimates for each level of Harmonic Motion for Classical rating

are scattered round the relatively high estimated mean. Therefore, we can conclude that *I-V-vi* has the strongest association with classical ratings among the levels of *Harmonic Motion*.

To determine if it matters whether the respondent is familiar with one or the other (or both) of the Pachelbel rants/comedy bits, we need to know if interaction of *HarmonyI-V-vi* and *KnowRob* or *KnowAxis* has strong association with classical ratings. First, we update our current model for classical ratings with the interaction term of *Harmony* and *KnowRob*. As indicated in Figure 14, the interaction of *HarmonyI-V-vi* and *KnowRob* has t value greater than 2, which shows that whether the respondent is familiar with the Pachelbel rants matters to the association between *HarmonyI-V-vi* and classical ratings. We also compare AIC between the two models with and without the interaction term of *HarmonyI-V-vi* and *KnowRob*. Since the model with the interaction has lower AIC, we hold the interaction term. Next, we further decide whether interaction of *HarmonyI-V-vi* and *KnowAxis* should be added in the same way. The result shows that t value is not significant and AIC increases. Therefore, we drop the interaction term. We can conclude that whether the respondent is familiar with the comedy bits does not matter to the association between *HarmonyI-V-vi* and classical ratings. (For more details about adding the two interactions, please refer to the Appendix on page X.) We can refer from Figure 14 that the more one have heard Rob Paravonians Pachelbel Rant, the greater influence of *HarmonyI-V-vi* has on classical ratings.

To determine whether *contrary motion* have a strong (the strongest?) association with classical ratings among the levels of *Voice Leading*, we look at the fixed effects table for each level of Voice Leading for Classical rating. We find that all levels of *Voice Leading* have t value greater than 1.96, which means all levels have strong association with Classical rating. When we look into the estimated coefficients, *contrary motion* has the largest effect size among the levels of *Voice Leading*. The occurrence of *contrary motion* will increase the predicted Classical rating by 3.80 unit. Therefore, we can conclude that *contrary motion* has the strongest association with classical

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Classical ~ Voice + Harmony + Instrument + (1 | Subject) + (Harmony |
    Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +      X16.minus.17 + APTheory.fc - 1
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9684.8   9888.0  -4807.4   9614.8     2418

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.8300 -0.5793  0.0238  0.5765  4.0467

Random effects:
 Groups    Name            Variance Std.Dev. Corr
 Subject   (Intercept)     0.000000 0.00000
 Subject.1 (Intercept)     1.730704 1.31556
           HarmonyI-V-IV   0.043910 0.20955   0.38
           HarmonyI-V-VI   1.607069 1.26770  -0.54  0.32
           HarmonyIV-I-V   0.005513 0.07425   0.01  0.51 -0.21
 Subject.2 (Intercept)     1.333907 1.15495
           Instrumentpiano  1.654300 1.28620  -0.66
           Instrumentstring 3.336362 1.82657  -1.00  0.68
 Residual                  2.322252 1.52389
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                          Estimate Std. Error t value
Voicecontrary              3.79984    0.31535  12.050
Voicepar3rd                3.56581    0.31514  11.315
Voicepar5th                3.59856    0.31572  11.398
HarmonyI-V-IV              0.14513    0.15292   0.949
HarmonyI-V-VI              1.19214    0.21409   5.568
HarmonyIV-I-V             -0.12802    0.15048  -0.851
Instrumentpiano            1.33843    0.17152   7.803
Instrumentstring           3.05373    0.23135  13.200
ClsListen                  0.21309    0.08902   2.394
X16.minus.17              -0.07380    0.04659  -1.584
APTheory.fc1               0.52331    0.34252   1.528
Voicepar3rd:HarmonyI-V-IV -0.34459    0.21316  -1.617
Voicepar5th:HarmonyI-V-IV -0.20793    0.21378  -0.973
Voicepar3rd:HarmonyI-V-VI -0.76397    0.21324  -3.583
Voicepar5th:HarmonyI-V-VI -0.47521    0.21387  -2.222
Voicepar3rd:HarmonyIV-I-V  0.50067    0.21297   2.351
Voicepar5th:HarmonyIV-I-V  0.06482    0.21297   0.304
```

Figure 11: Fixed effects estimates for each level of Voice Leading for Classical rating

ratings among the levels of *Voice Leading*.

3.2 Are there differences in the way that musicians and non-musicians identify classical music?

One of the secondary hypotheses of the researchers is that people who self-identify as musicians may be influenced by things that do not influence non-musicians. We dichotomize *Selfdeclare* ("are you a musician?") so that about half the participants are categorized as self-declared musicians, and half not. Then we are to find differences in the way that musicians and non-musicians identify classical music by interacting the dichotomized musician variable with other predictors. Similar to how we reached the final best model in the last section, we first determine which individual covariates should be added to the model as fixed effects. Considering that the number of variables is not small, we apply an automatic selection method, backward stepwise BIC, which penalizes model complexity more heavily than AIC and results in a smaller model that interprets better. After selecting individual covariates, we consider whether random intercept helps using likelihood ratio test since the random intercept in a repeated measures model can account for personal biases in ratings. Given the extremely small p-value in the likelihood ratio test, we include the random intercept in our model. Then we re-examine the fixed effects and drop those are not statistically significant. Next, we go back and check whether there should be any change in the random effects by forwardfitting random effects by way of log-likelihood ratio testing. (For more details about the process that leads to our final model, please refer to the Appendix on page 43-50.) We end up with a model includes *ConsInstr*, *Instr.minus.Notes*, *ClsListen*, *CollegeMusic*, *APTheory*, *PianoPlay*, *GuitarPlay*, interaction of *Harmony* and *Selfdeclare*, interaction of *X16.minus.17* and *Selfdeclare*, interaction of *CollegeMusic* and *Selfdeclare*, interaction of *GuitarPlay* and *Selfdeclare* with random

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Popular ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
    Subject) + (Instrument | Subject) + Composing + X16.minus.17 +      KnowRob + X1990s2000s + GuitarPlay - 1
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9726.2   9906.1  -4832.1   9664.2     2422

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.0597 -0.5935  0.0281  0.5885  3.3094

Random effects:
 Groups    Name            Variance Std.Dev. Corr
 Subject   (Intercept)     0.25645  0.5064
 Subject.1 (Intercept)     0.27728  0.5266
           HarmonyI-V-IV   0.08388  0.2896    1.00
           HarmonyI-V-VI   0.86317  0.9291   -0.43 -0.43
           HarmonyIV-I-V   0.19836  0.4454   -0.66 -0.66 -0.40
 Subject.2 (Intercept)     1.02435  1.0121
           Instrumentpiano 1.37471  1.1725   -0.35
           Instrumentstring 3.29114 1.8141   -0.65  0.74
 Residual                  2.39490  1.5475
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                 Estimate Std. Error t value
HarmonyI-IV-V     5.52127    0.41668  13.251
HarmonyI-V-IV     5.47287    0.42346  12.924
HarmonyI-V-VI     5.22008    0.42428  12.303
HarmonyIV-I-V     5.28893    0.41461  12.756
Instrumentpiano  -0.95045    0.15996  -5.942
Instrumentstring -2.54783    0.23034 -11.061
Voicepar3rd       0.14361    0.07664   1.874
Voicepar5th       0.16736    0.07660   2.185
Composing         0.26507    0.11746   2.257
X16.minus.17      0.11389    0.04403   2.586
KnowRob           0.18387    0.08499   2.163
X1990s2000s       0.16457    0.08379   1.964
GuitarPlay       -0.24291    0.12157  -1.998
```

Figure 12: Fixed effects estimates for each level of Harmonic Motion for Popular rating

effect from intercept, *Harmony* and *Instrument*.

To determine whether there are differences in the way that musicians and non-musicians identify classical music, we look at the summary table of our final model. From the interaction terms with t value greater than 2 or less than -2, we find that the influence of *GuitarPlay* and the occurrence of *Harmony I-V-VI* are different between musicians and non-musicians. These factors indicate the differences in the way that musicians and non-musicians identify classical music.

- For those who declare themselves as musicians, the occurrence of *Harmony I-V-VI* results in 2.37 unit increase in classical ratings, while for non-musicians, the occurrence of *Harmony I-V-VI* results in only 0.29 unit increase in classical ratings.

- For those who declare themselves as musicians, one unit increase in *GuitarPlay* results in 0.06 unit decrease in classical ratings, while for non-musicians, one unit increase in *GuitarPlay* results in 1.15 unit increase in classical ratings.

Since I-V-VI is the beginning progression for Pachelbels Canon in D, which many people have heard, it is not surprising that I-V-VI is frequently rated as classical. This is in accordance with our final model for classical ratings that *HarmonyI-V-VI* has a significant positive influence on classical ratings. Perhaps musicians are more sensitive to levels of harmony motion, it is easier for them to distinguish *HarmonyI-V-VI* from other levels. Another interesting finding is that non-musicians who play guitar are more likely to identify the same music as classical than musicians. Those who self-declare as musicians may have a solid background knowledge in music so that they think of guitar as a mainstream instrument in popular music. Therefore, whenever they hear guitar in musical stimuli, they prone to classify it as popular music.

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Popular ~ Voice + Harmony + Instrument + (1 | Subject) + (Harmony |
    Subject) + (Instrument | Subject) + Composing + X16.minus.17 +      KnowRob + X1990s2000s + GuitarPlay - 1
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9726.2   9906.1  -4832.1   9664.2     2422

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.0571 -0.5924  0.0293  0.5882  3.3093

Random effects:
 Groups    Name              Variance Std.Dev. Corr
 Subject   (Intercept)       0.01785  0.1336
 Subject.1 (Intercept)       0.79417  0.8912
           HarmonyI-V-IV     0.08321  0.2885    0.61
           HarmonyI-V-VI     0.86399  0.9295   -0.25 -0.43
           HarmonyIV-I-V     0.19909  0.4462   -0.38 -0.66 -0.39
 Subject.2 (Intercept)       0.74052  0.8605
           Instrumentpiano   1.37471  1.1725   -0.41
           Instrumentstring  3.29112  1.8141   -0.77  0.74
 Residual                    2.39488  1.5475
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                 Estimate Std. Error t value
Voicecontrary     5.51848    0.41657  13.247
Voicepar3rd       5.66209    0.41649  13.595
Voicepar5th       5.68586    0.41667  13.646
HarmonyI-V-IV    -0.04846    0.09502  -0.510
HarmonyI-V-VI    -0.30122    0.14213  -2.119
HarmonyIV-I-V    -0.23238    0.10320  -2.252
Instrumentpiano  -0.95046    0.15996  -5.942
Instrumentstring -2.54783    0.23034 -11.061
Composing         0.26559    0.11745   2.261
X16.minus.17      0.11377    0.04403   2.584
KnowRob           0.18400    0.08499   2.165
X1990s2000s       0.16524    0.08378   1.972
GuitarPlay       -0.24343    0.12157  -2.002
```

Figure 13: Fixed effects estimates for each level of Voice Leading for Popular rating

3.3 Are there differences in the things that drive classical, vs. popular, ratings?

To figure out the differences in the things that drive classical, vs. popular, ratings, we refer to Figure 16 and Figure 17, summary tables of our final model for classical ratings and popular ratings. *ClsListen* has a positive association with classical ratings but is not included in the model for popular ratings. *X16.minus.17*, *Composing* and *KnowRob* have positive associations with popular ratings but none of them is in the model for classical ratings. All of these influences are statistically significant.

ClsListen represents how much classical music does the subject listen. If someone frequently listens to classical music, he may be pretty familiar with the elements that features classical music. Even if some elements are also common in popular music, he will be prone to classify the music as classical whenever he identifies these elements. *X16.minus.17* is an auxiliary measure of listeners ability to distinguish classical vs popular music. *Composing* represents if the subject has done any music composing. According to our common sense, those who have done large amount of music composing may have a good grasp of music knowledge. Therefore, these two variables *X16.minus.17* and *Composing* measure how well the subject knows music. *KnowRob* represents if the subject has heard Rob Paravonians Pachelbel Rant. Rob Paravonian have been written about harmony progression I-V-vi in popular music. Therefore, for those who have good music knowledge and are familiar with Rob Paravonians Pachelbel Rant, it might be easy for them to distinguish harmony progression in music. Whenever they hear I-V-vi, they tend to label the music as popular.

4 Discussion

We use the dataset from a designed experiment, containing 36 musical stimuli presented to 70 listeners to measure the influence of factors on listeners' identification of music as "classical" or "popular".

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Harmony |
    Subject) + (Instrument | Subject) + ClsListen + X16.minus.17 +      X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc +
    KnowRob + Voice:Harmony + Harmony:KnowRob
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9675.1   9913.1  -4796.5   9593.1     2412

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.8221 -0.5819  0.0213  0.5701  4.0679

Random effects:
 Groups    Name            Variance  Std.Dev.  Corr
 Subject   (Intercept)     9.709e-07 0.0009853
 Subject.1 (Intercept)     1.625e+00 1.2746101
           HarmonyI-V-IV   3.918e-02 0.1979470  0.58
           HarmonyI-V-VI   1.230e+00 1.1088586 -0.59  0.32
           HarmonyIV-I-V   1.798e-03 0.0424012  0.83  0.94 -0.03
 Subject.2 (Intercept)     1.240e+00 1.1137213
           Instrumentpiano 1.655e+00 1.2866582 -0.59
           Instrumentstring 3.339e+00 1.8271627 -0.99  0.68
 Residual                  2.323e+00 1.5241553
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                            Estimate Std. Error t value
(Intercept)                  4.46396    0.51487   8.670
Instrumentpiano              1.33831    0.17157   7.800
Instrumentstring             3.05343    0.23142  13.195
Voicepar3rd                 -0.23399    0.15014  -1.558
Voicepar5th                 -0.20107    0.15121  -1.330
HarmonyI-V-IV                0.13482    0.15762   0.855
HarmonyI-V-VI                0.92778    0.21235   4.369
HarmonyIV-I-V               -0.11418    0.15503  -0.736
ClsListen                    0.24384    0.09568   2.549
X16.minus.17                -0.06393    0.04558  -1.403
X1990s2000s                 -0.25429    0.10396  -2.446
X1990s2000s.minus.1960s1970s 0.17949    0.09009   1.992
APTheory.fc1                 0.59863    0.32980   1.815
KnowRob                     -0.12821    0.10281  -1.247
Voicepar3rd:HarmonyI-V-IV   -0.34512    0.21320  -1.619
Voicepar5th:HarmonyI-V-IV   -0.20817    0.21382  -0.974
Voicepar3rd:HarmonyI-V-VI   -0.76463    0.21326  -3.585
Voicepar5th:HarmonyI-V-VI   -0.47665    0.21390  -2.228
Voicepar3rd:HarmonyIV-I-V    0.49990    0.21300   2.347
Voicepar5th:HarmonyIV-I-V    0.06315    0.21301   0.296
HarmonyI-V-IV:KnowRob        0.01414    0.05415   0.261
HarmonyI-V-VI:KnowRob        0.37077    0.09514   3.897
HarmonyIV-I-V:KnowRob       -0.01828    0.05202  -0.351
```

Figure 14: Summary of model for classical ratings with interaction of Harmony and KnowRob

According to our analysis, we have figured out experimental factor's influence on ratings: Instrument exerts the strongest influence on classical ratings among the three design factors (Instrument, Harmonic Motion, VoiceLeading), but doesn't have the strongest influence on popular ratings; *I-V-vi* have the strongest association with classical ratings among the levels of *Harmony motion*; It seems that whether the respondent is familiar with the Pachelbel rants matters to the association between *I-V-vi* and classical ratings; Aomng the levels of Voice Leading, contrary motion have the strongest association with classical ratings.

Other than the influence of three experimental factors, *ClsListen* has a positive association with classical ratings, while *X16.minus.17*, *Composing* and *KnowRob* have positive associations with popular ratings.

If we classify all listeners to musicians and non-musicians, we find *Harmony I-V-VI* and *GuitarPlay* indicate the differences in the way that musicians and non-musicians identify classical music. Compared to non-musicians, the occurrence of *Harmony I-V-VI* and decrease in *GuitarPlay* let musicians identify music as classical more likely.

These results roughly meet researchers' main hypotheses and meet our common sense.

However, there are some drawbacks in the models. First, we treat the missing values using mode imputation, which will cause bias and reduce the variability of variables. We simply discard the miscoded crazy values. However, although the proportion of these values are extremely small, they may contain valuable information. In the future, we could communicate with those who collected the data and explore what caused the crazy values to give a more reliable dataset. Second, when deciding what contribute to the difference in the way that musicians and non-musicians identify classical music, we dichotomize all respondents to musicians and non-musicians. Dichotomy might not be a rigorous way to classify. We could further check if the results are sensitive to where we dichotomize. Third, dichotomizing all respondents to musicians and non-musicians based on *Selfdeclare* might be too subjective since *Selfdeclare* represents how much they consider themselves

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Classical ~ Voice + Harmony + Instrument + ConsInstr + Instr.minus.Notes +
    ClsListen + CollegeMusic + APTheory + PianoPlay + GuitarPlay +
    (1 | Subject) + (Harmony | Subject) + (Instrument | Subject) +
    Harmony:Selfdeclare + Selfdeclare:X16.minus.17 + CollegeMusic:Selfdeclare +      GuitarPlay:Selfdeclare
   Data: musician.ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9709.3   9947.3  -4813.6   9627.3     2412

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.7488 -0.5856  0.0224  0.5596  4.2944

Random effects:
 Groups    Name             Variance Std.Dev. Corr
 Subject   (Intercept)      0.000377 0.01942
 Subject.1 (Intercept)      1.264297 1.12441
           HarmonyI-V-IV    0.033257 0.18237   0.66
           HarmonyI-V-VI    1.255313 1.12041  -0.46  0.31
           HarmonyIV-I-V    0.002098 0.04581   0.84  0.79 -0.29
 Subject.2 (Intercept)      1.373039 1.17177
           Instrumentpiano  1.657202 1.28732  -0.69
           Instrumentstring 3.323099 1.82294  -1.00  0.68
 Residual                   2.365342 1.53797
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                           Estimate Std. Error t value
(Intercept)                 3.82776    0.50517   7.577
Voicepar3rd                -0.38397    0.07616  -5.042
Voicepar5th                -0.35404    0.07612  -4.651
HarmonyI-V-IV              -0.06899    0.11747  -0.587
HarmonyI-V-VI               0.29069    0.20737   1.402
HarmonyIV-I-V               0.06235    0.11388   0.547
Instrumentpiano             1.33784    0.17195   7.780
Instrumentstring            3.05192    0.23117  13.202
ConsInstr                  -0.11331    0.08828  -1.283
Instr.minus.Notes           0.12919    0.08452   1.529
ClsListen                   0.15795    0.08887   1.777
CollegeMusic                0.55362    0.44203   1.252
APTheory                    0.77282    0.39506   1.956
PianoPlay                  -0.06780    0.09572  -0.708
GuitarPlay                  1.14758    0.47447   2.419
HarmonyI-IV-V:Selfdeclare1  1.15931    0.68684   1.688
HarmonyI-V-IV:Selfdeclare1  1.23111    0.70017   1.758
HarmonyI-V-VI:Selfdeclare1  2.37430    0.69064   3.438
HarmonyIV-I-V:Selfdeclare1  1.14804    0.69081   1.662
Selfdeclare0:X16.minus.17  -0.06315    0.05515  -1.145
Selfdeclare1:X16.minus.17  -0.13861    0.08287  -1.673
CollegeMusic:Selfdeclare1  -1.35599    0.68165  -1.989
GuitarPlay:Selfdeclare1    -1.21288    0.49225  -2.464
```

Figure 15: Differences in the way that musicians and non-musicians identify classical music

as musicians. It might be a good idea to dichotomize according to *OMSI*, which is a more objective variable that indicates whether one is a musician or not. Last, in this study, we treat two-level variables as factor variables and all the other discrete variables as numeric, so there are only two factor variables *APTheory* and *CollegeMusic*. In the future, we could do further research and make a more informative decision on whether to treat those five-level variables as factor variables. That may produce a more complex model with multiple levels in variables. Due to the limited time to do this study, we could try more data cleaning and processing methods in the future.

5 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/.

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Classical ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
    Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +      X16.minus.17 + APTheory.fc
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9684.8   9888.0  -4807.4   9614.8     2418

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.8300 -0.5793  0.0238  0.5765  4.0467

Random effects:
 Groups    Name             Variance  Std.Dev.  Corr
 Subject   (Intercept)      1.534e-08 0.0001239
 Subject.1 (Intercept)      1.731e+00 1.3155632
           HarmonyI-V-IV    4.391e-02 0.2095484  0.38
           HarmonyI-V-VI    1.607e+00 1.2677021 -0.54  0.32
           HarmonyIV-I-V    5.513e-03 0.0742516  0.01  0.51 -0.21
 Subject.2 (Intercept)      1.334e+00 1.1549499
           Instrumentpiano  1.654e+00 1.2861962 -0.66
           Instrumentstring 3.336e+00 1.8265721 -1.00  0.68
 Residual                   2.322e+00 1.5238935
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                          Estimate Std. Error t value
(Intercept)                3.79984    0.31535  12.050
HarmonyI-V-IV              0.14513    0.15292   0.949
HarmonyI-V-VI              1.19214    0.21409   5.568
HarmonyIV-I-V             -0.12802    0.15048  -0.851
Instrumentpiano            1.33843    0.17152   7.803
Instrumentstring           3.05373    0.23135  13.200
Voicepar3rd               -0.23403    0.15012  -1.559
Voicepar5th               -0.20128    0.15118  -1.331
ClsListen                  0.21309    0.08902   2.394
X16.minus.17              -0.07380    0.04659  -1.584
APTheory.fc1               0.52331    0.34252   1.528
HarmonyI-V-IV:Voicepar3rd -0.34459    0.21316  -1.617
HarmonyI-V-VI:Voicepar3rd -0.76397    0.21324  -3.583
HarmonyIV-I-V:Voicepar3rd  0.50067    0.21297   2.351
HarmonyI-V-IV:Voicepar5th -0.20793    0.21378  -0.973
HarmonyI-V-VI:Voicepar5th -0.47521    0.21387  -2.222
HarmonyIV-I-V:Voicepar5th  0.06482    0.21297   0.304
```

Figure 16: Summary of final model for Classical ratings

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Popular ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
    Subject) + (Instrument | Subject) + Composing + X16.minus.17 +      KnowRob + X1990s2000s + GuitarPlay
   Data: ratings
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  9726.2   9906.1  -4832.1   9664.2     2422

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.0597 -0.5935  0.0281  0.5885  3.3094

Random effects:
 Groups    Name             Variance Std.Dev. Corr
 Subject   (Intercept)      0.23773  0.4876
 Subject.1 (Intercept)      0.27728  0.5266
           HarmonyI-V-IV    0.08388  0.2896    1.00
           HarmonyI-V-VI    0.86317  0.9291   -0.43 -0.43
           HarmonyIV-I-V    0.19836  0.4454   -0.66 -0.66 -0.40
 Subject.2 (Intercept)      1.04306  1.0213
           Instrumentpiano  1.37470  1.1725   -0.34
           Instrumentstring 3.29113  1.8141   -0.65  0.74
 Residual                   2.39490  1.5475
Number of obs: 2453, groups:  Subject, 70

Fixed effects:
                 Estimate Std. Error t value
(Intercept)       5.52127    0.41668  13.251
HarmonyI-V-IV    -0.04840    0.09507  -0.509
HarmonyI-V-VI    -0.30119    0.14209  -2.120
HarmonyIV-I-V    -0.23234    0.10315  -2.252
Instrumentpiano  -0.95045    0.15996  -5.942
Instrumentstring -2.54783    0.23034 -11.061
Voicepar3rd       0.14361    0.07664   1.874
Voicepar5th       0.16736    0.07660   2.185
Composing         0.26507    0.11746   2.257
X16.minus.17      0.11389    0.04403   2.586
KnowRob           0.18387    0.08499   2.163
X1990s2000s       0.16457    0.08379   1.964
GuitarPlay       -0.24291    0.12157  -1.998
```

Figure 17: Summary of final model for Popular ratings

Technical Appendix

Q1

```
library(corrplot)
library(GGally)


# clean data

mode <- function(x) {
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]
}

ratings <- read.csv("ratings.csv")
summary(ratings)

##        X               Subject        Harmony      Instrument        Voice
##  Min.   :   1.0   15      :  36   I-IV-V:630   guitar:840   contrary:840
##  1st Qu.: 630.8   16      :  36   I-V-IV:630   piano :840   par3rd  :840
##  Median :1260.5   17      :  36   I-V-VI:630   string:840   par5th  :840
##  Mean   :1260.5   18b     :  36   IV-I-V:630
##  3rd Qu.:1890.2   19      :  36
##  Max.   :2520.0   20      :  36
##                   (Other):2304
##   Selfdeclare        OMSI          X16.minus.17      ConsInstr
##  Min.   :1.000   Min.   : 11.0   Min.   :-4.000   Min.   :0.000
##  1st Qu.:2.000   1st Qu.: 49.0   1st Qu.: 0.000   1st Qu.:1.670
##  Median :2.000   Median :145.5   Median : 1.000   Median :3.000
##  Mean   :2.443   Mean   :225.9   Mean   : 1.721   Mean   :2.857
##  3rd Qu.:3.000   3rd Qu.:323.0   3rd Qu.: 3.000   3rd Qu.:4.330
##  Max.   :6.000   Max.   :970.0   Max.   : 9.000   Max.   :5.000
##
##    ConsNotes      Instr.minus.Notes   PachListen      ClsListen
##  Min.   :0.000   Min.   :-4.0000   Min.   :0.000   Min.   :0.000
##  1st Qu.:0.750   1st Qu.: 0.0000   1st Qu.:5.000   1st Qu.:1.000
##  Median :3.000   Median : 0.3350   Median :5.000   Median :3.000
##  Mean   :2.533   Mean   : 0.6857   Mean   :4.515   Mean   :2.159
##  3rd Qu.:5.000   3rd Qu.: 2.0000   3rd Qu.:5.000   3rd Qu.:3.000
##  Max.   :5.000   Max.   : 4.3300   Max.   :5.000   Max.   :5.000
##  NA's   :360                       NA's   :72      NA's   :36
##     KnowRob          KnowAxis        X1990s2000s
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:3.000
##  Median :0.0000   Median :0.0000   Median :5.000
##  Mean   :0.7692   Mean   :0.9032   Mean   :4.061
##  3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:5.000
##  Max.   :5.0000   Max.   :5.0000   Max.   :5.000
##  NA's   :180      NA's   :288      NA's   :144
```

```
##  X1990s2000s.minus.1960s1970s  CollegeMusic      NoClass
##  Min.   :-4.000               Min.   :0.000   Min.   :0.0000
##  1st Qu.: 0.000               1st Qu.:1.000   1st Qu.:0.0000
##  Median : 2.000               Median :1.000   Median :1.0000
##  Mean   : 2.015               Mean   :0.791   Mean   :0.9194
##  3rd Qu.: 3.000               3rd Qu.:1.000   3rd Qu.:1.0000
##  Max.   : 5.000               Max.   :1.000   Max.   :8.0000
##  NA's   :180                  NA's   :108     NA's   :288
##     APTheory        Composing    PianoPlay       GuitarPlay
##  Min.   :0.0000   Min.   :0    Min.   :0.000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0    1st Qu.:0.000   1st Qu.:0.0000
##  Median :0.0000   Median :0    Median :0.000   Median :0.0000
##  Mean   :0.2344   Mean   :1    Mean   :1.086   Mean   :0.6857
##  3rd Qu.:0.0000   3rd Qu.:2    3rd Qu.:1.000   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :5    Max.   :5.000   Max.   :5.0000
##  NA's   :216      NA's   :72
##    X1stInstr       X2ndInstr       first12       Classical
##  Min.   :1.000   Min.   :0.000   guitar: 720   Min.   : 0.000
##  1st Qu.:1.000   1st Qu.:1.000   piano : 720   1st Qu.: 4.000
##  Median :3.500   Median :1.000   string:1080   Median : 6.000
##  Mean   :2.786   Mean   :1.556                 Mean   : 5.783
##  3rd Qu.:4.000   3rd Qu.:2.000                 3rd Qu.: 8.000
##  Max.   :5.000   Max.   :4.000                 Max.   :19.000
##  NA's   :1512    NA's   :2196                  NA's   :27
##     Popular
##  Min.   : 0.000
##  1st Qu.: 4.000
##  Median : 5.000
##  Mean   : 5.381
##  3rd Qu.: 7.000
##  Max.   :19.000
##  NA's   :27
```

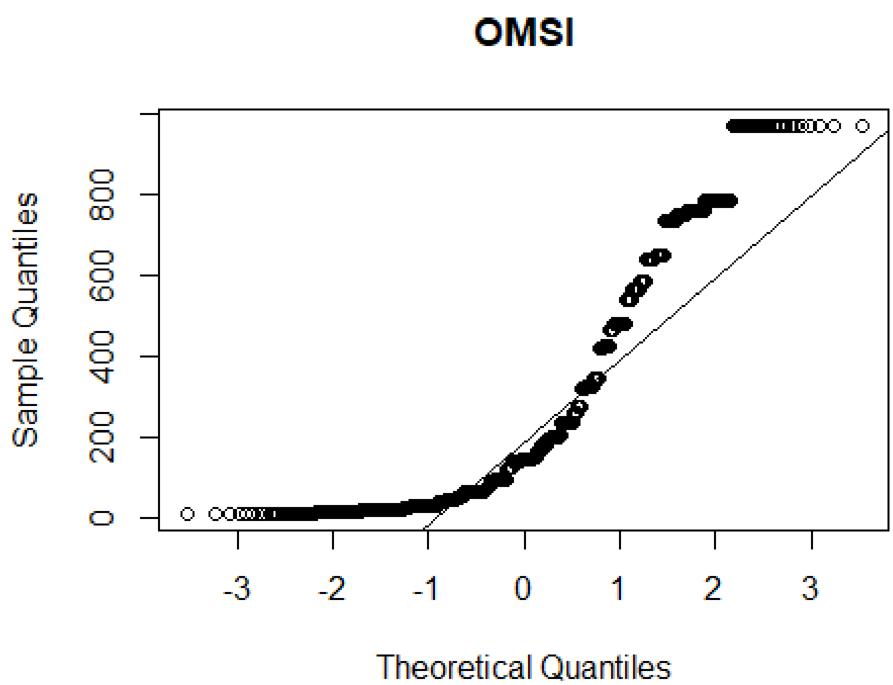
```
str(ratings)
```

```
## 'data.frame':    2520 obs. of  28 variables:
##  $ X                             : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Subject                       : Factor w/ 70 levels "15","16","17",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ Harmony                       : Factor w/ 4 levels "I-IV-V","I-V-IV",..:
1 1 1 1 1 1 1 1 1 2 ...
##  $ Instrument                    : Factor w/ 3 levels "guitar","piano",..: 1
 1 1 2 2 2 3 3 3 1 ...
##  $ Voice                         : Factor w/ 3 levels "contrary","par3rd",..
: 1 2 3 1 2 3 1 2 3 1 ...
##  $ Selfdeclare                   : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ OMSI                          : int  734 734 734 734 734 734 734 734 734
734 ...
##  $ X16.minus.17                  : num  5 5 5 5 5 5 5 5 5 5 ...
##  $ ConsInstr                     : num  4.33 4.33 4.33 4.33 4.33 4.33 4.33 4
```

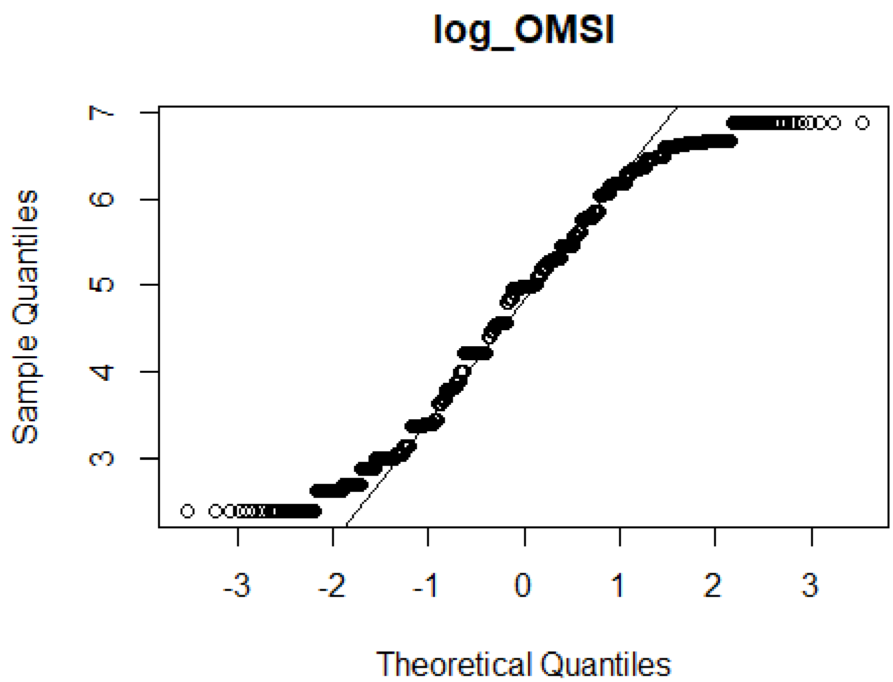
```
.33 4.33 4.33 ...
##  $ ConsNotes                    : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ Instr.minus.Notes            : num  -0.67 -0.67 -0.67 -0.67 -0.67 -0.67
-0.67 -0.67 -0.67 -0.67 ...
##  $ PachListen                   : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ ClsListen                    : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ KnowRob                      : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ KnowAxis                     : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X1990s2000s                  : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ X1990s2000s.minus.1960s1970s: int  2 2 2 2 2 2 2 2 2 2 ...
##  $ CollegeMusic                 : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ NoClass                      : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ APTheory                     : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ Composing                    : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ PianoPlay                    : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ GuitarPlay                   : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ X1stInstr                    : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ X2ndInstr                    : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ first12                      : Factor w/ 3 levels "guitar","piano",..: 3
 3 3 3 3 3 3 3 3 3 ...
##  $ Classical                    : num  3 3 1 3 2 8 10 6 5 1 ...
##  $ Popular                      : num  9 7 8 7 8 3 1 4 5 8 ...
```

```
ratings <- ratings[!is.na(ratings$Classical), ]
ratings <- ratings[!is.na(ratings$Popular), ]
ratings <- ratings[ratings$Classical <= 10 & ratings$Classical >= 1 &
                     ratings$Popular <= 10 & ratings$Popular >= 1, ]
ratings <- ratings[ratings$Classical == as.integer(ratings$Classical) &
                     ratings$Popular == as.integer(ratings$Popular), ]
ratings$ConsNotes[is.na(ratings$ConsNotes)] <-
  mode(ratings$ConsNotes[!is.na(ratings$ConsNotes)])
ratings$PachListen[is.na(ratings$PachListen)] <-
  mode(ratings$PachListen[!is.na(ratings$PachListen)])
ratings$ClsListen[is.na(ratings$ClsListen)] <-
  mode(ratings$ClsListen[!is.na(ratings$ClsListen)])
ratings$KnowRob[is.na(ratings$KnowRob)] <-
  mode(ratings$KnowRob[!is.na(ratings$KnowRob)])
ratings$KnowAxis[is.na(ratings$KnowAxis)] <-
  mode(ratings$KnowAxis[!is.na(ratings$KnowAxis)])
ratings$X1990s2000s[is.na(ratings$X1990s2000s)] <-
  mode(ratings$X1990s2000s[!is.na(ratings$X1990s2000s)])
ratings$X1990s2000s.minus.1960s1970s[is.na(ratings$X1990s2000s.minus.1960s197
0s)] <-
  mode(ratings$X1990s2000s.minus.1960s1970s[!is.na(ratings$X1990s2000s.minus.
1960s1970s)])
ratings$CollegeMusic[is.na(ratings$CollegeMusic)] <-
  mode(ratings$CollegeMusic[!is.na(ratings$CollegeMusic)])
ratings$NoClass[is.na(ratings$NoClass)] <-
  mode(ratings$NoClass[!is.na(ratings$NoClass)])
ratings$APTheory[is.na(ratings$APTheory)] <-
```

```r
  mode(ratings$APTheory[!is.na(ratings$APTheory)])
ratings$Composing[is.na(ratings$Composing)] <-
  mode(ratings$Composing[!is.na(ratings$Composing)])



# EDA
qqnorm(ratings[,"OMSI"],main="OMSI"); qqline(ratings[,"OMSI"])
```




```r
qqnorm(log(ratings$OMSI), main = "log_OMSI"); qqline(log(ratings$OMSI))
```




```
ratings$log_OMSI <- log(ratings$OMSI)
str(ratings)

## 'data.frame':    2453 obs. of  29 variables:
##  $ X                         : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Subject                   : Factor w/ 70 levels "15","16","17",..: 1
1 1 1 1 1 1 1 1 1 ...
##  $ Harmony                   : Factor w/ 4 levels "I-IV-V","I-V-IV",..:
1 1 1 1 1 1 1 1 1 2 ...
##  $ Instrument                : Factor w/ 3 levels "guitar","piano",..: 1
 1 1 2 2 2 3 3 3 1 ...
##  $ Voice                     : Factor w/ 3 levels "contrary","par3rd",..
: 1 2 3 1 2 3 1 2 3 1 ...
##  $ Selfdeclare               : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ OMSI                      : int  734 734 734 734 734 734 734 734 734
734 ...
##  $ X16.minus.17              : num  5 5 5 5 5 5 5 5 5 5 ...
##  $ ConsInstr                 : num  4.33 4.33 4.33 4.33 4.33 4.33 4.33 4
.33 4.33 4.33 ...
##  $ ConsNotes                 : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ Instr.minus.Notes         : num  -0.67 -0.67 -0.67 -0.67 -0.67 -0.67
-0.67 -0.67 -0.67 -0.67 ...
##  $ PachListen                : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ ClsListen                 : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ KnowRob                   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ KnowAxis                  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X1990s2000s               : int  5 5 5 5 5 5 5 5 5 5 ...
```

```
##  $ X1990s2000s.minus.1960s1970s: int  2 2 2 2 2 2 2 2 2 2 ...
##  $ CollegeMusic                : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ NoClass                     : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ APTheory                    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ Composing                   : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ PianoPlay                   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ GuitarPlay                  : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ X1stInstr                   : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ X2ndInstr                   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ first12                     : Factor w/ 3 levels "guitar","piano",..: 3
 3 3 3 3 3 3 3 3 3 ...
##  $ Classical                   : num  3 3 1 3 2 8 10 6 5 1 ...
##  $ Popular                     : num  9 7 8 7 8 3 1 4 5 8 ...
##  $ log_OMSI                    : num  6.6 6.6 6.6 6.6 6.6 ...

ratings_mat <- cor(ratings[, names(ratings[c(6:17,19,21:23)])])
corrplot.mixed(ratings_mat)
```

```
par(mfrow=c(4,4))
for (i in names(ratings[c(6:17,19,21:23)])) {
  boxplot(ratings[,i] ~ APTheory, data = ratings, ylab = i)
  }

par(mfrow=c(4,4))
for (i in names(ratings[c(6:17,19,21:23)])) {
  boxplot(ratings[,i] ~ CollegeMusic, data = ratings, ylab = i)
}



library(MASS)
lm.2 <- lm(Classical ~ Harmony + Instrument + Voice + Harmony:Instrument + Ha
rmony:Voice + Instrument:Voice + Harmony:Instrument:Voice, data = ratings)
fit2 <- stepAIC(lm.2, trace = 0)
formula(fit2)

## Classical ~ Harmony + Instrument + Voice + Harmony:Voice

summary(fit2)

##
## Call:
## lm(formula = Classical ~ Harmony + Instrument + Voice + Harmony:Voice,
##     data = ratings)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8152 -1.7368 -0.0223  1.6753  6.1881
##
```

```
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 4.24877    0.17052  24.917  < 2e-16 ***
## HarmonyI-V-IV               0.14665    0.22410   0.654   0.5129
## HarmonyI-V-VI               1.21005    0.22438   5.393  7.6e-08 ***
## HarmonyIV-I-V              -0.13107    0.22327  -0.587   0.5572
## Instrumentpiano             1.34142    0.11195  11.982  < 2e-16 ***
## Instrumentstring            3.04699    0.11198  27.209  < 2e-16 ***
## Voicepar3rd                -0.22871    0.22300  -1.026   0.3052
## Voicepar5th                -0.19369    0.22466  -0.862   0.3887
## HarmonyI-V-IV:Voicepar3rd -0.35478    0.31654  -1.121   0.2625
## HarmonyI-V-VI:Voicepar3rd -0.77804    0.31673  -2.456   0.0141 *
## HarmonyIV-I-V:Voicepar3rd  0.50344    0.31635   1.591   0.1116
## HarmonyI-V-IV:Voicepar5th -0.21846    0.31771  -0.688   0.4918
## HarmonyI-V-VI:Voicepar5th -0.49694    0.31752  -1.565   0.1177
## HarmonyIV-I-V:Voicepar5th  0.05127    0.31636   0.162   0.8713
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.266 on 2439 degrees of freedom
## Multiple R-squared:  0.2541, Adjusted R-squared:  0.2501
## F-statistic:  63.9 on 13 and 2439 DF,  p-value: < 2.2e-16

par(mfrow= c(2,2))
plot(fit2)
```

```
library(lme4)
library(RLRsim)
library(dplyr)
source("residual-functions.r")

# add random intercept
M0 <- lmer(Classical ~ Harmony + Instrument + Voice + Harmony:Voice + (1|Subj
ect), REML = F, data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(M0)

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony + Instrument + Voice + Harmony:Voice + (1 |
##     Subject)
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  10206.1  10298.9  -5087.0  10174.1     2437
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9332 -0.6298 -0.0251  0.6462  3.9629
##
```

```
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.670    1.292
##  Residual             3.411    1.847
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                  4.24049    0.20786  20.400
## HarmonyI-V-IV                0.14292    0.18271   0.782
## HarmonyI-V-VI                1.20869    0.18295   6.607
## HarmonyIV-I-V               -0.12648    0.18201  -0.695
## Instrumentpiano              1.34952    0.09137  14.770
## Instrumentstring             3.04137    0.09143  33.264
## Voicepar3rd                 -0.23155    0.18180  -1.274
## Voicepar5th                 -0.20475    0.18315  -1.118
## HarmonyI-V-IV:Voicepar3rd -0.34753    0.25809  -1.347
## HarmonyI-V-VI:Voicepar3rd -0.77238    0.25822  -2.991
## HarmonyIV-I-V:Voicepar3rd  0.49275    0.25790   1.911
## HarmonyI-V-IV:Voicepar5th -0.21109    0.25899  -0.815
## HarmonyI-V-VI:Voicepar5th -0.47720    0.25889  -1.843
## HarmonyIV-I-V:Voicepar5th  0.06393    0.25794   0.248
```

```
AIC(M0) - AIC(fit2)
```

```
## [1] -784.3421
```

```
BIC(M0) - BIC(fit2)
```

```
## [1] -778.5371
```

```
# DIC btw AIC and BIC

summary(M0)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony + Instrument + Voice + Harmony:Voice + (1 |
##     Subject)
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  10206.1  10298.9  -5087.0  10174.1     2437
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9332 -0.6298 -0.0251  0.6462  3.9629
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.670    1.292
```

```
##  Residual             3.411    1.847
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                  4.24049    0.20786  20.400
## HarmonyI-V-IV                0.14292    0.18271   0.782
## HarmonyI-V-VI                1.20869    0.18295   6.607
## HarmonyIV-I-V               -0.12648    0.18201  -0.695
## Instrumentpiano              1.34952    0.09137  14.770
## Instrumentstring             3.04137    0.09143  33.264
## Voicepar3rd                 -0.23155    0.18180  -1.274
## Voicepar5th                 -0.20475    0.18315  -1.118
## HarmonyI-V-IV:Voicepar3rd -0.34753    0.25809  -1.347
## HarmonyI-V-VI:Voicepar3rd -0.77238    0.25822  -2.991
## HarmonyIV-I-V:Voicepar3rd  0.49275    0.25790   1.911
## HarmonyI-V-IV:Voicepar5th -0.21109    0.25899  -0.815
## HarmonyI-V-VI:Voicepar5th -0.47720    0.25889  -1.843
## HarmonyIV-I-V:Voicepar5th  0.06393    0.25794   0.248
```

```
summary(fit2)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony + Instrument + Voice + Harmony:Voice,
##     data = ratings)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8152 -1.7368 -0.0223  1.6753  6.1881
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  4.24877    0.17052  24.917  < 2e-16 ***
## HarmonyI-V-IV                0.14665    0.22410   0.654   0.5129
## HarmonyI-V-VI                1.21005    0.22438   5.393  7.6e-08 ***
## HarmonyIV-I-V               -0.13107    0.22327  -0.587   0.5572
## Instrumentpiano              1.34142    0.11195  11.982  < 2e-16 ***
## Instrumentstring             3.04699    0.11198  27.209  < 2e-16 ***
## Voicepar3rd                 -0.22871    0.22300  -1.026   0.3052
## Voicepar5th                 -0.19369    0.22466  -0.862   0.3887
## HarmonyI-V-IV:Voicepar3rd -0.35478    0.31654  -1.121   0.2625
## HarmonyI-V-VI:Voicepar3rd -0.77804    0.31673  -2.456   0.0141 *
## HarmonyIV-I-V:Voicepar3rd  0.50344    0.31635   1.591   0.1116
## HarmonyI-V-IV:Voicepar5th -0.21846    0.31771  -0.688   0.4918
## HarmonyI-V-VI:Voicepar5th -0.49694    0.31752  -1.565   0.1177
## HarmonyIV-I-V:Voicepar5th  0.05127    0.31636   0.162   0.8713
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 2.266 on 2439 degrees of freedom
## Multiple R-squared:  0.2541, Adjusted R-squared:  0.2501
## F-statistic:  63.9 on 13 and 2439 DF,  p-value: < 2.2e-16

# Conditional residuals
ratings.res <- ratings %>% mutate(r.cond = r.cond(M0),
                          fitted = fitted(M0))
ggplot(ratings.res, aes(x=fitted,y=r.cond)) +
  facet_wrap( ~ Subject, as.table=F) + geom_point(pch=1) +
  geom_hline(yintercept = 0, col = "red", linetype="dashed")
```

```
library(LMERConvenienceFunctions)

vars <- attr(terms(formula(M0)),"term.labels")
vars <- vars[-length(vars)]

ratings.2c1 <- fitLMER.fnc(M0,
                         ran.effects=
                             list(slopes=c("Harmony","Instrument","Voice"),
                                  by.vars="Subject",
                                  corr=rep(1,length(vars))),method="AIC")

## ======================================================
## ===              backfitting fixed effects         ===
## ======================================================
## setting REML to FALSE
## processing model terms of interaction level 2
##   iteration 1
##     p-value for term "Harmony:Voice" = 2e-04 >= 0
##     not part of higher-order interaction
##     AIC simple = 10220; AIC complex = 10206; decrease = 14 >= 5
##     skipping term
## length = 1
## processing model terms of interaction level 1
##   iteration 2
##     p-value for term "Voice" = 0 >= 0
##     part of higher-order interaction
##     skipping term
##   iteration 3
##     p-value for term "Harmony" = 0 >= 0
##     part of higher-order interaction
##     skipping term
##   iteration 4
##     p-value for term "Instrument" = 0 >= 0
##     not part of higher-order interaction
##     AIC simple = 11114; AIC complex = 10206; decrease = 908 >= 5
##     skipping term
## length = 1
## pruning random effects structure ...
```

```
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects       ===
## ======================================================
##  ===       random slopes       ===
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 1.492751e-77
##  adding (Instrument | Subject) to model
## evaluating addition of (Voice|Subject) to model
## ======================================================
## ===            re-backfitting fixed effects        ===
## ======================================================
## setting REML to FALSE
## processing model terms of interaction level 2
##   iteration 1
##     p-value for term "Harmony:Voice" = 0 >= 0
##     not part of higher-order interaction
##     AIC simple = 9867; AIC complex = 9845; decrease = 22 >= 5
##     skipping term
## length = 1
## processing model terms of interaction level 1
##   iteration 2
##     p-value for term "Voice" = 0 >= 0
##     part of higher-order interaction
##     skipping term
##   iteration 3
##     p-value for term "Harmony" = 0 >= 0
##     part of higher-order interaction
##     skipping term
##   iteration 4
##     p-value for term "Instrument" = 0 >= 0
##     not part of higher-order interaction
##     AIC simple = 9929; AIC complex = 9845; decrease = 84 >= 5
##     skipping term
## length = 1
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/fitLMER_log_Su
n_Dec_08_01-18-00_2019.txt
```

```
formula(ratings.2c1)
```

```
## Classical ~ Harmony + Instrument + Voice + (1 | Subject) + (Instrument |
##     Subject) + Harmony:Voice
```

```
summary(M0)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony + Instrument + Voice + Harmony:Voice + (1 |
```

```
##     Subject)
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  10206.1  10298.9  -5087.0  10174.1     2437
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9332 -0.6298 -0.0251  0.6462  3.9629
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.670    1.292
##  Residual             3.411    1.847
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                  4.24049    0.20786  20.400
## HarmonyI-V-IV                0.14292    0.18271   0.782
## HarmonyI-V-VI                1.20869    0.18295   6.607
## HarmonyIV-I-V               -0.12648    0.18201  -0.695
## Instrumentpiano              1.34952    0.09137  14.770
## Instrumentstring             3.04137    0.09143  33.264
## Voicepar3rd                 -0.23155    0.18180  -1.274
## Voicepar5th                 -0.20475    0.18315  -1.118
## HarmonyI-V-IV:Voicepar3rd -0.34753    0.25809  -1.347
## HarmonyI-V-VI:Voicepar3rd -0.77238    0.25822  -2.991
## HarmonyIV-I-V:Voicepar3rd  0.49275    0.25790   1.911
## HarmonyI-V-IV:Voicepar5th -0.21109    0.25899  -0.815
## HarmonyI-V-VI:Voicepar5th -0.47720    0.25889  -1.843
## HarmonyIV-I-V:Voicepar5th  0.06393    0.25794   0.248
```

```
summary(ratings.2c1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Classical ~ Harmony + Instrument + Voice + (1 | Subject) + (Instrument |
##     Subject) + Harmony:Voice
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
## REML criterion at convergence: 9831
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4527 -0.5821  0.0113  0.5854  3.8823
##
## Random effects:
```

```
##  Groups     Name            Variance Std.Dev. Corr
##  Subject    (Intercept)     0.7983   0.8935
##  Subject.1 (Intercept)      1.9123   1.3829
##            Instrumentpiano  1.6217   1.2735   -0.51
##            Instrumentstring 3.3390   1.8273   -0.78  0.68
##  Residual                  2.6624   1.6317
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                           Estimate Std. Error t value
## (Intercept)                4.24376    0.23209  18.285
## HarmonyI-V-IV              0.14494    0.16149   0.898
## HarmonyI-V-VI              1.20954    0.16175   7.478
## HarmonyIV-I-V             -0.12658    0.16084  -0.787
## Instrumentpiano            1.33883    0.17259   7.757
## Instrumentstring           3.05842    0.23325  13.112
## Voicepar3rd               -0.23540    0.16073  -1.465
## Voicepar5th               -0.20324    0.16187  -1.256
## HarmonyI-V-IV:Voicepar3rd -0.34342    0.22821  -1.505
## HarmonyI-V-VI:Voicepar3rd -0.77139    0.22824  -3.380
## HarmonyIV-I-V:Voicepar3rd  0.49904    0.22802   2.189
## HarmonyI-V-IV:Voicepar5th -0.20689    0.22889  -0.904
## HarmonyI-V-VI:Voicepar5th -0.47863    0.22893  -2.091
## HarmonyIV-I-V:Voicepar5th  0.06489    0.22800   0.285
```

```
ratings$CollegeMusic.fc <- as.factor(ratings$CollegeMusic)
ratings$APTheory.fc <- as.factor(ratings$APTheory)

lmer.ranef.Harmony.Instrument <- ratings.2c1


# add the first fixed effect
AIC_ls_num = list()
for (i in c("Selfdeclare","X16.minus.17","ConsInstr","ConsNotes",
            "Instr.minus.Notes","PachListen","ClsListen","KnowRob",
            "KnowAxis","X1990s2000s","X1990s2000s.minus.1960s1970s",
            "NoClass","Composing","PianoPlay","GuitarPlay","log_OMSI",
            "CollegeMusic.fc","APTheory.fc")){
  lm.HA.1 <- update(lmer.ranef.Harmony.Instrument, . ~. + eval(parse(text=i))
)
  anov <- anova(lmer.ranef.Harmony.Instrument, lm.HA.1, test = "Chisq")
  AIC_ls_num[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls_num[which.min(AIC_ls_num)]
```

```
## $ClsListen
```

```
## [1] -3.144231
```

```r
names(AIC_ls_num)[which.min(AIC_ls_num)]

## [1] "ClsListen"

lmer.Cls.ranef.Har.Instr <- update(lmer.ranef.Harmony.Instrument, . ~. + eval
(parse(text=names(AIC_ls_num)[which.min(AIC_ls_num)])))

# add the second fixed effect
AIC_ls2_num = list()
for (i in c("Selfdeclare","ConsInstr","ConsNotes",
            "Instr.minus.Notes","PachListen","KnowRob","KnowAxis",
            "X1990s2000s","X1990s2000s.minus.1960s1970s","NoClass",
            "Composing","PianoPlay","GuitarPlay","log_OMSI","CollegeMusic.fc"
,
            "APTheory.fc")){
  lm.HA.2 <- update(lmer.Cls.ranef.Har.Instr, . ~. + eval(parse(text=i)))
  anov <- anova(lmer.Cls.ranef.Har.Instr, lm.HA.2, test = "Chisq")
  AIC_ls2_num[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls2_num[which.min(AIC_ls2_num)]

## $X16.minus.17

## [1] -0.5604761

names(AIC_ls2_num)[which.min(AIC_ls2_num)]

## [1] "X16.minus.17"

lmer.Cls.X16min17.ranef.Har.Instr <-
  update(lmer.Cls.ranef.Har.Instr, . ~. +
           eval(parse(text=names(AIC_ls2_num)[which.min(AIC_ls2_num)])))

# add the third fixed effect
AIC_ls3_num = list()
for (i in c("Selfdeclare","X16.minus.17","ConsInstr","ConsNotes",
            "Instr.minus.Notes","PachListen","KnowRob","KnowAxis",
            "X1990s2000s","X1990s2000s.minus.1960s1970s","NoClass",
            "Composing","PianoPlay","GuitarPlay","log_OMSI",
            "CollegeMusic.fc","APTheory.fc")){
  lm.HA.3 <- update(lmer.Cls.X16min17.ranef.Har.Instr, . ~. + eval(parse(text
=i)))
  anov <- anova(lmer.Cls.X16min17.ranef.Har.Instr, lm.HA.3, test = "Chisq")
  AIC_ls3_num[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls3_num[which.min(AIC_ls3_num)]

## $APTheory.fc
```

```
## [1] -0.1266359

names(AIC_ls3_num)[which.min(AIC_ls3_num)]

## [1] " APTheory.fc"

lmer.Cls.X16min17.AP.ranef.Har.Instr <- update(lmer.Cls.X16min17.ranef.Har.In
str, . ~. + eval(parse(text=names(AIC_ls3_num)[which.min(AIC_ls3_num)])))

# add the fourth fixed effect
AIC_ls4_num = list()
for (i in c("Selfdeclare","ConsInstr","ConsNotes","Instr.minus.Notes",
            "PachListen","KnowRob","KnowAxis","X1990s2000s",
            "X1990s2000s.minus.1960s1970s","NoClass","Composing",
            "PianoPlay","GuitarPlay","log_OMSI","CollegeMusic.fc")){
  lm.HA.4 <-
    update(lmer.Cls.X16min17.AP.ranef.Har.Instr, . ~. + eval(parse(text=i)))
  anov <- anova(lmer.Cls.X16min17.AP.ranef.Har.Instr, lm.HA.4, test = "Chisq"
)
  AIC_ls4_num[[i]] <- anov[2,2] - anov[1,2]
}


AIC_ls4_num[which.min(AIC_ls4_num)] # >0 stop adding fixed effects

## $X1990s2000s

## [1] 0.001833417

names(AIC_ls4_num)[which.min(AIC_ls4_num)]

## [1] "X1990s2000s"

lmer.add.fixef <- lmer(Classical ~ Harmony + Instrument + Voice + (1 | Subjec
t) +
                         (Harmony | Subject) + (Instrument | Subject) +
                         Harmony:Voice + ClsListen + X16.minus.17 + APTheory.
fc,
                       REML = F, data = ratings, control = lmerControl(optimi
zer = "bobyqa"))
formula(lmer.add.fixef)

## Classical ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
##     X16.minus.17 + APTheory.fc

# check if there should be any changes in random effects
vars <- attr(terms(formula(lmer.add.fixef)),"term.labels")
vars <- vars[-c(4,5,6,10)]
lmer.add.fixef.0 <- ffRanefLMER.fnc(lmer.add.fixef,
                         ran.effects=
```

```
                    list(slopes=vars, by.vars="Subject",
                corr=rep(0,length(vars))))

##  ===       random slopes       ===
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
## evaluating addition of (0 + ClsListen|Subject) to model
## evaluating addition of (0 + X16.minus.17|Subject) to model
## evaluating addition of (APTheory.fc|Subject) to model
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/ffRanefLMER_lo
g_Sun_Dec_08_01-19-22_2019.txt

lmer.add.fixef.1 <- ffRanefLMER.fnc(lmer.add.fixef,
                ran.effects=
                    list(slopes=vars, by.vars="Subject",
                        corr=rep(1,length(vars))))

##  ===       random slopes       ===
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
## evaluating addition of (1 + ClsListen|Subject) to model
## evaluating addition of (1 + X16.minus.17|Subject) to model
## evaluating addition of (APTheory.fc|Subject) to model
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/ffRanefLMER_lo
g_Sun_Dec_08_01-20-42_2019.txt

anova(lmer.add.fixef.0, lmer.add.fixef.1)

## Data: ratings
## Models:
## lmer.add.fixef.0: Classical ~ Harmony + Instrument + Voice + (1 | Subject)
 + (Harmony |
## lmer.add.fixef.0:     Subject) + (Instrument | Subject) + Harmony:Voice +
ClsListen +
## lmer.add.fixef.0:     X16.minus.17 + APTheory.fc
## lmer.add.fixef.1: Classical ~ Harmony + Instrument + Voice + (1 | Subject)
 + (Harmony |
## lmer.add.fixef.1:     Subject) + (Instrument | Subject) + Harmony:Voice +
ClsListen +
## lmer.add.fixef.1:     X16.minus.17 + APTheory.fc
##                  Df    AIC  BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## lmer.add.fixef.0 35 9684.8 9888 -4807.4   9614.8
## lmer.add.fixef.1 35 9684.8 9888 -4807.4   9614.8     0      0          1

## the models turn out to be the same!
```

```
## automatic
## lm + stepAIC
lm.full.fixef <- lm(Classical ~ Harmony + Instrument + Voice + Harmony:Voice
+
                      Selfdeclare + log_OMSI + X16.minus.17 + ConsInstr +
                      ConsNotes + Instr.minus.Notes + PachListen + ClsListen
+
                      KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.19
60s1970s +
                      CollegeMusic.fc + NoClass + APTheory.fc + Composing + P
ianoPlay +
                      GuitarPlay, data=ratings)

lm.add.fixef <- stepAIC(lm.full.fixef, direction = "backward", trace = 0)
summary(lm.add.fixef)

##
## Call:
## lm(formula = Classical ~ Harmony + Instrument + Voice + Selfdeclare +
##     log_OMSI + X16.minus.17 + PachListen + ClsListen + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + APTheory.fc + Composing +
##     PianoPlay + Harmony:Voice, data = ratings)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.1769 -1.6134  0.0101  1.6126  6.1695
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    3.78578    0.33053  11.454  < 2e-16 ***
## HarmonyI-V-IV                  0.14572    0.21513   0.677   0.4983
## HarmonyI-V-VI                  1.21494    0.21541   5.640 1.90e-08 ***
## HarmonyIV-I-V                 -0.12950    0.21434  -0.604   0.5458
## Instrumentpiano                1.34971    0.10750  12.556  < 2e-16 ***
## Instrumentstring               3.05015    0.10751  28.370  < 2e-16 ***
## Voicepar3rd                   -0.23051    0.21408  -1.077   0.2817
## Voicepar5th                   -0.19297    0.21567  -0.895   0.3710
## Selfdeclare                   -0.57427    0.05915  -9.709  < 2e-16 ***
## log_OMSI                       0.23057    0.05235   4.404 1.11e-05 ***
## X16.minus.17                  -0.10959    0.01605  -6.828 1.08e-11 ***
## PachListen                     0.08031    0.04502   1.784   0.0746 .
## ClsListen                      0.23024    0.03551   6.483 1.08e-10 ***
## X1990s2000s                   -0.14872    0.03495  -4.255 2.17e-05 ***
## X1990s2000s.minus.1960s1970s   0.14915    0.03164   4.713 2.57e-06 ***
## APTheory.fc1                   0.21665    0.12266   1.766   0.0775 .
## Composing                      0.23650    0.03983   5.937 3.31e-09 ***
## PianoPlay                      0.07998    0.03249   2.461   0.0139 *
## HarmonyI-V-IV:Voicepar3rd     -0.34782    0.30387  -1.145   0.2525
## HarmonyI-V-VI:Voicepar3rd     -0.77518    0.30406  -2.549   0.0109 *
## HarmonyIV-I-V:Voicepar3rd      0.49578    0.30369   1.633   0.1027
```

```
## HarmonyI-V-IV:Voicepar5th    -0.22136    0.30499  -0.726   0.4680
## HarmonyI-V-VI:Voicepar5th    -0.49297    0.30483  -1.617   0.1060
## HarmonyIV-I-V:Voicepar5th     0.06116    0.30371   0.201   0.8404
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.175 on 2429 degrees of freedom
## Multiple R-squared:  0.3154, Adjusted R-squared:  0.3089
## F-statistic: 48.66 on 23 and 2429 DF,  p-value: < 2.2e-16
```

```
formula(lm.add.fixef)
```

```
## Classical ~ Harmony + Instrument + Voice + Selfdeclare + log_OMSI +
##     X16.minus.17 + PachListen + ClsListen + X1990s2000s + X1990s2000s.minu
s.1960s1970s +
##     APTheory.fc + Composing + PianoPlay + Harmony:Voice
```

```
lmer.autoadd.fixef <- lmer(Classical ~ Harmony + Instrument + Voice + Selfdec
lare +
                             log_OMSI + X16.minus.17 + Instr.minus.Notes + Cl
sListen +
                             X1990s2000s + X1990s2000s.minus.1960s1970s + NoC
lass +
                             APTheory.fc + Composing + PianoPlay + Harmony:Vo
ice +
                             (1 | Subject) + (Harmony | Subject) + (Instrumen
t | Subject),
                           REML = F, data = ratings, control = lmerControl(op
timizer = "bobyqa"))


formula(lmer.autoadd.fixef)
```

```
## Classical ~ Harmony + Instrument + Voice + Selfdeclare + log_OMSI +
##     X16.minus.17 + Instr.minus.Notes + ClsListen + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + NoClass + APTheory.fc + Composing +
##     PianoPlay + Harmony:Voice + (1 | Subject) + (Harmony | Subject) +
##     (Instrument | Subject)
```

```
anova(lmer.autoadd.fixef, lmer.add.fixef)
```

```
## Data: ratings
## Models:
## lmer.add.fixef: Classical ~ Harmony + Instrument + Voice + (1 | Subject) +
 (Harmony |
## lmer.add.fixef:     Subject) + (Instrument | Subject) + Harmony:Voice + Cl
sListen +
## lmer.add.fixef:     X16.minus.17 + APTheory.fc
## lmer.autoadd.fixef: Classical ~ Harmony + Instrument + Voice + Selfdeclare
 + log_OMSI +
## lmer.autoadd.fixef:     X16.minus.17 + Instr.minus.Notes + ClsListen + X19
```

```
90s2000s +
## lmer.autoadd.fixef:     X1990s2000s.minus.1960s1970s + NoClass + APTheory.
fc + Composing +
## lmer.autoadd.fixef:     PianoPlay + Harmony:Voice + (1 | Subject) + (Harmo
ny | Subject) +
## lmer.autoadd.fixef:     (Instrument | Subject)
##                     Df    AIC  BIC  logLik deviance  Chisq Chi Df
## lmer.add.fixef      35 9684.8 9888 -4807.4   9614.8
## lmer.autoadd.fixef  43 9687.4 9937 -4800.7   9601.4 13.464      8
##                     Pr(>Chisq)
## lmer.add.fixef
## lmer.autoadd.fixef    0.09684 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# go back and check fixed effects (lower AIC)
anova(lmer.add.fixef, ratings.2c1)
```

```
## Data: ratings
## Models:
## ratings.2c1: Classical ~ Harmony + Instrument + Voice + (1 | Subject) + (I
nstrument |
## ratings.2c1:     Subject) + Harmony:Voice
## lmer.add.fixef: Classical ~ Harmony + Instrument + Voice + (1 | Subject) +
 (Harmony |
## lmer.add.fixef:     Subject) + (Instrument | Subject) + Harmony:Voice + Cl
sListen +
## lmer.add.fixef:     X16.minus.17 + APTheory.fc
##                 Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## ratings.2c1    22 9844.7 9972.4 -4900.4   9800.7
## lmer.add.fixef 35 9684.8 9888.0 -4807.4   9614.8 185.87     13  < 2.2e-16
##
## ratings.2c1
## lmer.add.fixef ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(lmer.add.fixef)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
##     X16.minus.17 + APTheory.fc
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9684.8   9888.0  -4807.4   9614.8     2418
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8300 -0.5793  0.0238  0.5765  4.0467
##
## Random effects:
##  Groups    Name              Variance  Std.Dev.  Corr
##  Subject   (Intercept)       1.534e-08 0.0001239
##  Subject.1 (Intercept)       1.731e+00 1.3155632
##            HarmonyI-V-IV     4.391e-02 0.2095484  0.38
##            HarmonyI-V-VI     1.607e+00 1.2677021 -0.54  0.32
##            HarmonyIV-I-V     5.513e-03 0.0742516  0.01  0.51 -0.21
##  Subject.2 (Intercept)       1.334e+00 1.1549499
##            Instrumentpiano   1.654e+00 1.2861962 -0.66
##            Instrumentstring  3.336e+00 1.8265721 -1.00  0.68
##  Residual                    2.322e+00 1.5238935
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                  3.79984    0.31535  12.050
## HarmonyI-V-IV                0.14513    0.15292   0.949
## HarmonyI-V-VI                1.19214    0.21409   5.568
## HarmonyIV-I-V               -0.12802    0.15048  -0.851
## Instrumentpiano              1.33843    0.17152   7.803
## Instrumentstring             3.05373    0.23135  13.200
## Voicepar3rd                 -0.23403    0.15012  -1.559
## Voicepar5th                 -0.20128    0.15118  -1.331
## ClsListen                    0.21309    0.08902   2.394
## X16.minus.17                -0.07380    0.04659  -1.584
## APTheory.fc1                 0.52331    0.34252   1.528
## HarmonyI-V-IV:Voicepar3rd -0.34459    0.21316  -1.617
## HarmonyI-V-VI:Voicepar3rd -0.76397    0.21324  -3.583
## HarmonyIV-I-V:Voicepar3rd  0.50067    0.21297   2.351
## HarmonyI-V-IV:Voicepar5th -0.20793    0.21378  -0.973
## HarmonyI-V-VI:Voicepar5th -0.47521    0.21387  -2.222
## HarmonyIV-I-V:Voicepar5th  0.06482    0.21297   0.304
## convergence code: 0
## boundary (singular) fit: see ?isSingular

# Conditional residuals
ratings.res <- ratings %>% mutate(r.cond = r.cond(lmer.add.fixef),
                          fitted = fitted(lmer.add.fixef))
ggplot(ratings.res, aes(x=fitted,y=r.cond)) +
  facet_wrap( ~ Subject, as.table=F) + geom_point(pch=1) +
  geom_hline(yintercept = 0, col = "red", linetype="dashed")
```

```r
q1.c <- lmer(Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Harm
ony |
    Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
    X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc,
    REML = F, data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.c)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
##     X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     APTheory.fc
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9683.9   9898.7  -4805.0   9609.9     2416
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8291 -0.5827  0.0238  0.5720  4.0561
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.00000  0.00000
##  Subject.1 (Intercept)      1.68004  1.29616
##            HarmonyI-V-IV    0.04295  0.20723   0.50
##            HarmonyI-V-VI    1.60682  1.26760  -0.59  0.33
##            HarmonyIV-I-V    0.00307  0.05541   0.64  0.69 -0.31
##  Subject.2 (Intercept)      1.29775  1.13919
##            Instrumentpiano  1.65495  1.28645  -0.60
##            Instrumentstring 3.33691  1.82672  -0.99  0.68
##  Residual                   2.32292  1.52411
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                   4.33162    0.51788   8.364
## Instrumentpiano               1.33826    0.17155   7.801
## Instrumentstring              3.05306    0.23137  13.196
## Voicepar3rd                  -0.23379    0.15014  -1.557
## Voicepar5th                  -0.20092    0.15120  -1.329
## HarmonyI-V-IV                 0.14533    0.15289   0.951
## HarmonyI-V-VI                 1.19158    0.21410   5.566
## HarmonyIV-I-V                -0.12756    0.15038  -0.848
## ClsListen                     0.25666    0.09324   2.753
## X16.minus.17                 -0.07261    0.04590  -1.582
## X1990s2000s                  -0.23896    0.10476  -2.281
## X1990s2000s.minus.1960s1970s  0.16542    0.09084   1.821
```

```
## APTheory.fc1                  0.56990    0.33232   1.715
## Voicepar3rd:HarmonyI-V-IV    -0.34514    0.21319  -1.619
## Voicepar5th:HarmonyI-V-IV    -0.20823    0.21381  -0.974
## Voicepar3rd:HarmonyI-V-VI    -0.76420    0.21326  -3.583
## Voicepar5th:HarmonyI-V-VI    -0.47531    0.21389  -2.222
## Voicepar3rd:HarmonyIV-I-V     0.50026    0.21300   2.349
## Voicepar5th:HarmonyIV-I-V     0.06373    0.21300   0.299
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# all levels of Instrument
q1.1 <- lmer(Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Harm
ony |
    Subject) + (Instrument - 1 | Subject) + Harmony:Voice + ClsListen +
    X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc -
 1,
    REML = F, data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.1)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Harmony |
##     Subject) + (Instrument - 1 | Subject) + Harmony:Voice + ClsListen +
##     X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     APTheory.fc - 1
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9683.9   9898.7  -4805.0   9609.9     2416
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8291 -0.5827  0.0238  0.5720  4.0561
##
## Random effects:
##  Groups    Name                Variance  Std.Dev.  Corr
##  Subject   (Intercept)         1.363e-09 3.692e-05
##  Subject.1 (Intercept)         1.680e+00 1.296e+00
##            HarmonyI-V-IV       4.294e-02 2.072e-01  0.50
##            HarmonyI-V-VI       1.607e+00 1.268e+00 -0.59  0.33
##            HarmonyIV-I-V       3.070e-03 5.540e-02  0.64  0.69 -0.31
##  Subject.2 Instrumentguitar    1.298e+00 1.139e+00
##            Instrumentpiano     1.201e+00 1.096e+00  0.34
##            Instrumentstring    4.959e-01 7.042e-01 -0.96 -0.07
##  Residual                      2.323e+00 1.524e+00
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                                 Estimate Std. Error t value
```

```
## Instrumentguitar              4.33162    0.51788   8.364
## Instrumentpiano               5.66989    0.51664  10.974
## Instrumentstring              7.38468    0.50760  14.548
## Voicepar3rd                  -0.23379    0.15014  -1.557
## Voicepar5th                  -0.20092    0.15120  -1.329
## HarmonyI-V-IV                 0.14533    0.15289   0.951
## HarmonyI-V-VI                 1.19158    0.21410   5.566
## HarmonyIV-I-V                -0.12756    0.15038  -0.848
## ClsListen                     0.25666    0.09324   2.753
## X16.minus.17                 -0.07261    0.04590  -1.582
## X1990s2000s                  -0.23896    0.10476  -2.281
## X1990s2000s.minus.1960s1970s  0.16542    0.09084   1.821
## APTheory.fc1                  0.56990    0.33232   1.715
## Voicepar3rd:HarmonyI-V-IV    -0.34514    0.21319  -1.619
## Voicepar5th:HarmonyI-V-IV    -0.20823    0.21381  -0.974
## Voicepar3rd:HarmonyI-V-VI    -0.76420    0.21326  -3.583
## Voicepar5th:HarmonyI-V-VI    -0.47531    0.21389  -2.222
## Voicepar3rd:HarmonyIV-I-V     0.50026    0.21300   2.349
## Voicepar5th:HarmonyIV-I-V     0.06373    0.21300   0.299
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# all levels of Harmony
q1.2 <- lmer(Classical ~ Harmony + Voice + Instrument + (1 | Subject) + (Harm
ony - 1 |
    Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
    X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc -
 1,
    REML = F, data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Harmony + Voice + Instrument + (1 | Subject) + (Harmony -
##     1 | Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
##     X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     APTheory.fc - 1
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9683.9   9898.7  -4805.0   9609.9     2416
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8291 -0.5827  0.0238  0.5720  4.0561
##
## Random effects:
##  Groups    Name            Variance  Std.Dev.  Corr
##  Subject   (Intercept)     8.674e-11 9.314e-06
```

```
##  Subject.1 HarmonyI-IV-V    1.680e+00 1.296e+00
##            HarmonyI-V-IV    1.990e+00 1.411e+00 0.99
##            HarmonyI-V-VI    1.350e+00 1.162e+00 0.47  0.57
##            HarmonyIV-I-V    1.775e+00 1.332e+00 1.00  0.99  0.48
##  Subject.2 (Intercept)      1.298e+00 1.139e+00
##            Instrumentpiano  1.655e+00 1.286e+00 -0.60
##            Instrumentstring 3.337e+00 1.827e+00 -0.99  0.68
##  Residual                   2.323e+00 1.524e+00
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                               Estimate Std. Error t value
## HarmonyI-IV-V                  4.33162    0.51788   8.364
## HarmonyI-V-IV                  4.47695    0.52228   8.572
## HarmonyI-V-VI                  5.52321    0.51295  10.768
## HarmonyIV-I-V                  4.20407    0.51886   8.103
## Voicepar3rd                   -0.23379    0.15014  -1.557
## Voicepar5th                   -0.20092    0.15120  -1.329
## Instrumentpiano                1.33826    0.17155   7.801
## Instrumentstring               3.05306    0.23137  13.196
## ClsListen                      0.25666    0.09324   2.753
## X16.minus.17                  -0.07261    0.04590  -1.582
## X1990s2000s                   -0.23896    0.10476  -2.281
## X1990s2000s.minus.1960s1970s  0.16542    0.09084   1.821
## APTheory.fc1                   0.56990    0.33232   1.715
## HarmonyI-V-IV:Voicepar3rd     -0.34514    0.21319  -1.619
## HarmonyI-V-VI:Voicepar3rd     -0.76420    0.21326  -3.583
## HarmonyIV-I-V:Voicepar3rd      0.50026    0.21300   2.349
## HarmonyI-V-IV:Voicepar5th     -0.20823    0.21381  -0.974
## HarmonyI-V-VI:Voicepar5th     -0.47531    0.21389  -2.222
## HarmonyIV-I-V:Voicepar5th      0.06373    0.21300   0.299
## convergence code: 1
## boundary (singular) fit: see ?isSingular
```

```
q1.2.2.Rob <- update(q1.c, . ~. + Harmony*KnowRob)
anova(q1.c, q1.2.2.Rob)
```

```
## Data: ratings
## Models:
## q1.c: Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Harmony
|
## q1.c:     Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
## q1.c:     X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s +
## q1.c:     APTheory.fc
## q1.2.2.Rob: Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Ha
rmony |
## q1.2.2.Rob:     Subject) + (Instrument | Subject) + ClsListen + X16.minus.
17 +
## q1.2.2.Rob:     X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc +
```

```
## q1.2.2.Rob:     KnowRob + Voice:Harmony + Harmony:KnowRob
##             Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## q1.c        37 9683.9 9898.7 -4805.0   9609.9
## q1.2.2.Rob 41 9675.1 9913.1 -4796.5   9593.1 16.866      4   0.002052 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
q1.2.2.Axis <- update(q1.c, . ~. + Harmony*KnowAxis)
anova(q1.2.2.Rob, q1.2.2.Axis)
```

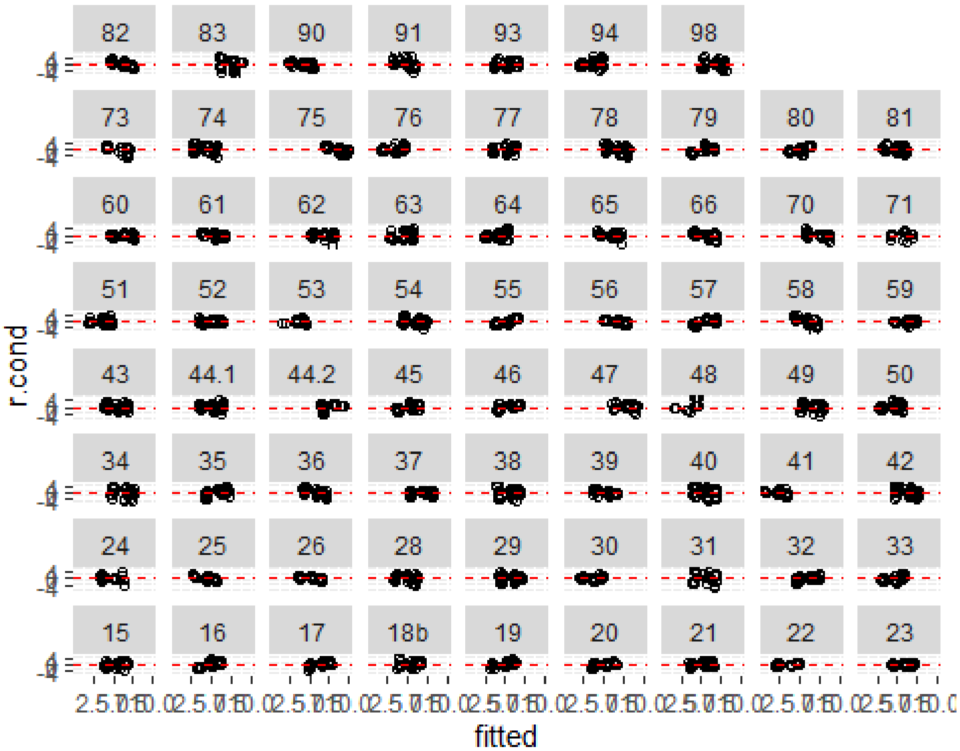
```
## Data: ratings
## Models:
## q1.2.2.Rob: Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (Ha
rmony |
## q1.2.2.Rob:     Subject) + (Instrument | Subject) + ClsListen + X16.minus.
17 +
## q1.2.2.Rob:     X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc +

## q1.2.2.Rob:     KnowRob + Voice:Harmony + Harmony:KnowRob
## q1.2.2.Axis: Classical ~ Instrument + Voice + Harmony + (1 | Subject) + (H
armony |
## q1.2.2.Axis:     Subject) + (Instrument | Subject) + ClsListen + X16.minus
.17 +
## q1.2.2.Axis:     X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc
+
## q1.2.2.Axis:     KnowAxis + Voice:Harmony + Harmony:KnowAxis
##              Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## q1.2.2.Rob  41 9675.1 9913.1 -4796.5   9593.1
## q1.2.2.Axis 41 9690.8 9928.8 -4804.4   9608.8     0      0          1
```

```
# all levels of Voice
q1.3 <- lmer(Classical ~ Voice + Instrument + Harmony + (1 | Subject) + (Harm
ony |
    Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
    X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s + APTheory.fc -
 1,
    REML = F, data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.3)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Voice + Instrument + Harmony + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Harmony:Voice + ClsListen +
##     X16.minus.17 + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     APTheory.fc - 1
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9683.9   9898.7  -4805.0   9609.9     2416
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8291 -0.5827  0.0238  0.5720  4.0561
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      5.842e-08 0.0002417
##  Subject.1 (Intercept)      1.680e+00 1.2961648
##            HarmonyI-V-IV    4.295e-02 0.2072320  0.50
##            HarmonyI-V-VI    1.607e+00 1.2676046 -0.59  0.33
##            HarmonyIV-I-V    3.070e-03 0.0554095  0.64  0.69 -0.31
##  Subject.2 (Intercept)      1.298e+00 1.1391895
##            Instrumentpiano  1.655e+00 1.2864474 -0.60
##            Instrumentstring 3.337e+00 1.8267211 -0.99  0.68
##  Residual                   2.323e+00 1.5241120
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                              Estimate Std. Error t value
## Voicecontrary                 4.33162    0.51788   8.364
## Voicepar3rd                   4.09784    0.51744   7.919
## Voicepar5th                   4.13071    0.51792   7.976
## Instrumentpiano               1.33826    0.17155   7.801
## Instrumentstring              3.05306    0.23137  13.196
## HarmonyI-V-IV                 0.14533    0.15289   0.951
## HarmonyI-V-VI                 1.19158    0.21410   5.566
## HarmonyIV-I-V                -0.12756    0.15038  -0.848
## ClsListen                     0.25665    0.09324   2.753
## X16.minus.17                 -0.07261    0.04590  -1.582
## X1990s2000s                  -0.23896    0.10476  -2.281
## X1990s2000s.minus.1960s1970s  0.16542    0.09084   1.821
## APTheory.fc1                  0.56990    0.33232   1.715
## Voicepar3rd:HarmonyI-V-IV    -0.34514    0.21319  -1.619
## Voicepar5th:HarmonyI-V-IV    -0.20823    0.21381  -0.974
## Voicepar3rd:HarmonyI-V-VI    -0.76420    0.21326  -3.583
## Voicepar5th:HarmonyI-V-VI    -0.47531    0.21389  -2.222
## Voicepar3rd:HarmonyIV-I-V     0.50026    0.21300   2.349
## Voicepar5th:HarmonyIV-I-V     0.06373    0.21300   0.299
## convergence code: 0
## boundary (singular) fit: see ?isSingular


# Popular ratings

library(MASS)
library(lme4)
lm.2.p <- lm(Popular ~ Harmony + Instrument + Voice + Harmony:Instrument + Ha
rmony:Voice + Instrument:Voice + Harmony:Instrument:Voice, data = ratings)
```

```r
fit2.p <- stepAIC(lm.2.p, trace = 0)
formula(fit2.p)

## Popular ~ Harmony + Instrument

# add random intercept
M0.p <- lmer(Popular ~ Harmony + Instrument + Voice + (1|Subject), REML = F,
data = ratings, control = lmerControl(optimizer = "bobyqa"))

anova(M0.p, fit2.p)

## Data: ratings
## Models:
## fit2.p: Popular ~ Harmony + Instrument
## M0.p: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
##        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## fit2.p  7 10849 10890 -5417.6    10835
## M0.p   10 10160 10218 -5069.8    10140 695.59      3  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

formula(M0.p)

## Popular ~ Harmony + Instrument + Voice + (1 | Subject)

# Conditional residuals
ratings.res <- ratings %>% mutate(r.cond = r.cond(M0.p),
                                  fitted = fitted(M0.p))
ggplot(ratings.res, aes(x=fitted,y=r.cond)) +
  facet_wrap( ~ Subject, as.table=F) + geom_point(pch=1) +
  geom_hline(yintercept = 0, col = "red", linetype="dashed")
```




```
# add random slope
vars.p <- attr(terms(formula(M0.p)),"term.labels")
vars.p <- vars.p[-length(vars.p)]

ratings.2c1.p <- ffRanefLMER.fnc(M0.p, ran.effects= list(slopes=c("Harmony","
Instrument","Voice"), by.vars="Subject", corr=rep(1,length(vars.p))))

##  ===        random slopes       ===
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/ffRanefLMER_lo
g_Sun_Dec_08_01-23-06_2019.txt

formula(ratings.2c1.p)

## Popular ~ Harmony + Instrument + Voice + (1 | Subject)

summary(ratings.2c1.p)

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  10159.6  10217.6  -5069.8  10139.6     2443
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8033 -0.6421  0.0432  0.6782  3.2143
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.460    1.208
##  Residual             3.375    1.837
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                  Estimate Std. Error t value
## (Intercept)       6.61619    0.17843  37.081
## HarmonyI-V-IV    -0.04531    0.10499  -0.432
## HarmonyI-V-VI    -0.30914    0.10497  -2.945
## HarmonyIV-I-V    -0.23613    0.10478  -2.254
## Instrumentpiano  -0.95094    0.09088 -10.464
## Instrumentstring -2.54959    0.09094 -28.037
## Voicepar3rd       0.14830    0.09091   1.631
## Voicepar5th       0.17382    0.09086   1.913
##
## Correlation of Fixed Effects:
##             (Intr) HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r
## HrmnyI-V-IV -0.293
## HrmnyI-V-VI -0.293  0.499
## HrmnyIV-I-V -0.294  0.500  0.500
## Instrumntpn -0.254  0.002  0.002  0.000
## Instrmntstr -0.252  0.000  0.000 -0.002  0.496
## Voicepar3rd -0.255  0.000 -0.001  0.004  0.001    -0.003
## Voicepar5th -0.253 -0.004 -0.006 -0.005 -0.001     0.000     0.500
```

```
##3(a) add fixed effect
# add the first fixed effect
AIC_ls_num.p = list()
for (i in c("Selfdeclare","X16.minus.17","ConsInstr","ConsNotes","Instr.minus
.Notes",
            "PachListen","ClsListen","KnowRob","KnowAxis","X1990s2000s",
            "X1990s2000s.minus.1960s1970s","NoClass","Composing","PianoPlay",
"GuitarPlay",
            "log_OMSI","CollegeMusic.fc","APTheory.fc")){
  lm.HA.1.p <- update(ratings.2c1.p, . ~. + eval(parse(text=i)))
  anov <- anova(ratings.2c1.p, lm.HA.1.p, test = "Chisq")
  AIC_ls_num.p[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls_num.p[which.min(AIC_ls_num.p)]
```

```
## $Composing
```

```
## [1] -2.113404
```

```
 names(AIC_ls_num.p)[which.min(AIC_ls_num.p)] # Composing

## [1] "Composing"

lmer.p.1 <-
  update(ratings.2c1.p, . ~. +
           eval(parse(text=names(AIC_ls_num.p)[which.min(AIC_ls_num.p)])))

# add the second fixed effect
AIC_ls2_num.p = list()
for (i in c("Selfdeclare","X16.minus.17","ConsInstr","ConsNotes","Instr.minus
.Notes",
            "PachListen","ClsListen","KnowRob","KnowAxis","X1990s2000s",
            "X1990s2000s.minus.1960s1970s","NoClass","PianoPlay","GuitarPlay"
,
            "log_OMSI","CollegeMusic.fc","APTheory.fc")){
  lm.HA.2.p <- update(lmer.p.1, . ~. + eval(parse(text=i)))
  anov <- anova(lmer.p.1, lm.HA.2.p, test = "Chisq")
  AIC_ls2_num.p[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls2_num.p[which.min(AIC_ls2_num.p)]

## $X16.minus.17

## [1] -0.5169771

names(AIC_ls2_num.p)[which.min(AIC_ls2_num.p)] # X16.minus.17

## [1] "X16.minus.17"

lmer.p.2 <-
  update(lmer.p.1, . ~. +
           eval(parse(text=names(AIC_ls2_num.p)[which.min(AIC_ls2_num.p)])))

# add the third fixed effect
AIC_ls3_num.p = list()
for (i in c("Selfdeclare","ConsInstr","ConsNotes","Instr.minus.Notes","PachLi
sten",
            "ClsListen","KnowRob","KnowAxis","X1990s2000s","X1990s2000s.minus
.1960s1970s",
            "NoClass","PianoPlay","GuitarPlay","log_OMSI","CollegeMusic.fc","
APTheory.fc")){
  lm.HA.3.p <- update(lmer.p.2, . ~. + eval(parse(text=i)))
  anov <- anova(lmer.p.2, lm.HA.3.p, test = "Chisq")
  AIC_ls3_num.p[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls3_num.p[which.min(AIC_ls3_num.p)]

## $KnowRob
```

```
## [1] -0.4344615

names(AIC_ls3_num.p)[which.min(AIC_ls3_num.p)] # KnowRob

## [1] "KnowRob"

lmer.p.3 <-
  update(lmer.p.2, . ~. + eval(parse(text=names(AIC_ls3_num.p)[which.min(AIC_
ls3_num.p)])))


# add the fourth fixed effect
AIC_ls4_num.p = list()
for (i in c("Selfdeclare","ConsInstr","ConsNotes","Instr.minus.Notes","PachLi
sten",
            "ClsListen","KnowAxis","X1990s2000s","X1990s2000s.minus.1960s1970
s",
            "NoClass","PianoPlay","GuitarPlay","log_OMSI","CollegeMusic.fc",
            "APTheory.fc")){
  lm.HA.4.p <- update(lmer.p.3, . ~. + eval(parse(text=i)))
  anov <- anova(lmer.p.3, lm.HA.4.p, test = "Chisq")
  AIC_ls4_num.p[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls4_num.p[which.min(AIC_ls4_num.p)]

## $X1990s2000s

## [1] -0.08335643

names(AIC_ls4_num.p)[which.min(AIC_ls4_num.p)] # X1990s2000s

## [1] "X1990s2000s"

lmer.p.4 <-
  update(lmer.p.3, . ~. + eval(parse(text=names(AIC_ls4_num.p)[which.min(AIC_
ls4_num.p)])))


# add the fifth fixed effect
AIC_ls5_num.p = list()
for (i in c("Selfdeclare","ConsInstr","ConsNotes","Instr.minus.Notes","PachLi
sten",
            "ClsListen","KnowAxis","X1990s2000s.minus.1960s1970s","NoClass","
PianoPlay",
            "GuitarPlay","log_OMSI","CollegeMusic.fc","APTheory.fc")){
  lm.HA.5.p <- update(lmer.p.4, . ~. + eval(parse(text=i)))
  anov <- anova(lmer.p.4, lm.HA.5.p, test = "Chisq")
  AIC_ls5_num.p[[i]] <- anov[2,2] - anov[1,2]
}
```

```
AIC_ls5_num.p[which.min(AIC_ls5_num.p)]

## $GuitarPlay

## [1] -1.374219

names(AIC_ls5_num.p)[which.min(AIC_ls5_num.p)] # GuitarPlay

## [1] "GuitarPlay"

lmer.p.5 <-
  update(lmer.p.4, . ~. + eval(parse(text=names(AIC_ls5_num.p)[which.min(AIC_
ls5_num.p)])))


# add the sixth fixed effect
AIC_ls6_num.p = list()
for (i in c("Selfdeclare","ConsInstr","ConsNotes","Instr.minus.Notes","PachLi
sten",
            "ClsListen","KnowAxis","X1990s2000s.minus.1960s1970s","NoClass","
PianoPlay",
            "log_OMSI","CollegeMusic.fc","APTheory.fc")){
  lm.HA.6.p <- update(lmer.p.5, . ~. + eval(parse(text=i)))
  anov <- anova(lmer.p.5, lm.HA.6.p, test = "Chisq")
  AIC_ls6_num.p[[i]] <- anov[2,2] - anov[1,2]
}

AIC_ls6_num.p[which.min(AIC_ls6_num.p)] # >0, stop adding fixed effects

## $PachListen

## [1] 0.3728064

lmer.add.fixef.p <- lmer(Popular ~ Harmony + Instrument + Voice + (1 | Subjec
t) +
                           (Harmony | Subject) + (Instrument | Subject) + Com
posing +
                           X16.minus.17 + KnowRob + X1990s2000s + GuitarPlay,
 REML = F,
                         data = ratings, control = lmerControl(optimizer = "b
obyqa"))
summary(lmer.add.fixef.p)

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Composing + X16.minus.17 +
##     KnowRob + X1990s2000s + GuitarPlay
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
```

```
##
##      AIC      BIC   logLik deviance df.resid
##   9726.2   9906.1  -4832.1   9664.2     2422
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0597 -0.5935  0.0281  0.5885  3.3094
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.23773  0.4876
##  Subject.1 (Intercept)      0.27728  0.5266
##            HarmonyI-V-IV    0.08388  0.2896    1.00
##            HarmonyI-V-VI    0.86317  0.9291   -0.43 -0.43
##            HarmonyIV-I-V    0.19836  0.4454   -0.66 -0.66 -0.40
##  Subject.2 (Intercept)      1.04306  1.0213
##            Instrumentpiano  1.37470  1.1725   -0.34
##            Instrumentstring 3.29113  1.8141   -0.65  0.74
##  Residual                   2.39490  1.5475
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                  Estimate Std. Error t value
## (Intercept)       5.52127    0.41668  13.251
## HarmonyI-V-IV    -0.04840    0.09507  -0.509
## HarmonyI-V-VI    -0.30119    0.14209  -2.120
## HarmonyIV-I-V    -0.23234    0.10315  -2.252
## Instrumentpiano  -0.95045    0.15996  -5.942
## Instrumentstring -2.54783    0.23034 -11.061
## Voicepar3rd       0.14361    0.07664   1.874
## Voicepar5th       0.16736    0.07660   2.185
## Composing         0.26507    0.11746   2.257
## X16.minus.17      0.11389    0.04403   2.586
## KnowRob           0.18387    0.08499   2.163
## X1990s2000s       0.16457    0.08379   1.964
## GuitarPlay       -0.24291    0.12157  -1.998
## convergence code: 0
## boundary (singular) fit: see ?isSingular

## automatic
lm.full.fixef.p <- lm(Popular ~ Harmony + Instrument + Voice + Selfdeclare + 
log_OMSI +
                        X16.minus.17 + ConsInstr + ConsNotes + Instr.minus.No
tes +
                        PachListen + ClsListen + KnowRob + KnowAxis + X1990s2
000s +
                        X1990s2000s.minus.1960s1970s + CollegeMusic.fc + NoCl
ass +
                        APTheory.fc + Composing + PianoPlay + GuitarPlay, dat
a=ratings)
```

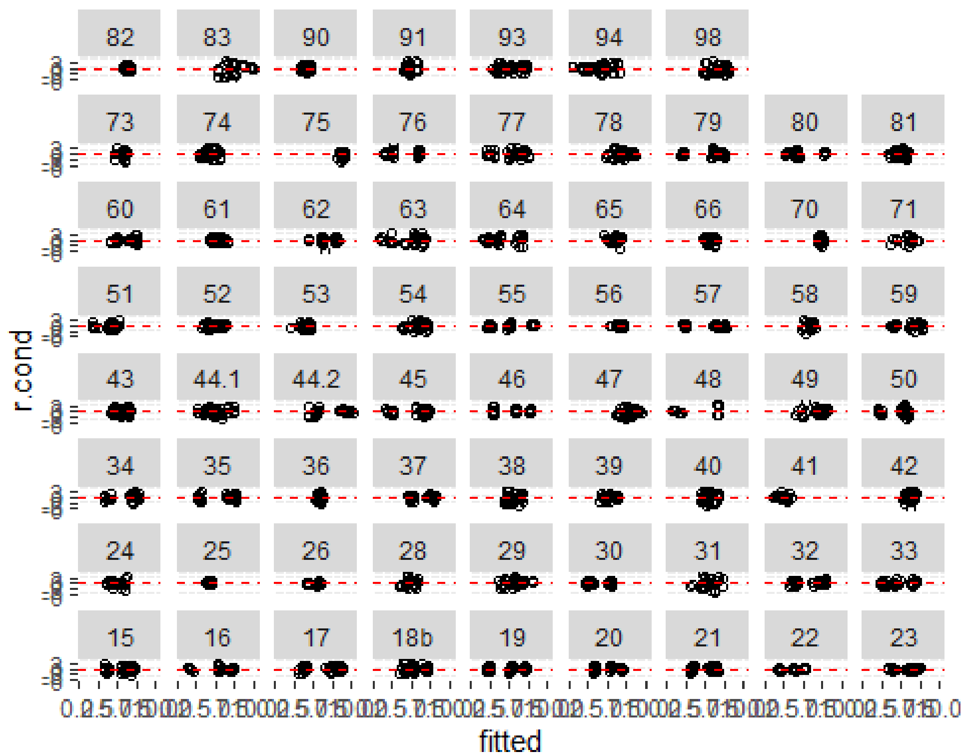
```r
lm.autoadd.fixef.p <- stepAIC(lm.full.fixef.p, direction = "backward", trace
= 0)
summary(lm.autoadd.fixef.p)

##
## Call:
## lm(formula = Popular ~ Harmony + Instrument + Selfdeclare + X16.minus.17 +

##     ConsInstr + ConsNotes + Instr.minus.Notes + PachListen +
##     ClsListen + KnowRob + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     APTheory.fc + Composing + GuitarPlay, data = ratings)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7729 -1.5239  0.1073  1.5425  5.6917
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.77554    0.28177  24.047  < 2e-16 ***
## HarmonyI-V-IV                 -0.04000    0.12143  -0.329 0.741871
## HarmonyI-V-VI                 -0.29818    0.12138  -2.457 0.014097 *
## HarmonyIV-I-V                 -0.23796    0.12118  -1.964 0.049680 *
## Instrumentpiano               -0.95172    0.10500  -9.064  < 2e-16 ***
## Instrumentstring              -2.53005    0.10502 -24.090  < 2e-16 ***
## Selfdeclare                    0.19978    0.05461   3.658 0.000259 ***
## X16.minus.17                   0.11212    0.01662   6.746 1.89e-11 ***
## ConsInstr                      0.16856    0.04638   3.634 0.000285 ***
## ConsNotes                     -0.20783    0.04850  -4.285 1.90e-05 ***
## Instr.minus.Notes             -0.19742    0.04743  -4.163 3.26e-05 ***
## PachListen                    -0.21495    0.04366  -4.923 9.08e-07 ***
## ClsListen                     -0.10244    0.03403  -3.011 0.002635 **
## KnowRob                        0.23170    0.03068   7.551 6.05e-14 ***
## X1990s2000s                    0.14303    0.03672   3.895 0.000101 ***
## X1990s2000s.minus.1960s1970s -0.05308    0.03427  -1.549 0.121488
## APTheory.fc1                   0.25837    0.11897   2.172 0.029969 *
## Composing                      0.10097    0.04572   2.208 0.027324 *
## GuitarPlay                    -0.19027    0.04914  -3.872 0.000111 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.125 on 2434 degrees of freedom
## Multiple R-squared:  0.247,  Adjusted R-squared:  0.2414
## F-statistic: 44.35 on 18 and 2434 DF,  p-value: < 2.2e-16

formula(lm.autoadd.fixef.p)

## Popular ~ Harmony + Instrument + Selfdeclare + X16.minus.17 +
##     ConsInstr + ConsNotes + Instr.minus.Notes + PachListen +
```

```
##     ClsListen + KnowRob + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     APTheory.fc + Composing + GuitarPlay

lmer.autoadd.fixef.p <- lmer(Popular ~ Harmony + Instrument + Voice + Selfdec
lare +
                               X16.minus.17 + ConsInstr + ConsNotes + Instr.m
inus.Notes +
                               PachListen + ClsListen + KnowRob + X1990s2000s
 +
                               X1990s2000s.minus.1960s1970s + APTheory.fc + C
omposing +
                               GuitarPlay + (1 | Subject) + (Harmony | Subjec
t) +
                               (Instrument | Subject), REML = F, data = ratin
gs,
                             control = lmerControl(optimizer = "bobyqa"))
formula(lmer.autoadd.fixef.p)

## Popular ~ Harmony + Instrument + Voice + Selfdeclare + X16.minus.17 +
##     ConsInstr + ConsNotes + Instr.minus.Notes + PachListen +
##     ClsListen + KnowRob + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     APTheory.fc + Composing + GuitarPlay + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject)

anova(lmer.autoadd.fixef.p, lmer.add.fixef.p)

## Data: ratings
## Models:
## lmer.add.fixef.p: Popular ~ Harmony + Instrument + Voice + (1 | Subject) +
 (Harmony |
## lmer.add.fixef.p:     Subject) + (Instrument | Subject) + Composing + X16.
minus.17 +
## lmer.add.fixef.p:     KnowRob + X1990s2000s + GuitarPlay
## lmer.autoadd.fixef.p: Popular ~ Harmony + Instrument + Voice + Selfdeclare
 + X16.minus.17 +
## lmer.autoadd.fixef.p:     ConsInstr + ConsNotes + Instr.minus.Notes + Pach
Listen +
## lmer.autoadd.fixef.p:     ClsListen + KnowRob + X1990s2000s + X1990s2000s.
minus.1960s1970s +
## lmer.autoadd.fixef.p:     APTheory.fc + Composing + GuitarPlay + (1 | Subj
ect) + (Harmony |
## lmer.autoadd.fixef.p:     Subject) + (Instrument | Subject)
##                      Df    AIC    BIC  logLik deviance  Chisq Chi Df
## lmer.add.fixef.p     31 9726.2 9906.1 -4832.1   9664.2
## lmer.autoadd.fixef.p 39 9735.0 9961.4 -4828.5   9657.0 7.1912      8
##                      Pr(>Chisq)
## lmer.add.fixef.p
## lmer.autoadd.fixef.p     0.5161

# p = 0.52, we do not have enough evidence to reject the simpler model
```

```r
##3(b) go back and check random effects
vars <- attr(terms(formula(lmer.add.fixef.p)),"term.labels")
vars <- vars[-c(4,5,6)]
lmer.add.fixef.p.0 <- ffRanefLMER.fnc(lmer.add.fixef.p,
                         ran.effects=
                             list(slopes=vars, by.vars="Subject",
                         corr=rep(0,length(vars))))

##  ===       random slopes       ===
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
## evaluating addition of (0 + Composing|Subject) to model
## evaluating addition of (0 + X16.minus.17|Subject) to model
## evaluating addition of (0 + KnowRob|Subject) to model
## evaluating addition of (0 + X1990s2000s|Subject) to model
## evaluating addition of (0 + GuitarPlay|Subject) to model
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/ffRanefLMER_lo
g_Sun_Dec_08_01-23-20_2019.txt

lmer.add.fixef.p.1 <- ffRanefLMER.fnc(lmer.add.fixef.p,
                         ran.effects=
                             list(slopes=vars, by.vars="Subject",
                                  corr=rep(1,length(vars))))

##  ===       random slopes       ===
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
## evaluating addition of (1 + Composing|Subject) to model
## evaluating addition of (1 + X16.minus.17|Subject) to model
## evaluating addition of (1 + KnowRob|Subject) to model
## evaluating addition of (1 + X1990s2000s|Subject) to model
## evaluating addition of (1 + GuitarPlay|Subject) to model
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/ffRanefLMER_lo
g_Sun_Dec_08_01-25-05_2019.txt

anova(lmer.add.fixef.p.0, lmer.add.fixef.p.1)

## Data: ratings
## Models:
## lmer.add.fixef.p.0: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
 + (Harmony |
## lmer.add.fixef.p.0:     Subject) + (Instrument | Subject) + Composing + X1
6.minus.17 +
## lmer.add.fixef.p.0:     KnowRob + X1990s2000s + GuitarPlay
## lmer.add.fixef.p.1: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
 + (Harmony |
## lmer.add.fixef.p.1:     Subject) + (Instrument | Subject) + Composing + X1
6.minus.17 +
## lmer.add.fixef.p.1:     KnowRob + X1990s2000s + GuitarPlay
```

```
##                     Df    AIC    BIC  logLik deviance Chisq Chi Df
## lmer.add.fixef.p.0 31 9726.2 9906.1 -4832.1   9664.2
## lmer.add.fixef.p.1 31 9726.2 9906.1 -4832.1   9664.2     0      0
##                    Pr(>Chisq)
## lmer.add.fixef.p.0
## lmer.add.fixef.p.1          1

## the models turn out to be the same!
# not add additional random effects

# go back and check fixed effects (lower AIC)
anova(lmer.add.fixef.p, ratings.2c1.p)

## Data: ratings
## Models:
## ratings.2c1.p: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
## lmer.add.fixef.p: Popular ~ Harmony + Instrument + Voice + (1 | Subject) +
 (Harmony |
## lmer.add.fixef.p:     Subject) + (Instrument | Subject) + Composing + X16.
minus.17 +
## lmer.add.fixef.p:     KnowRob + X1990s2000s + GuitarPlay
##                  Df     AIC     BIC  logLik deviance  Chisq Chi Df
## ratings.2c1.p    10 10159.6 10217.6 -5069.8  10139.6
## lmer.add.fixef.p 31  9726.2  9906.1 -4832.1   9664.2 475.39     21
##                  Pr(>Chisq)
## ratings.2c1.p
## lmer.add.fixef.p  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(lmer.add.fixef.p)

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Composing + X16.minus.17 +
##     KnowRob + X1990s2000s + GuitarPlay
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9726.2   9906.1  -4832.1   9664.2     2422
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0597 -0.5935  0.0281  0.5885  3.3094
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.23773  0.4876
##  Subject.1 (Intercept)      0.27728  0.5266
```

```
##              HarmonyI-V-IV    0.08388  0.2896    1.00
##              HarmonyI-V-VI    0.86317  0.9291   -0.43 -0.43
##              HarmonyIV-I-V    0.19836  0.4454   -0.66 -0.66 -0.40
##  Subject.2 (Intercept)       1.04306  1.0213
##              Instrumentpiano  1.37470  1.1725   -0.34
##              Instrumentstring 3.29113  1.8141   -0.65  0.74
##  Residual                    2.39490  1.5475
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)        5.52127    0.41668  13.251
## HarmonyI-V-IV     -0.04840    0.09507  -0.509
## HarmonyI-V-VI     -0.30119    0.14209  -2.120
## HarmonyIV-I-V     -0.23234    0.10315  -2.252
## Instrumentpiano   -0.95045    0.15996  -5.942
## Instrumentstring  -2.54783    0.23034 -11.061
## Voicepar3rd        0.14361    0.07664   1.874
## Voicepar5th        0.16736    0.07660   2.185
## Composing          0.26507    0.11746   2.257
## X16.minus.17       0.11389    0.04403   2.586
## KnowRob            0.18387    0.08499   2.163
## X1990s2000s        0.16457    0.08379   1.964
## GuitarPlay        -0.24291    0.12157  -1.998
## convergence code: 0
## boundary (singular) fit: see ?isSingular

# Conditional residuals
ratings.res <- ratings %>% mutate(r.cond = r.cond(lmer.add.fixef.p),
                          fitted = fitted(lmer.add.fixef.p))
ggplot(ratings.res, aes(x=fitted,y=r.cond)) +
  facet_wrap( ~ Subject, as.table=F) + geom_point(pch=1) +
  geom_hline(yintercept = 0, col = "red", linetype="dashed")
```




```
# summary
q1.p <- lmer(Popular ~ Harmony + Instrument + Voice + (1 | Subject) +
                 (Harmony | Subject) + (Instrument | Subject) + Composing +
                 X16.minus.17 + KnowRob + X1990s2000s + GuitarPlay, REML = F,

               data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.p)

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Composing + X16.minus.17 +
##     KnowRob + X1990s2000s + GuitarPlay
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9726.2   9906.1  -4832.1   9664.2     2422
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0597 -0.5935  0.0281  0.5885  3.3094
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.23773  0.4876
##  Subject.1 (Intercept)      0.27728  0.5266
```

```
##            HarmonyI-V-IV    0.08388  0.2896    1.00
##            HarmonyI-V-VI    0.86317  0.9291   -0.43 -0.43
##            HarmonyIV-I-V    0.19836  0.4454   -0.66 -0.66 -0.40
##  Subject.2 (Intercept)      1.04306  1.0213
##            Instrumentpiano  1.37470  1.1725   -0.34
##            Instrumentstring 3.29113  1.8141   -0.65  0.74
##  Residual                   2.39490  1.5475
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                  Estimate Std. Error t value
## (Intercept)       5.52127    0.41668  13.251
## HarmonyI-V-IV    -0.04840    0.09507  -0.509
## HarmonyI-V-VI    -0.30119    0.14209  -2.120
## HarmonyIV-I-V    -0.23234    0.10315  -2.252
## Instrumentpiano  -0.95045    0.15996  -5.942
## Instrumentstring -2.54783    0.23034 -11.061
## Voicepar3rd       0.14361    0.07664   1.874
## Voicepar5th       0.16736    0.07660   2.185
## Composing         0.26507    0.11746   2.257
## X16.minus.17      0.11389    0.04403   2.586
## KnowRob           0.18387    0.08499   2.163
## X1990s2000s       0.16457    0.08379   1.964
## GuitarPlay       -0.24291    0.12157  -1.998
## convergence code: 0
## boundary (singular) fit: see ?isSingular

# all levels of Instrument
q1.1.p <- lmer(Popular ~ Instrument + Harmony + Voice + (1 | Subject) +
                 (Harmony | Subject) + (Instrument - 1 | Subject) + Composing
 +
                 X16.minus.17 + KnowRob + X1990s2000s + GuitarPlay - 1, REML
= F,
               data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.1.p)

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Harmony |
##     Subject) + (Instrument - 1 | Subject) + Composing + X16.minus.17 +
##     KnowRob + X1990s2000s + GuitarPlay - 1
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9726.2   9906.1  -4832.1   9664.2     2422
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0570 -0.5923  0.0293  0.5882  3.3093
```

```
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.08041  0.2836
##  Subject.1 (Intercept)      0.86320  0.9291
##            HarmonyI-V-IV    0.08319  0.2884    0.58
##            HarmonyI-V-VI    0.86398  0.9295   -0.24 -0.43
##            HarmonyIV-I-V    0.19911  0.4462   -0.37 -0.66 -0.39
##  Subject.2 Instrumentguitar 0.60877  0.7802
##            Instrumentpiano  1.16463  1.0792    0.24
##            Instrumentstring 1.50517  1.2269   -0.61  0.43
##  Residual                  2.39489  1.5475
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  5.51842    0.41658  13.247
## Instrumentpiano   4.56795    0.42644  10.712
## Instrumentstring  2.97058    0.43244   6.869
## HarmonyI-V-IV    -0.04846    0.09501  -0.510
## HarmonyI-V-VI    -0.30122    0.14213  -2.119
## HarmonyIV-I-V    -0.23238    0.10320  -2.252
## Voicepar3rd       0.14361    0.07664   1.874
## Voicepar5th       0.16737    0.07660   2.185
## Composing         0.26561    0.11745   2.261
## X16.minus.17      0.11376    0.04403   2.583
## KnowRob           0.18401    0.08499   2.165
## X1990s2000s       0.16525    0.08378   1.972
## GuitarPlay       -0.24344    0.12157  -2.002
## convergence code: 1
## boundary (singular) fit: see ?isSingular
```

```
# all levels of Harmony
q1.2.p <- lmer(Popular ~ Harmony + Instrument + Voice + (1 | Subject) +
                 (Harmony - 1 | Subject) + (Instrument | Subject) + Composing
 +
                 X16.minus.17 + KnowRob + X1990s2000s + GuitarPlay - 1, REML
= F,
               data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.2.p)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Harmony + Instrument + Voice + (1 | Subject) + (Harmony -
##     1 | Subject) + (Instrument | Subject) + Composing + X16.minus.17 +
##     KnowRob + X1990s2000s + GuitarPlay - 1
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
```

```
##   9726.2   9906.1  -4832.1   9664.2     2422
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0572 -0.5925  0.0292  0.5882  3.3093
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.1435   0.3789
##  Subject.1 HarmonyI-IV-V    0.6981   0.8355
##            HarmonyI-V-IV    1.0933   1.0456   0.98
##            HarmonyI-V-VI    1.1495   1.0721   0.55  0.48
##            HarmonyIV-I-V    0.5921   0.7695   0.85  0.77  0.21
##  Subject.2 (Intercept)      0.7112   0.8433
##            Instrumentpiano  1.3747   1.1725   -0.41
##            Instrumentstring 3.2912   1.8142   -0.78  0.74
##  Residual                   2.3949   1.5475
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                  Estimate Std. Error t value
## HarmonyI-IV-V     5.51860    0.41659  13.247
## HarmonyI-V-IV     5.47014    0.42348  12.917
## HarmonyI-V-VI     5.21738    0.42426  12.298
## HarmonyIV-I-V     5.28622    0.41462  12.750
## Instrumentpiano  -0.95046    0.15995  -5.942
## Instrumentstring -2.54783    0.23034 -11.061
## Voicepar3rd       0.14361    0.07664   1.874
## Voicepar5th       0.16737    0.07660   2.185
## Composing         0.26558    0.11746   2.261
## X16.minus.17      0.11377    0.04403   2.584
## KnowRob           0.18400    0.08499   2.165
## X1990s2000s       0.16521    0.08379   1.972
## GuitarPlay       -0.24341    0.12157  -2.002
## convergence code: 1
## boundary (singular) fit: see ?isSingular
```

```r
# all levels of Voice
q1.3.p <- lmer(Popular ~ Voice + Harmony + Instrument + (1 | Subject) +
                 (Harmony | Subject) + (Instrument | Subject) + Composing +
                 X16.minus.17 + KnowRob + X1990s2000s + GuitarPlay - 1, REML
= F,
               data = ratings, control = lmerControl(optimizer = "bobyqa"))
summary(q1.3.p)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Voice + Harmony + Instrument + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Composing + X16.minus.17 +
##     KnowRob + X1990s2000s + GuitarPlay - 1
```

```
##    Data: ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9726.2   9906.1  -4832.1   9664.2     2422
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0571 -0.5924  0.0293  0.5882  3.3093
##
## Random effects:
##  Groups    Name              Variance Std.Dev. Corr
##  Subject   (Intercept)       0.01785  0.1336
##  Subject.1 (Intercept)       0.79417  0.8912
##            HarmonyI-V-IV     0.08321  0.2885    0.61
##            HarmonyI-V-VI     0.86399  0.9295   -0.25 -0.43
##            HarmonyIV-I-V     0.19909  0.4462   -0.38 -0.66 -0.39
##  Subject.2 (Intercept)       0.74052  0.8605
##            Instrumentpiano   1.37471  1.1725   -0.41
##            Instrumentstring  3.29112  1.8141   -0.77  0.74
##  Residual                    2.39488  1.5475
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                  Estimate Std. Error t value
## Voicecontrary     5.51848    0.41657  13.247
## Voicepar3rd       5.66209    0.41649  13.595
## Voicepar5th       5.68586    0.41667  13.646
## HarmonyI-V-IV    -0.04846    0.09502  -0.510
## HarmonyI-V-VI    -0.30122    0.14213  -2.119
## HarmonyIV-I-V    -0.23238    0.10320  -2.252
## Instrumentpiano  -0.95046    0.15996  -5.942
## Instrumentstring -2.54783    0.23034 -11.061
## Composing         0.26559    0.11745   2.261
## X16.minus.17      0.11377    0.04403   2.584
## KnowRob           0.18400    0.08499   2.165
## X1990s2000s       0.16524    0.08378   1.972
## GuitarPlay       -0.24343    0.12157  -2.002
## convergence code: 1
## boundary (singular) fit: see ?isSingular
```

Q2

```
is.musician <- as.numeric(ratings$Selfdeclare)
is.musician[which(is.musician <= 2)] <- 0
is.musician[which(is.musician > 2)] <- 1
is.musician  <- as.factor(is.musician)
musician.ratings <- ratings
musician.ratings$Selfdeclare <- is.musician
```

```r
# Step BIC on Fixed effect
mlm.final.fix <-
  step(lm(Classical ~ Selfdeclare*(Harmony + Instrument + Voice + log_OMSI +
                                     ClsListen + X16.minus.17 + ConsInstr +
                                     ConsNotes + Instr.minus.Notes + PachList
en +
                                     KnowRob + KnowAxis + APTheory.fc + X1990
s2000s +
                                     X1990s2000s.minus.1960s1970s + X1990s200
0s +
                                     CollegeMusic.fc + NoClass + Composing +
                                     PianoPlay + GuitarPlay), data = musician
.ratings),
                  direction = "backward", trace = 0,
                  k = log(nrow(musician.ratings)))
summary(mlm.final.fix)
```

```
##
## Call:
## lm(formula = Classical ~ Selfdeclare + Harmony + Instrument +
##     Voice + ClsListen + X16.minus.17 + ConsInstr + ConsNotes +
##     Instr.minus.Notes + KnowRob + KnowAxis + APTheory.fc + X1990s2000s +
##     CollegeMusic.fc + Composing + PianoPlay + GuitarPlay + Selfdeclare:Har
mony +
##     Selfdeclare:ClsListen + Selfdeclare:X16.minus.17 + Selfdeclare:Instr.m
inus.Notes +
##     Selfdeclare:KnowRob + Selfdeclare:APTheory.fc + Selfdeclare:CollegeMus
ic.fc +
##     Selfdeclare:PianoPlay + Selfdeclare:GuitarPlay, data = musician.rating
s)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7997 -1.4969  0.0304  1.4014  7.2923
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      4.274505   0.270420  15.807  < 2e-16 ***
## Selfdeclare1                     1.487254   0.338824   4.389 1.19e-05 ***
## HarmonyI-V-IV                   -0.074421   0.154223  -0.483 0.629456
## HarmonyI-V-VI                    0.297395   0.154237   1.928 0.053951 .
## HarmonyIV-I-V                    0.057598   0.153803   0.374 0.708072
## Instrumentpiano                  1.339457   0.102903  13.017  < 2e-16 ***
## Instrumentstring                 3.036443   0.102908  29.506  < 2e-16 ***
## Voicepar3rd                     -0.380936   0.103013  -3.698 0.000222 ***
## Voicepar5th                     -0.359658   0.102952  -3.493 0.000485 ***
## ClsListen                       -0.027528   0.039108  -0.704 0.481568
## X16.minus.17                     0.014929   0.020447   0.730 0.465368
## ConsInstr                       -0.346743   0.049645  -6.984 3.68e-12 ***
## ConsNotes                        0.191911   0.053957   3.557 0.000383 ***
```

```
## Instr.minus.Notes                0.352026   0.055895   6.298 3.57e-10 ***
## KnowRob                          0.249412   0.064009   3.897 0.000100 ***
## KnowAxis                         0.136515   0.029657   4.603 4.38e-06 ***
## APTheory.fc1                     1.937835   0.260055   7.452 1.28e-13 ***
## X1990s2000s                     -0.193493   0.033954  -5.699 1.35e-08 ***
## CollegeMusic.fc1                 0.985775   0.166660   5.915 3.79e-09 ***
## Composing                        0.224405   0.046337   4.843 1.36e-06 ***
## PianoPlay                       -0.500381   0.076931  -6.504 9.45e-11 ***
## GuitarPlay                       2.210948   0.197350  11.203  < 2e-16 ***
## Selfdeclare1:HarmonyI-V-IV       0.076793   0.242369   0.317 0.751390
## Selfdeclare1:HarmonyI-V-VI       1.194450   0.242229   4.931 8.73e-07 ***
## Selfdeclare1:HarmonyIV-I-V      -0.009465   0.241951  -0.039 0.968798
## Selfdeclare1:ClsListen           0.250711   0.074402   3.370 0.000764 ***
## Selfdeclare1:X16.minus.17       -0.383053   0.038792  -9.874  < 2e-16 ***
## Selfdeclare1:Instr.minus.Notes -0.210059   0.062548  -3.358 0.000796 ***
## Selfdeclare1:KnowRob            -0.374644   0.076195  -4.917 9.38e-07 ***
## Selfdeclare1:APTheory.fc1       -1.775444   0.325110  -5.461 5.22e-08 ***
## Selfdeclare1:CollegeMusic.fc1   -1.487293   0.245393  -6.061 1.57e-09 ***
## Selfdeclare1:PianoPlay           0.473797   0.085188   5.562 2.96e-08 ***
## Selfdeclare1:GuitarPlay         -2.281522   0.196080 -11.636  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.082 on 2420 degrees of freedom
## Multiple R-squared:  0.3752, Adjusted R-squared:  0.367
## F-statistic: 45.42 on 32 and 2420 DF,  p-value: < 2.2e-16

as.formula(mlm.final.fix)

## Classical ~ Selfdeclare + Harmony + Instrument + Voice + ClsListen +
##     X16.minus.17 + ConsInstr + ConsNotes + Instr.minus.Notes +
##     KnowRob + KnowAxis + APTheory.fc + X1990s2000s + CollegeMusic.fc +
##     Composing + PianoPlay + GuitarPlay + Selfdeclare:Harmony +
##     Selfdeclare:ClsListen + Selfdeclare:X16.minus.17 + Selfdeclare:Instr.m
inus.Notes +
##     Selfdeclare:KnowRob + Selfdeclare:APTheory.fc + Selfdeclare:CollegeMus
ic.fc +
##     Selfdeclare:PianoPlay + Selfdeclare:GuitarPlay

# random intercept
mlm.semi <- lmer(as.formula(paste("Classical ~",
                                   paste(as.character(formula(mlm.final.fix))
[3],
                                         "(1 | Subject)",
                                         sep = "+"))),
                  lmerControl(optimizer = "bobyqa"),  REML = F, data = musici
an.ratings)

exactLRT(mlm.semi, mlm.final.fix)
```

```
##
##  simulated finite sample distribution of LRT. (p-value based on
##  10000 simulated values)
##
## data:
## LRT = 410.3, p-value < 2.2e-16
```

```
summary(mlm.semi)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Selfdeclare + Harmony + Instrument + Voice + ClsListen +
##     X16.minus.17 + ConsInstr + ConsNotes + Instr.minus.Notes +
##     KnowRob + KnowAxis + APTheory.fc + X1990s2000s + CollegeMusic.fc +
##     Composing + PianoPlay + GuitarPlay + Selfdeclare:Harmony +
##     Selfdeclare:ClsListen + Selfdeclare:X16.minus.17 + Selfdeclare:Instr.m
inus.Notes +
##     Selfdeclare:KnowRob + Selfdeclare:APTheory.fc + Selfdeclare:CollegeMus
ic.fc +
##     Selfdeclare:PianoPlay + Selfdeclare:GuitarPlay + (1 | Subject)
##    Data: musician.ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  10185.3  10388.5  -5057.7  10115.3     2418
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2308 -0.6216 -0.0154  0.6412  3.7578
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.8822   0.9393
##  Residual             3.3866   1.8403
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                                Estimate Std. Error t value
## (Intercept)                    4.214966   0.677893   6.218
## Selfdeclare1                   1.495439   0.875074   1.709
## HarmonyI-V-IV                 -0.076385   0.136340  -0.560
## HarmonyI-V-VI                  0.311125   0.136393   2.281
## HarmonyIV-I-V                  0.061116   0.135965   0.449
## Instrumentpiano                1.348845   0.091033  14.817
## Instrumentstring               3.039591   0.091089  33.369
## Voicepar3rd                   -0.384239   0.091064  -4.219
## Voicepar5th                   -0.357336   0.091014  -3.926
## ClsListen                     -0.019069   0.109282  -0.174
## X16.minus.17                  -0.003133   0.054330  -0.058
## ConsInstr                     -0.354193   0.140482  -2.521
```

```
## ConsNotes                          0.194722   0.152434   1.277
## Instr.minus.Notes                  0.336549   0.157157   2.141
## KnowRob                            0.244597   0.181828   1.345
## KnowAxis                           0.137128   0.084072   1.631
## APTheory.fc1                       1.946655   0.736875   2.642
## X1990s2000s                       -0.178893   0.094812  -1.887
## CollegeMusic.fc1                   0.985621   0.473023   2.084
## Composing                          0.223801   0.130234   1.718
## PianoPlay                         -0.489821   0.217257  -2.255
## GuitarPlay                         2.194759   0.558415   3.930
## Selfdeclare1:HarmonyI-V-IV         0.081159   0.214251   0.379
## Selfdeclare1:HarmonyI-V-VI         1.182255   0.214152   5.521
## Selfdeclare1:HarmonyIV-I-V        -0.012886   0.213880  -0.060
## Selfdeclare1:ClsListen             0.246706   0.209987   1.175
## Selfdeclare1:X16.minus.17         -0.359163   0.107309  -3.347
## Selfdeclare1:Instr.minus.Notes -0.190110   0.175599  -1.083
## Selfdeclare1:KnowRob              -0.368477   0.216710  -1.700
## Selfdeclare1:APTheory.fc1         -1.768150   0.917616  -1.927
## Selfdeclare1:CollegeMusic.fc1  -1.500499   0.696278  -2.155
## Selfdeclare1:PianoPlay             0.459623   0.240378   1.912
## Selfdeclare1:GuitarPlay           -2.271040   0.555482  -4.088
```

```
as.formula(mlm.semi)
```

```
## Classical ~ Selfdeclare + Harmony + Instrument + Voice + ClsListen +
##     X16.minus.17 + ConsInstr + ConsNotes + Instr.minus.Notes +
##     KnowRob + KnowAxis + APTheory.fc + X1990s2000s + CollegeMusic.fc +
##     Composing + PianoPlay + GuitarPlay + Selfdeclare:Harmony +
##     Selfdeclare:ClsListen + Selfdeclare:X16.minus.17 + Selfdeclare:Instr.m
inus.Notes +
##     Selfdeclare:KnowRob + Selfdeclare:APTheory.fc + Selfdeclare:CollegeMus
ic.fc +
##     Selfdeclare:PianoPlay + Selfdeclare:GuitarPlay + (1 | Subject)
```

```
# drop insignificant fixed effects
mlm.semi.2 <- lme4::lmer(Classical ~ Voice + Harmony + Instrument +
                     ConsInstr  + Instr.minus.Notes + ClsListen  + CollegeMus
ic +
                     APTheory + PianoPlay + GuitarPlay + Selfdeclare:Harmony
+
                     Selfdeclare:X16.minus.17 + Selfdeclare:CollegeMusic +
                     Selfdeclare:GuitarPlay + (1 | Subject),
                  lmerControl(optimizer = "bobyqa"),  REML = F, data = musici
an.ratings)
exactRLRT(mlm.semi.2) # include random intercept
```

```
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
```

```
## data:
## RLRT = 589.43, p-value < 2.2e-16
```

```
# Random slope
vars <- attr(terms(formula(mlm.semi.2)),"term.labels")
vars <- vars[c(1:10)]

mlm.final.0 <- ffRanefLMER.fnc(mlm.semi.2,
                         ran.effects=
                             list(slopes=vars, by.vars="Subject",
                         corr=rep(0,length(vars))))
```

```
##  ===       random slopes       ===
## evaluating addition of (Voice|Subject) to model
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 2.481122e-78
##  adding (Instrument | Subject) to model
## evaluating addition of (0 + ConsInstr|Subject) to model
## evaluating addition of (0 + Instr.minus.Notes|Subject) to model
## evaluating addition of (0 + ClsListen|Subject) to model
## evaluating addition of (0 + CollegeMusic|Subject) to model
## evaluating addition of (0 + APTheory|Subject) to model
## evaluating addition of (0 + PianoPlay|Subject) to model
## evaluating addition of (0 + GuitarPlay|Subject) to model
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/ffRanefLMER_lo
g_Sun_Dec_08_01-27-19_2019.txt
```

```
as.formula(mlm.final.0)
```

```
## Classical ~ Voice + Harmony + Instrument + ConsInstr + Instr.minus.Notes +

##     ClsListen + CollegeMusic + APTheory + PianoPlay + GuitarPlay +
##     (1 | Subject) + (Instrument | Subject) + Harmony:Selfdeclare +
##     Selfdeclare:X16.minus.17 + CollegeMusic:Selfdeclare + GuitarPlay:Selfd
eclare
```

```
mlm.final.1 <- ffRanefLMER.fnc(mlm.semi.2,
                         ran.effects=
                             list(slopes=vars, by.vars="Subject",
                                  corr=rep(1,length(vars))))
```

```
##  ===       random slopes       ===
## evaluating addition of (Voice|Subject) to model
## evaluating addition of (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 2.481122e-78
##  adding (Instrument | Subject) to model
## evaluating addition of (1 + ConsInstr|Subject) to model
## evaluating addition of (1 + Instr.minus.Notes|Subject) to model
## evaluating addition of (1 + ClsListen|Subject) to model
```

```
## evaluating addition of (1 + CollegeMusic|Subject) to model
##  log-likelihood ratio test p-value = 0.9027076
##  not adding (1+CollegeMusic | Subject) to model
## evaluating addition of (1 + APTheory|Subject) to model
##  log-likelihood ratio test p-value = 0.9975724
##  not adding (1+APTheory | Subject) to model
## evaluating addition of (1 + PianoPlay|Subject) to model
## evaluating addition of (1 + GuitarPlay|Subject) to model
## log file is C:\Users\YINGZH~1\AppData\Local\Temp\Rtmpofgae6/ffRanefLMER_lo
g_Sun_Dec_08_01-27-28_2019.txt
```

```
as.formula(mlm.final.1)
```

```
## Classical ~ Voice + Harmony + Instrument + ConsInstr + Instr.minus.Notes +

##     ClsListen + CollegeMusic + APTheory + PianoPlay + GuitarPlay +
##     (1 | Subject) + (Instrument | Subject) + Harmony:Selfdeclare +
##     Selfdeclare:X16.minus.17 + CollegeMusic:Selfdeclare + GuitarPlay:Selfd
eclare
```

```
anova(mlm.final.0, mlm.final.1)
```

```
## Data: musician.ratings
## Models:
## mlm.final.0: Classical ~ Voice + Harmony + Instrument + ConsInstr + Instr.
minus.Notes +
## mlm.final.0:     ClsListen + CollegeMusic + APTheory + PianoPlay + GuitarP
lay +
## mlm.final.0:     (1 | Subject) + (Instrument | Subject) + Harmony:Selfdecl
are +
## mlm.final.0:     Selfdeclare:X16.minus.17 + CollegeMusic:Selfdeclare + Gui
tarPlay:Selfdeclare
## mlm.final.1: Classical ~ Voice + Harmony + Instrument + ConsInstr + Instr.
minus.Notes +
## mlm.final.1:     ClsListen + CollegeMusic + APTheory + PianoPlay + GuitarP
lay +
## mlm.final.1:     (1 | Subject) + (Instrument | Subject) + Harmony:Selfdecl
are +
## mlm.final.1:     Selfdeclare:X16.minus.17 + CollegeMusic:Selfdeclare + Gui
tarPlay:Selfdeclare
##             Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## mlm.final.0 31 9818.7 9998.7 -4878.4   9756.7
## mlm.final.1 31 9818.7 9998.7 -4878.4   9756.7     0      0          1
```

```
summary(mlm.final.1)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Voice + Harmony + Instrument + ConsInstr + Instr.minus.Notes +

##     ClsListen + CollegeMusic + APTheory + PianoPlay + GuitarPlay +
```

```
##     (1 | Subject) + (Instrument | Subject) + Harmony:Selfdeclare +
##     Selfdeclare:X16.minus.17 + CollegeMusic:Selfdeclare + GuitarPlay:Selfd
eclare
##    Data: musician.ratings
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   9818.7   9998.7  -4878.4   9756.7     2422
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3049 -0.5960  0.0274  0.5765  3.9624
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.1888   0.4345
##  Subject.1 (Intercept)      2.1477   1.4655
##            Instrumentpiano  1.6126   1.2699   -0.55
##            Instrumentstring 3.2820   1.8116   -0.75  0.68
##  Residual                   2.6233   1.6197
## Number of obs: 2453, groups:  Subject, 70
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                   3.93467    0.51496   7.641
## Voicepar3rd                  -0.38537    0.08019  -4.806
## Voicepar5th                  -0.35451    0.08014  -4.424
## HarmonyI-V-IV                -0.06867    0.12006  -0.572
## HarmonyI-V-VI                 0.31244    0.12013   2.601
## HarmonyIV-I-V                 0.06286    0.11970   0.525
## Instrumentpiano               1.33832    0.17193   7.784
## Instrumentstring              3.05479    0.23128  13.208
## ConsInstr                    -0.11323    0.09152  -1.237
## Instr.minus.Notes             0.15774    0.08768   1.799
## ClsListen                     0.11480    0.09213   1.246
## CollegeMusic                  0.49834    0.45852   1.087
## APTheory                      0.55752    0.40969   1.361
## PianoPlay                    -0.10005    0.09928  -1.008
## GuitarPlay                    1.27373    0.49222   2.588
## HarmonyI-IV-V:Selfdeclare1  1.35393    0.70213   1.928
## HarmonyI-V-IV:Selfdeclare1  1.42632    0.70214   2.031
## HarmonyI-V-VI:Selfdeclare1  2.53438    0.70226   3.609
## HarmonyIV-I-V:Selfdeclare1  1.34232    0.70220   1.912
## Selfdeclare0:X16.minus.17  -0.05006    0.05709  -0.877
## Selfdeclare1:X16.minus.17  -0.22446    0.08599  -2.610
## CollegeMusic:Selfdeclare1  -1.18048    0.70735  -1.669
## GuitarPlay:Selfdeclare1    -1.30009    0.51065  -2.546
```