36-763 HLM (F13) hw5

BJ Kim (bjkim@cmu.edu)

Question 1

1. The three main experimental factors.

1-(a)

Examine the influence of the three main experimental factors (Instrument, Harmony & Voice) on Classical ratings, using conventional linear models and/or analysis of variance models. Comment briefly on your findings, providing suitable brief evidence for each result. Hint: To determine whether Harmony is important, for example, one might compare the fit of a model with Harmony in it, to one without Harmony. To determine how particular kinds of harmony affect ratings, one might begin by looking at fixed effects estimates in a suitable model. Etc.

```
ratings<-read.csv("ratings.csv")
str(ratings)
attach(ratings)
m1<-lm(Classical ~ Instrument + Voice)
Coefficients:
  Estimate Std. Error t value Pr(>|t|)
(Intercept)        4.5367     0.1038  43.724  < 2e-16 ***
  Instrumentpiano    1.3730     0.1141  12.035  < 2e-16 ***
  Instrumentstring   3.1334     0.1134  27.631  < 2e-16 ***
  Voicepar3rd       -0.4134     0.1138  -3.633 0.000286 ***
  Voicepar5th       -0.3690     0.1137  -3.244 0.001193 **

m2<-lm(Classical ~ Instrument + Harmony + Voice)
Coefficients:
  Estimate Std. Error t value Pr(>|t|)
  (Intercept)       4.34016    0.12987  33.420  < 2e-16 ***
  Instrumentpiano   1.37359    0.11298  12.158  < 2e-16 ***
  Instrumentstring  3.13312    0.11230  27.899  < 2e-16 ***
  HarmonyI-V-IV    -0.03108    0.13008  -0.239 0.811168
  HarmonyI-V-VI     0.76909    0.13008   5.913 3.83e-09 ***
  HarmonyIV-I-V     0.05007    0.12997   0.385 0.700092
  Voicepar3rd      -0.41247    0.11271  -3.660 0.000258 ***
  Voicepar5th      -0.37058    0.11264  -3.290 0.001016 **
  ---

anova(m1, m2)
Model 1: Classical ~ Instrument + Voice
Model 2: Classical ~ Instrument + Harmony + Voice
Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2488 13381
2   2485 13108  3    273.65 17.293 4.107e-11 ***

is.factor(Subject)

m3<-lm(Classical ~ Instrument + Harmony + Voice + Subject)
```

```
Coefficients:
  Estimate Std. Error t value Pr(>|t|)
  (Intercept)       4.06324    0.33091  12.279  < 2e-16 ***
  Instrumentpiano   1.37737    0.09319  14.780  < 2e-16 ***
  Instrumentstring  3.13154    0.09257  33.828  < 2e-16 ***
  HarmonyI-V-IV    -0.03262    0.10718  -0.304 0.760924
  HarmonyI-V-VI     0.77106    0.10718   7.194 8.36e-13 ***
  HarmonyIV-I-V     0.04988    0.10709   0.466 0.641432
  Voicepar3rd      -0.41523    0.09287  -4.471 8.14e-06 ***
  Voicepar5th      -0.37464    0.09281  -4.037 5.59e-05 ***
  Subject16         0.36111    0.44603   0.810 0.418241
  Subject17        -1.13889    0.44603  -2.553 0.010728 *
  ....
  Subject94        -0.47222    0.44603  -1.059 0.289829
  Subject98        -0.47222    0.44603  -1.059 0.289829
  ---

anova(m2, m3)
Model 1: Classical ~ Instrument + Harmony + Voice
Model 2: Classical ~ Instrument + Harmony + Voice + Subject
Res.Df     RSS Df Sum of Sq      F    Pr(>F)
1   2485 13107.5
2   2416  8651.5 69    4455.9 18.034 < 2.2e-16 ***
  ---

anova(m1, m2, m3)
Analysis of Variance Table

Model 1: Classical ~ Instrument + Voice
Model 2: Classical ~ Instrument + Harmony + Voice
Model 3: Classical ~ Instrument + Harmony + Voice + Subject
  Res.Df     RSS Df Sum of Sq      F    Pr(>F)
1   2488 13381.1
2   2485 13107.5  3     273.6 25.473 3.202e-16 ***
3   2416  8651.5 69    4455.9 18.034 < 2.2e-16 ***
```

As shown above, model 2 including `Harmony` variable is significantly different from model 1. And one of the `Harmony` dummy is significant in model 2. Thus model 2 is better than model 1. And we can interpret model 2 as that using piano and strings as well as a particular harmony (I-V-VI) have positive association with classical ratings, while Voice has negative association. When we add `Subject` for controlling for individual fixed effect, all the statistically significant coefficients have similar effect size with same direction. In `anova()`, we also found that model 3 is better than model 2. Overall, we saw that `Harmony` variable needs to be included in the model.

1-(b)

Since we have approximately 36 ratings from each participant, we can fit a random intercept for each participant if we wish. Such a model is called a "repeated measures" model.

1-(b)-i

Carefully write this model in mathematical terms as a hierarchical linear model.

$$ratings_i = \alpha_{0j[i]} + \alpha_1 Harmony_i + \alpha_2 Instrument_i + \alpha_3 Voice_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2)$$
$$\alpha_{0j} = \beta_0 + \eta_j, \quad \eta_j \sim N(0, \tau^2)$$

Thus

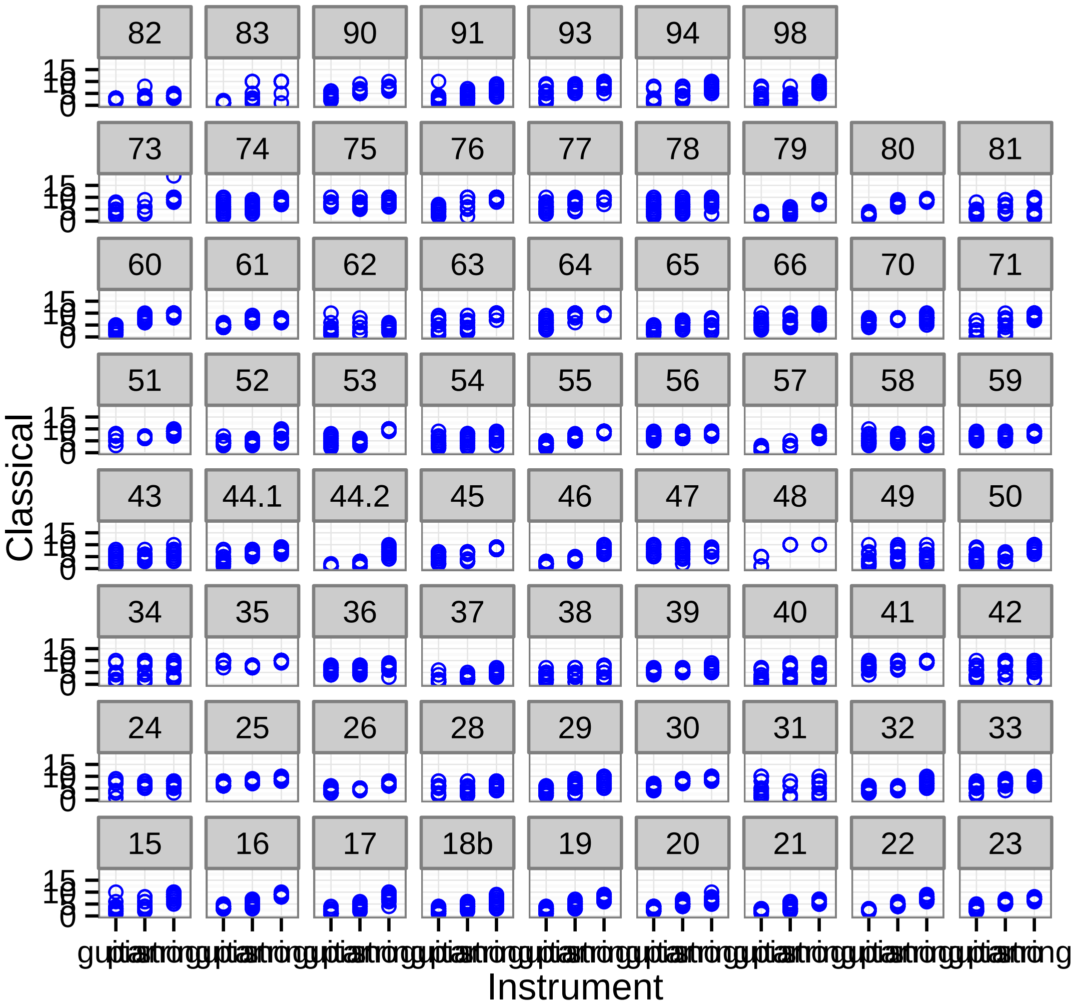
$$ratings_i = \beta_0 + \alpha_1 Harmony_i + \alpha_2 Instrument_i + \alpha_3 Voice_i + \eta_{j[i]} + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2)$$
$$\eta_j \sim N(0, \tau^2)$$

1-(b)-ii

Use at least two different methods to test whether the random intercept is needed in the model. Is the random effect needed? Justify your answer with evidence from your tests. [1]

```
library(ggplot2); theme_set(theme_bw())
library(arm)
library(lme4)
attach(ratings)

# comparing the models

# pooled regression graph
plot(Classical ~ 1,ylab="Classical ratings")
abline(lm(Classical ~ 1), col="red")

# unpooled, means only, with Instrument as X variable
ggplot(ratings,aes(x=Instrument,y=Classical)) +
  geom_point(pch=1,color="blue") +
  geom_smooth(method="lm", formula = Classical ~ 1, se=F, size=0.5,
              fullrange=T, color="black") +
  #scale_x_continuous(labels=NULL) +
  facet_wrap( ~ Subject,as.table=F)

# unpooled, means only, with Instrument as X variable
ggplot(ratings,aes(x=Instrument,y=Classical)) +
  geom_point(pch=1,color="blue") +
  geom_smooth(method="lm", formula = Classical ~ 1, se=F, size=0.5,
              fullrange=T, color="black") +
  #scale_x_continuous(labels=NULL) +
  facet_wrap( ~ Subject,as.table=F)

# comparing the models
summary(newm1<-lm(Classical ~ 1))
is.factor(Subject)
[1] TRUE
summary(newm2<-lm(Classical ~ Subject))

anova(newm1, newm2)
```

[1] 03-, 04-.r from lecture materials



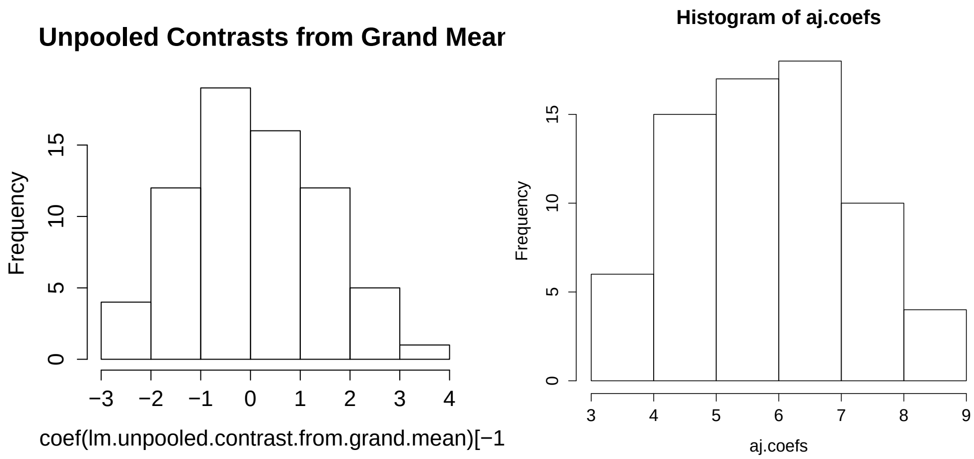

```
Model 1: Classical ~ 1
Model 2: Classical ~ Subject
Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    2492 17595
2    2423 13132 69    4462.1 11.931 < 2.2e-16 ***

is.factor(Subject)
contrasts(Subject) <- contr.sum(70)
lm.unpooled.contrast.from.grand.mean <- lm(Classical ~  Subject)

summary(newm1)$coef

summary(lm.unpooled.contrast.from.grand.mean)$coef

anova(newm1,lm.unpooled.contrast.from.grand.mean)
Analysis of Variance Table
```

```
Model 1: Classical ~ 1
Model 2: Classical ~ Subject
Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2492 17595
2   2423 13132 69    4462.1 11.931 < 2.2e-16 ***

hist(coef(lm.unpooled.contrast.from.grand.mean)[-1],
     main="Unpooled Contrasts from Grand Mean")

############################################################

# How many Subjects have Classical means significantly different from
# the grand mean?

# forces Subject coefficients to sum to zero, so their
# values show how different each county mean is from the
# grand mean...

lm.unpooled.contrast.from.grand.mean <- lm(Classical ~ Subject)
summary(lm.unpooled.contrast.from.grand.mean)

length(unique(Subject))
[1] 70

sum(coef(summary(lm.unpooled.contrast.from.grand.mean))[,4]<0.05)
[1] 43

43/70
[1] 0.6142857

# hierarchical structure

aj.coefs <- NULL
for (Subj in sort(unique(Subject))) {
  aj.coefs <- c(aj.coefs,coef(lm(Classical ~ 1,subset=(Subject==Subj))))
}

hist(aj.coefs)
```

1-(b)-iii

Re-examine the influence of the three main experimental factors (Instrument, Harmony & Voice) on Classical ratings, using the repeated-measures model with the random intercept for participants.

```
lmer.intercept.only <- lmer( Classical ~ 1 + ( 1 | Subject ) )

summary(lmer.intercept.only)
Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ 1 + (1 | Subject)

REML criterion at convergence: 11462.08
```




```
Random effects:
  Groups   Name        Variance Std.Dev.
Subject  (Intercept) 1.654    1.286
Residual             5.420    2.328
Number of obs: 2493, groups: Subject, 70

Fixed effects:
  Estimate Std. Error t value
(Intercept)   5.7872     0.1607   36.02

fixef(lmer.intercept.only)
(Intercept)
5.787247

ranef(lmer.intercept.only)


lmer1<-lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject))
display(lmer1)
lmer(formula = Classical ~ Instrument + Harmony + Voice + (1 |
    Subject))
                 coef.est coef.se
(Intercept)       4.34     0.19
Instrumentpiano   1.38     0.09
Instrumentstring  3.13     0.09
HarmonyI-V-IV    -0.03     0.11
HarmonyI-V-VI     0.77     0.11
HarmonyIV-I-V     0.05     0.11
Voicepar3rd      -0.42     0.09
Voicepar5th      -0.37     0.09

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

```
 Subject  (Intercept) 1.30
 Residual             1.89
---
number of obs: 2493, groups: Subject, 70
AIC = 10491.5, DIC = 10426.2
deviance = 10448.9
```

1-(c)

The random intercept in a repeated measures model can account for "personal biases" in ratings: 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. Alternatively, 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". This suggests, e.g., a random effect of the form `(1 | Subject:Instrument)`: a random draw is made from a single normal distribution, for each person/instrument combination. One could argue for a similar random effect for each person/harmony combination, and for each person/voice leading combination.

1-(c)-i

Determine whether a model with all three new random effect terms (but not the original single random intercept) is better or worse than each of the models in problems 1a and 1b. Provide suitable evidence to justify your answer.

```
lmer1<-lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject))
lmer2<-lmer(Classical ~ Instrument + Harmony + Voice + (Instrument + Harmony + Voice |
    Subject))
lmer3<-lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument))
lmer4<-lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject:Harmony))
lmer5<-lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject:Voice))

anova(lmer1, lmer2, lmer3, lmer4, lmer5)
Data:
Models:
lmer1: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
lmer3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument)
lmer4: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Harmony)
lmer5: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Voice)
lmer2: Classical ~ Instrument + Harmony + Voice + (Instrument + Harmony + Voice | Subject
    )
      Df     AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer1 10 10468.9 10527 -5224.4  10448.9
lmer3 10 10153.3 10212 -5066.6  10133.3 315.61      0     <2e-16 ***
lmer4 10 10613.4 10672 -5296.7  10593.4   0.00      0          1
lmer5 10 10691.7 10750 -5335.8  10671.7   0.00      0          1
lmer2 45  9971.1 10233 -4940.6   9881.1 790.57     35     <2e-16 ***
---

cbind(
  AIC=sapply(list(lmer1=lmer1, lmer2=lmer2, lmer3=lmer3, lmer4=lmer4, lmer5=lmer5, lmer2b
    =lmer2b, m3),AIC)
  ,
```

```
  DIC=sapply(list(lmer1=lmer1, lmer2=lmer2, lmer3=lmer3, lmer4=lmer4, lmer5=lmer5, lmer2b
    =lmer2b, m3), invisible(function(x) display(x)$DIC))
  ,
  BIC=sapply(list(lmer1=lmer1, lmer2=lmer2, lmer3=lmer3, lmer4=lmer4, lmer5=lmer5, lmer2b
    =lmer2b, m3),BIC)
)
       AIC      DIC      BIC
lmer1  10491.51 10426.21 10549.73
lmer2  10062.51 9789.693 10324.46
lmer3  10173.45 10113.06 10231.66
lmer4  10632.19 10574.56 10690.4
lmer5  10711.53 10651.81 10769.74
lmer2b 10034.59 9817.616 10296.54
       10332.74 NULL     10786.8
```

As shown AIC and BIC in the ANOVA test above, `lmer2` using random effects is better than `lmer1` from Question 1 (b) and other three models (lmer 3, 4, 5). (Although BIC for lmer3 is lowest, considering deviance and DIC, I would choose lmer2 as the best model among those.)

However `lmer2` contains random effect of intercept. So we need to remove it.

```
lmer2a <- lmer(Classical ~ Instrument + Harmony + Voice + (Instrument + Harmony + Voice -
     1 | Subject))
lmer2b <- lmer(Classical ~ Instrument + Harmony + Voice + (0 + Instrument + Harmony +
    Voice | Subject))
display(lmer2a)
display(lmer2b)

lmer(formula = Classical ~ Instrument + Harmony + Voice + (0 +
    Instrument + Harmony + Voice | Subject))
                 coef.est coef.se
(Intercept)       4.34     0.28
Instrumentpiano   1.37     0.19
Instrumentstring  3.13     0.28
HarmonyI-V-IV    -0.03     0.10
HarmonyI-V-VI     0.77     0.18
HarmonyIV-I-V     0.05     0.11
Voicepar3rd      -0.41     0.11
Voicepar5th      -0.37     0.10

Error terms:
 Groups   Name             Std.Dev. Corr
 Subject  Instrumentguitar 2.23
          Instrumentpiano  1.91      0.77
          Instrumentstring 1.77      0.40  0.55
          HarmonyI-V-IV    0.47      0.51  0.28  0.47
          HarmonyI-V-VI    1.29      0.32  0.10 -0.21  0.18
          HarmonyIV-I-V    0.55     -0.02  0.00  0.48  0.28 -0.08
          Voicepar3rd      0.69     -0.20  0.02  0.22 -0.17 -0.37  0.76
          Voicepar5th      0.55     -0.16  0.03  0.08 -0.40 -0.19  0.61  0.92
 Residual                  1.52
---
number of obs: 2493, groups: Subject, 70
AIC = 10034.6, DIC = 9817.6
```

```
deviance = 9881.1

anova(lmer1,lmer2,lmer3,lmer4,lmer5,lmer2a)
Data:
Models:
lmer1: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
lmer3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument)
lmer4: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Harmony)
lmer5: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Voice)
lmer2: Classical ~ Instrument + Harmony + Voice + (Instrument + Harmony +
lmer2:     Voice | Subject)
lmer2a: Classical ~ Instrument + Harmony + Voice + (Instrument + Harmony +
lmer2a:     Voice - 1 | Subject)
       Df     AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer1  10 10468.9 10527 -5224.4  10448.9
lmer3  10 10153.3 10212 -5066.6  10133.3 315.61      0     <2e-16 ***
lmer4  10 10613.4 10672 -5296.7  10593.4   0.00      0          1
lmer5  10 10691.7 10750 -5335.8  10671.7   0.00      0          1
lmer2  45  9971.1 10233 -4940.6   9881.1 790.57     35     <2e-16 ***
lmer2a 45  9971.1 10233 -4940.6   9881.1   0.00      0          1
---
Signif. codes:  0  ***   0.001   **   0.01   *   0.05   .   0.1        1
```

`lmer2a` and `lmer2b` provides the same results in ANOVA. However, variance components are different between `lmer2` and `lmer2a / lmer2b` because `lmer2a / lmer2b` removed random intercept effect. Thus we will use `lmer2b` (which is the same as `lmer2a`) hereafter.

In addition, we compare `lme2b` with my final linear model from the question 1-(a), `m3`.

```
# Comparing lme() model with lm() model
# my lm() model is m3; my lmer() model based on the results above is lmer2
# Compare lmer2b with m3

LRT.observed <- as.numeric(2*(logLik(lmer2b) - logLik(m3)))
nsim <- 9
LRT.sim <- numeric(nsim)
for (i in 1:nsim) {
  y <- unlist(simulate(m3))
  nullmod <- lm(Classical ~ Instrument + Harmony + Voice + Subject)
  altmod <- lmer(Classical ~ Instrument + Harmony + Voice + (0 + Instrument + Harmony +
    Voice | Subject))
  LRT.sim[i] <- as.numeric(2*(logLik(altmod) - logLik(nullmod)))
}
mean(LRT.sim > LRT.observed) #pvalue
```

1-(c)-ii

Re-examine the influence of the three main experimental factors (Instrument, Harmony & Voice) on Classical ratings, using the model with all three new random effect terms in it. Comment briefly on your findings, providing suitable brief evidence for each result. In addition, comment on the sizes of the three estimated variance components, with respect to each other and with respect to the estimated residual variance.

```
display(lmer2b)
lmer(formula = Classical ~ Instrument + Harmony + Voice + (0 +
```

```
    Instrument + Harmony + Voice | Subject))
                 coef.est coef.se
(Intercept)       4.34     0.28
Instrumentpiano   1.37     0.19
Instrumentstring  3.13     0.28
HarmonyI-V-IV    -0.03     0.10
HarmonyI-V-VI     0.77     0.18
HarmonyIV-I-V     0.05     0.11
Voicepar3rd      -0.41     0.11
Voicepar5th      -0.37     0.10

Error terms:
 Groups   Name             Std.Dev. Corr
 Subject  Instrumentguitar 2.23
          Instrumentpiano  1.91      0.77
          Instrumentstring 1.77      0.40  0.55
          HarmonyI-V-IV    0.47      0.51  0.28  0.47
          HarmonyI-V-VI    1.29      0.32  0.10 -0.21  0.18
          HarmonyIV-I-V    0.55     -0.02  0.00  0.48  0.28 -0.08
          Voicepar3rd      0.69     -0.20  0.02  0.22 -0.17 -0.37  0.76
          Voicepar5th      0.55     -0.16  0.03  0.08 -0.40 -0.19  0.61  0.92
 Residual                  1.52
---
number of obs: 2493, groups: Subject, 70
AIC = 10034.6, DIC = 9817.6
deviance = 9881.1
```

As shown above, in this `lmer2b` model, the coefficients are similar to previous `lm()` model. All the `Instrument` and `Voice` variables are significant, and one of the `Harmony` variable is significant. Two insignificant Harmony variables are shown to be highly correlated in terms of their error terms. Also we can see that all three Instrument variables (piano and string) have larger standard deviation than residual standard deviation, and Harmony (I-V-VI) has also large standard deviation.

1-(c)-iii

Carefully write this model in mathematical terms as a hierarchical linear model. Because they are design variables in the experiment, the three experimental factors, Instrument, Harmony, and Voice, should be included in all models for the remainder of this homework, regardless of what you found about their influence or lack of influence on ratings.

$$ratings_i = \alpha_0 + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i + \alpha_{3j[i]} Voice_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2)$$
$$\alpha_{1j[i]} = \beta_1 + \eta_{1j}, \quad \eta_{1j} \sim N(0, \tau_1^2)$$
$$\alpha_{2j[i]} = \beta_2 + \eta_{2j}, \quad \eta_{2j} \sim N(0, \tau_2^2)$$
$$\alpha_{3j[i]} = \beta_3 + \eta_{3j}, \quad \eta_{3j} \sim N(0, \tau_3^2)$$

Thus

$$
\begin{aligned}
ratings_i =& \alpha_0 + \beta_1 Instrument_i + \beta_2 Harmony_i + \beta_3 Voice_i + \eta_1 Instrument_i + \eta_2 Harmony_i + \eta_3 Voice_i + \epsilon_i \\
&, \epsilon_i \sim N(0, \sigma^2) \\
&\eta_{1j} \sim N(0, \tau_1^2) \\
&\eta_{2j} \sim N(0, \tau_2^2) \\
&\eta_{3j} \sim N(0, \tau_3^2)
\end{aligned}
$$

Since the three experiment variables (`Instrument, Harmony, Voice`) are factors, we can rewrite equation above as follow:

Let $Instrumentpiano = I_1$, $Instrumentstring = I_2$, $HarmonyI - V - IV = H_1$, $HarmonyI - V - VI = H_2$, $HarmonyIV - I - V = H_3$, $Voicepar3rd = V_1$, $Voicepar5th = V_2$, then

$$
\begin{aligned}
ratings_i &= \alpha_0 + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i + \alpha_{3j[i]} Voice_i + \epsilon_i \\
&= \alpha_0 + \alpha_{1j[i]} I_{1i} + \alpha_{2j[i]} I_{2i} + \alpha_{3j[i]} H_{1i} + \alpha_{4j[i]} H_{2i} + \alpha_{5j[i]} H_{3i} + \alpha_{6j[i]} V_{1i} + \alpha_{7j[i]} V_{2i} + \epsilon_i
\end{aligned}
$$

$$
\begin{aligned}
\alpha_{1j[i]} &= \beta_1 + \eta_{1j}, \quad \eta_{1j} \sim N(0, \tau_1^2) \\
\alpha_{2j[i]} &= \beta_2 + \eta_{2j}, \quad \eta_{2j} \sim N(0, \tau_2^2) \\
\alpha_{3j[i]} &= \beta_3 + \eta_{3j}, \quad \eta_{3j} \sim N(0, \tau_3^2) \\
\alpha_{4j[i]} &= \beta_4 + \eta_{4j}, \quad \eta_{4j} \sim N(0, \tau_4^2) \\
\alpha_{5j[i]} &= \beta_5 + \eta_{5j}, \quad \eta_{5j} \sim N(0, \tau_5^2) \\
\alpha_{6j[i]} &= \beta_6 + \eta_{6j}, \quad \eta_{6j} \sim N(0, \tau_6^2) \\
\alpha_{7j[i]} &= \beta_7 + \eta_{7j}, \quad \eta_{7j} \sim N(0, \tau_7^2)
\end{aligned}
$$

so that we can match $\hat{\eta_1} - \hat{\eta_7}$ to the random effects shown in `lmer2b` results above. The R code will be as follows:

```
# R code:
lmer2b <- lmer(Classical ~ Instrument + Harmony + Voice + (0 + Instrument + Harmony +
    Voice | Subject))
```

Question 2

Individual covariates. For this problem, begin with your best model from problem 1.

2-(a)

Determine which individual covariates should be added to the model as fixed effects. Show a suitable summary of your work, and list the final set of variables that you would include in the model. Hint: Some covariates that are actually factor variables are coded as numeric. Be careful to treat them as factors!

```
# Model selection!
names(ratings)
"OMSI"  Score on a test of musical knowledge
```

```
"X16.minus.17" Auxiliary measure of listeners ability to distinguish classical vs popular
     music
"X1stInstr" ow proficient are you at your first musical instrument (0-5, 0=not at all)
"Selfdeclare"    Are you a musician? (1-6, 1=not at all)

tmp<-lm(Classical~Instrument + Harmony + Voice + OMSI + X16.minus.17 + X1stInstr +
    Selfdeclare)
summary(tmp)

# OMSI, X16.minus.17, and Selfdeclare variables look like having association with
    experiment variables and response variable (Classical).

lmer.covariate.1a <- update(lmer2b, . ~ . + OMSI)
lmer.covariate.1b <- update(lmer2b, . ~ . + X16.minus.17)
lmer.covariate.1c <- update(lmer2b, . ~ . + Selfdeclare)
lmer.covariate.1d <- update(lmer2b, . ~ . + OMSI + X16.minus.17 + Selfdeclare)

anova(lmer2b,lmer.covariate.1a, lmer.covariate.1b, lmer.covariate.1c, lmer.covariate.1d)
Df    AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2b            45 9971.1 10233 -4940.6   9881.1
lmer.covariate.1a 46 9972.9 10241 -4940.4   9880.9 0.2336      1     0.6288
lmer.covariate.1b 46 9969.8 10238 -4938.9   9877.8 3.0585      0     <2e-16 ***
lmer.covariate.1c 46 9973.1 10241 -4940.5   9881.1 0.0000      0     1.0000
lmer.covariate.1d 48 9975.5 10255 -4939.8   9879.5 1.5587      2     0.4587

a<-cbind(
  AIC=sapply(list(lmer2b=lmer2b, lmer.covariate.1a=lmer.covariate.1a, lmer.covariate.1b=
    lmer.covariate.1b, lmer.covariate.1c=lmer.covariate.1c, lmer.covariate.1d=lmer.
    covariate.1d),AIC)
  ,
  DIC=sapply(list(lmer2b=lmer2b, lmer.covariate.1a=lmer.covariate.1a, lmer.covariate.1b=
    lmer.covariate.1b, lmer.covariate.1c=lmer.covariate.1c, lmer.covariate.1d=lmer.
    covariate.1d), invisible(function(x) display(x)$DIC))
  ,
  BIC=sapply(list(lmer2b=lmer2b, lmer.covariate.1a=lmer.covariate.1a, lmer.covariate.1b=
    lmer.covariate.1b, lmer.covariate.1c=lmer.covariate.1c, lmer.covariate.1d=lmer.
    covariate.1d),BIC)
)
                       AIC      DIC      BIC
lmer2b            10034.59 9817.616 10296.54
lmer.covariate.1a 10036.91 9816.823 10304.69
lmer.covariate.1b 10017.19 9830.425 10284.97
lmer.covariate.1c 10026.39 9827.768 10294.16
lmer.covariate.1d 10045.24 9809.799 10324.66

# AIC and BIC shows weak preference of new models (only for lmer.covariate.1b) with
    additional covariates.
# Try with other covariates

tmp2<-lm(Classical~Instrument + Harmony + Voice + factor(CollegeMusic) + factor(APTheory)
    )
```

```
summary(tmp2)
tmp3<-lm(Classical~Instrument + Harmony + Voice + factor(APTheory) + X1stInstr + NoClass)
summary(tmp3)

lmer.covariate.2a <- update(lmer2b, . ~ . + factor(APTheory))
lmer.covariate.2b <- update(lmer2b, . ~ . + X1stInstr)
lmer.covariate.2c <- update(lmer2b, . ~ . + NoClass)
lmer.covariate.2d <- update(lmer2b, . ~ . + factor(APTheory) + X1stInstr + NoClass)

anova(lmer2b,lmer.covariate.1b, lmer.covariate.2a, lmer.covariate.2b, lmer.covariate.2c,
    lmer.covariate.2d)
Df    AIC     BIC  logLik deviance     Chisq Chi Df Pr(>Chisq)
lmer2b            45 9971.1 10233.1 -4940.6   9881.1
lmer.covariate.1b 46 9969.8 10237.6 -4938.9   9877.8    3.2921      1    0.06961 .
lmer.covariate.2a 46 9172.7  9436.6 -4540.4   9080.7  797.0726      0    < 2e-16 ***
lmer.covariate.2b 46 3996.3  4222.1 -1952.1   3904.3 5176.4428      0    < 2e-16 ***
lmer.covariate.2c 46 8863.0  9125.4 -4385.5   8771.0    0.0000      0    1.00000
lmer.covariate.2d 48 3569.8  3799.9 -1736.9   3473.8 5297.2223      2    < 2e-16 ***
---

a<-cbind(
  AIC=sapply(list(lmer2b=lmer2b,lmer.covariate.1b=lmer.covariate.1b, lmer.covariate.2a=
    lmer.covariate.2a,lmer.covariate.2b=lmer.covariate.2b,lmer.covariate.2c=lmer.
    covariate.2c,lmer.covariate.2d=lmer.covariate.2d),AIC)
  ,
  DIC=sapply(list(lmer2b=lmer2b,lmer.covariate.1b=lmer.covariate.1b, lmer.covariate.2a=
    lmer.covariate.2a,lmer.covariate.2b=lmer.covariate.2b,lmer.covariate.2c=lmer.
    covariate.2c,lmer.covariate.2d=lmer.covariate.2d), invisible(function(x) display(x)
    $DIC))
  ,
  BIC=sapply(list(lmer2b=lmer2b,lmer.covariate.1b=lmer.covariate.1b, lmer.covariate.2a=
    lmer.covariate.2a,lmer.covariate.2b=lmer.covariate.2b,lmer.covariate.2c=lmer.
    covariate.2c,lmer.covariate.2d=lmer.covariate.2d),BIC)
)
                        AIC      DIC       BIC
lmer2b            10034.585 9817.616 10296.541
lmer.covariate.1b 10017.193 9830.425 10284.970
lmer.covariate.2a  9295.816 8957.656  9559.666
lmer.covariate.2b  4026.692 3873.895  4252.449
lmer.covariate.2c  8939.388 8694.586  9201.768
lmer.covariate.2d  3610.515 3433.014  3840.601

# AIC prefer lmer.covariate.2d a lot.

formula(lmer.covariate.2d)
Classical ~ Instrument + Harmony + Voice + (0 + Instrument + Harmony + Voice | Subject) +
     factor(APTheory) + X1stInstr + NoClass

lmer.ranef.0<-lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass + (0 + Instrument + Harmony + Voice | Subject))
```

```
lmer.ranef.1 <- lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass + (0 + Instrument + Harmony + Voice + factor(APTheory) | Subject)
    )
lmer.ranef.2 <- lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass + (0 + Instrument + Harmony + Voice + X1stInstr | Subject))
lmer.ranef.3 <- lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass + (0 + Instrument + Harmony + Voice + NoClass | Subject))
lmer.ranef.4 <- lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass + (0 + Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass | Subject))

anova(lmer.ranef.0, lmer.ranef.1, lmer.ranef.2, lmer.ranef.3, lmer.ranef.4)
             Df    AIC    BIC  logLik deviance   Chisq Chi Df Pr(>Chisq)
lmer.ranef.0 48 3569.8 3799.9 -1736.9   3473.8
lmer.ranef.1 57 3581.7 3854.9 -1733.9   3467.7  6.0502      9     0.7349
lmer.ranef.2 57 3577.7 3850.9 -1731.8   3463.7  4.0121      0     <2e-16 ***
lmer.ranef.3 57 3580.1 3853.3 -1733.0   3466.1  0.0000      0     1.0000
lmer.ranef.4 78 3608.6 3982.5 -1726.3   3452.6 13.4691     21     0.8913

a<-cbind(
  AIC=sapply(list(lmer.ranef.0=lmer.ranef.0, lmer.ranef.1=lmer.ranef.1, lmer.ranef.2=lmer
    .ranef.2, lmer.ranef.3=lmer.ranef.3, lmer.ranef.4=lmer.ranef.4),AIC)
  ,
  DIC=sapply(list(lmer.ranef.0=lmer.ranef.0, lmer.ranef.1=lmer.ranef.1, lmer.ranef.2=lmer
    .ranef.2, lmer.ranef.3=lmer.ranef.3, lmer.ranef.4=lmer.ranef.4), invisible(function(x
    ) display(x)$DIC))
  ,
  BIC=sapply(list(lmer.ranef.0=lmer.ranef.0, lmer.ranef.1=lmer.ranef.1, lmer.ranef.2=lmer
    .ranef.2, lmer.ranef.3=lmer.ranef.3, lmer.ranef.4=lmer.ranef.4),BIC)
)

                  AIC      DIC      BIC
lmer.ranef.0 3610.515 3433.014 3840.601
lmer.ranef.1 3635.228 3414.201 3908.455
lmer.ranef.2 3634.225 3407.180 3907.452
lmer.ranef.3 3623.942 3422.180 3897.170
lmer.ranef.4 3674.838 3386.346 4048.728

# AIC prefers lmer.ranef.0

formula(lmer.ranef.0)
Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass + (0 + Instrument + Harmony + Voice |
    Subject)
```

2-(b)

Once the fixed effects are settled, go back and check to see whether there should be any change in the random effects. Provide suitable evidence to justify your answer.

```
# As shown in Question 2-(a), lmer.ranef.1~3 were compared to lmer.ranef.0 model. And AIC
     prefer lmer.ranef.0, which has no random effects for additional covariates.
```

2-(c)

Briefly interpret the effect of each variable kept in the final model, on Classical ratings.

```
summary(lmer.ranef.0)
Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ Instrument + Harmony + Voice + factor(APTheory) +      X1stInstr +
    NoClass + (0 + Instrument + Harmony + Voice |      Subject)

REML criterion at convergence: 3514.515

Random effects:
 Groups   Name             Variance Std.Dev. Corr
 Subject  Instrumentguitar 1.1821   1.0872
          Instrumentpiano  2.2885   1.5128    0.73
          Instrumentstring 4.8691   2.2066    0.33  0.67
          HarmonyI-V-IV    0.1921   0.4383    0.25 -0.30  0.11
          HarmonyI-V-VI    1.9360   1.3914   -0.18 -0.29 -0.35 -0.01
          HarmonyIV-I-V    0.5412   0.7357    0.18 -0.07  0.32  0.68 -0.01
          Voicepar3rd      0.3667   0.6055   -0.10  0.07  0.08 -0.18 -0.28  0.32
          Voicepar5th      0.8405   0.9168    0.23  0.48  0.28 -0.34  0.05  0.30  0.78
 Residual                  2.2482   1.4994
Number of obs: 892, groups: Subject, 25

Fixed effects:
                  Estimate Std. Error t value
(Intercept)        3.72638    0.56667   6.576
Instrumentpiano    1.89656    0.24168   7.847
Instrumentstring   3.59833    0.44004   8.177
HarmonyI-V-IV      0.12874    0.16691   0.771
HarmonyI-V-VI      1.01150    0.31246   3.237
HarmonyIV-I-V      0.03325    0.20452   0.163
Voicepar3rd       -0.50448    0.17270  -2.921
Voicepar5th       -0.40667    0.22081  -1.842
factor(APTheory)1  1.46045    0.51471   2.837
X1stInstr          0.02702    0.13612   0.199
NoClass           -0.09016    0.35412  -0.255

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r Vcpr5t f(APT) X1stIn
Instrumntpn -0.051
Instrmntstr -0.093  0.612
HrmnyI-V-IV -0.057 -0.314    -0.005
HrmnyI-V-VI -0.118 -0.185    -0.238     0.186
HrmnyIV-I-V -0.036 -0.180     0.166     0.551  0.151
Voicepar3rd -0.104  0.121     0.088    -0.066 -0.175  0.165
Voicepar5th  0.013  0.328     0.137    -0.151  0.039  0.177  0.656
fctr(APTh)1 -0.242  0.003     0.000     0.000  0.000  0.000  0.000 -0.002
X1stInstr   -0.736  0.001     0.000     0.000  0.000  0.000  0.000  0.000 -0.110
NoClass     -0.572 -0.001     0.000     0.000  0.000  0.000  0.000  0.001  0.249  0.169
```

Although the effect size has slightly changed, the original experimental variables has simlar effects on Classical ratings with the same direction and significance. In addition to these variables, taking AP music class has positive effect on Classical ratings, and its effect size is relatively large. But as shown above, the random effect of AP music class was insignificant and excluded from the model.

Question 3

Musicians vs. Non-musicians. 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. Dichotomize "Selfdeclare" ("are you a musician?") so that about half the participants are categorized as self-declared musicians, and half not. Examine and report on any interactions between the dichotomized musician variable and other predictors in the model. Provide suitable evidence for, and comment on, your results. [2]

```
describe(Selfdeclare)
Selfdeclare
      n missing  unique    Mean
   2520       0       6   2.443

            1   2   3   4  5  6
Frequency 576 936 468 432 72 36
%          23  37  19  17  3  1

musician <- ifelse(Selfdeclare > 2, 1, 0)
table (musician, Selfdeclare)
Selfdeclare
musician    1   2   3   4   5   6
       0 576 936   0   0   0   0
       1   0   0 468 432  72  36


lmer.ranef.0<-lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass + (0 + Instrument + Harmony + Voice | Subject))

lmer.interaction.1<-lmer(Classical ~ Instrument*musician + Harmony + Voice + factor(
    APTheory) + X1stInstr + NoClass + (0 + Instrument + Harmony + Voice | Subject))
lmer.interaction.2<-lmer(Classical ~ Instrument + Harmony*musician + Voice + factor(
    APTheory) + X1stInstr + NoClass + (0 + Instrument + Harmony + Voice | Subject))
lmer.interaction.3<-lmer(Classical ~ Instrument + Harmony + Voice*musician + factor(
    APTheory) + X1stInstr + NoClass + (0 + Instrument + Harmony + Voice | Subject))
lmer.interaction.4<-lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory)*
    musician + X1stInstr + NoClass + (0 + Instrument + Harmony + Voice | Subject))
lmer.interaction.5<-lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr*musician + NoClass + (0 + Instrument + Harmony + Voice | Subject))
lmer.interaction.6<-lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass*musician + (0 + Instrument + Harmony + Voice | Subject))

anova(lmer.interaction.1, lmer.interaction.2, lmer.interaction.3, lmer.interaction.4,
    lmer.interaction.5, lmer.interaction.6)
Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer.interaction.4 50 3572.2 3811.8 -1736.1   3472.2
lmer.interaction.5 50 3572.0 3811.7 -1736.0   3472.0 0.1418      0    < 2e-16 ***
lmer.interaction.6 50 3571.1 3810.7 -1735.5   3471.1 0.9475      0    < 2e-16 ***
```

[2] http://lme4.r-forge.r-project.org/slides/2009-07-21-Seewiesen/6Interactions-4a4.pdf

```
lmer.interaction.1 51 3574.8 3819.3 -1736.4   3472.8 0.0000      1    1.00000
lmer.interaction.3 51 3575.2 3819.7 -1736.6   3473.2 0.0000      0    1.00000
lmer.interaction.2 52 3574.1 3823.4 -1735.0   3470.1 3.1119      1    0.07772 .
---

# AIC weakly prefer lmer.interaction.6.

lmer.interaction.6a<-lmer(Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
    X1stInstr + NoClass*musician + (0 + Instrument + Harmony + Voice | Subject) + (0 +
    Instrument + Harmony + Voice | Subject:musician))

anova(lmer.interaction.6, lmer.interaction.6a)
                    Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer.interaction.6  50 3571.1 3810.7 -1735.5   3471.1
lmer.interaction.6a 86 3643.1 4055.3 -1735.5   3471.1     0     36          1

a<-cbind(
  AIC=sapply(list(lmer.interaction.1=lmer.interaction.1, lmer.interaction.2=lmer.
    interaction.2, lmer.interaction.3=lmer.interaction.3, lmer.interaction.4=lmer.
    interaction.4, lmer.interaction.5=lmer.interaction.5, lmer.interaction.6=lmer.
    interaction.6, lmer.interaction.6a=lmer.interaction.6a),AIC)
  ,
  DIC=sapply(list(lmer.interaction.1=lmer.interaction.1, lmer.interaction.2=lmer.
    interaction.2, lmer.interaction.3=lmer.interaction.3, lmer.interaction.4=lmer.
    interaction.4, lmer.interaction.5=lmer.interaction.5, lmer.interaction.6=lmer.
    interaction.6, lmer.interaction.6a=lmer.interaction.6a), invisible(function(x)
    display(x)$DIC))
  ,
  BIC=sapply(list(lmer.interaction.1=lmer.interaction.1, lmer.interaction.2=lmer.
    interaction.2, lmer.interaction.3=lmer.interaction.3, lmer.interaction.4=lmer.
    interaction.4, lmer.interaction.5=lmer.interaction.5, lmer.interaction.6=lmer.
    interaction.6, lmer.interaction.6a=lmer.interaction.6a),BIC)
)

                         AIC      DIC      BIC
lmer.interaction.1  3625.993 3421.582 3870.459
lmer.interaction.2  3611.161 3433.054 3860.422
lmer.interaction.3  3612.865 3435.574 3857.332
lmer.interaction.4  3605.961 3438.354 3845.634
lmer.interaction.5  3611.089 3432.943 3850.762
lmer.interaction.6  3616.229 3425.908 3855.902
lmer.interaction.6a 3729.928 3384.209 4142.166

# AIC and BIC weakly prefer model 4, while DIC prefer model 6a a lot.

summary(lmer.interaction.4)
Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ Instrument + Harmony + Voice + factor(APTheory)*musician + X1stInstr
     + NoClass + (0 + Instrument + Harmony +      Voice | Subject)

REML criterion at convergence: 3505.961
```

```
Random effects:
 Groups   Name              Variance Std.Dev. Corr
 Subject  Instrumentguitar 1.3727   1.1716
          Instrumentpiano  2.0191   1.4209    0.66
          Instrumentstring 2.0125   1.4186    0.02  0.47
          HarmonyI-V-IV    0.4692   0.6850    0.20  0.03  0.43
          HarmonyI-V-VI    2.5946   1.6108    0.32  0.15 -0.18  0.30
          HarmonyIV-I-V    0.6024   0.7762    0.20  0.12  0.38  0.88  0.39
          Voicepar3rd      0.2531   0.5031   -0.18 -0.04  0.09 -0.37 -0.42 -0.07
          Voicepar5th      0.4784   0.6917    0.21  0.53  0.10 -0.42 -0.02 -0.04  0.74
 Residual                  2.2572   1.5024
Number of obs: 892, groups: Subject, 25

Fixed effects:
                           Estimate Std. Error t value
(Intercept)                 4.61335    0.60117   7.674
Instrumentpiano             1.89804    0.25253   7.516
Instrumentstring            3.59833    0.38370   9.378
HarmonyI-V-IV               0.13236    0.19756   0.670
HarmonyI-V-VI               1.01493    0.35222   2.881
HarmonyIV-I-V               0.03563    0.21062   0.169
Voicepar3rd                -0.50185    0.15918  -3.153
Voicepar5th                -0.40630    0.18529  -2.193
factor(APTheory)1           0.37888    0.70336   0.539
musician                   -0.82547    0.50960  -1.620
X1stInstr                  -0.17422    0.15027  -1.159
NoClass                     0.04000    0.35881   0.111
factor(APTheory)1:musician  1.90609    1.00980   1.888

summary(lmer.interaction.6a)
Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ Instrument + Harmony + Voice + factor(APTheory) +      X1stInstr +
    NoClass * musician + (0 + Instrument + Harmony +      Voice | Subject) + (0 +
    Instrument + Harmony + Voice | Subject:musician)

REML criterion at convergence: 3557.928

Random effects:
 Groups           Name             Variance Std.Dev. Corr
 Subject          Instrumentguitar 2.3169   1.5221
                  Instrumentpiano  2.2895   1.5131    0.68
                  Instrumentstring 0.7517   0.8670    0.38  0.68
                  HarmonyI-V-IV    0.4379   0.6617   -0.43 -0.70 -0.14
                  HarmonyI-V-VI    2.1886   1.4794    0.31 -0.10  0.39  0.55
                  HarmonyIV-I-V    0.5296   0.7277   -0.45 -0.80 -0.22  0.92  0.49
                  Voicepar3rd      2.3565   1.5351    0.28  0.07  0.27  0.27  0.42  0.27
                  Voicepar5th      3.0322   1.7413    0.41  0.26  0.46  0.15  0.46  0.15
    0.97
 Subject:musician Instrumentguitar 1.9343   1.3908
                  Instrumentpiano  0.7366   0.8582   0.99
```

```
                 Instrumentstring 2.3042   1.5179   0.57  0.52
                 HarmonyI-V-IV    0.4915   0.7011   0.90  0.92  0.66
                 HarmonyI-V-VI    2.1330   1.4605   0.45  0.48  0.19  0.56
                 HarmonyIV-I-V    0.2599   0.5098   0.49  0.52  0.30  0.65  0.57
                 Voicepar3rd      0.4251   0.6520   0.24  0.20  0.51  0.31  0.39  0.65
                 Voicepar5th      0.7796   0.8829   0.42  0.39  0.47  0.44  0.60  0.74
    0.94
 Residual                         2.1570   1.4687
Number of obs: 892, groups: Subject, 25; Subject:musician, 25

Fixed effects:
                  Estimate Std. Error t value
(Intercept)        3.96263    0.79561   4.981
Instrumentpiano    1.89657    0.29114   6.514
Instrumentstring   3.59833    0.41295   8.714
HarmonyI-V-IV      0.13209    0.23780   0.555
HarmonyI-V-VI      1.01363    0.43846   2.312
HarmonyIV-I-V      0.03524    0.22572   0.156
Voicepar3rd       -0.50565    0.35472  -1.426
Voicepar5th       -0.40842    0.40866  -0.999
factor(APTheory)1  1.25411    0.64168   1.954
X1stInstr         -0.19494    0.18227  -1.070
NoClass            1.06724    0.58419   1.827
musician           1.00987    0.78963   1.279
NoClass:musician  -2.61946    0.83771  -3.127
```

As shown above, in model 4, interaction `APTheory:mucisian` is insignificant. But in model 6a, interaction `NoClass:mucisian` is significant. In model 6a, we can interpret the fixed effects regarding to interaction term as that: when the number of music class increase by 1 unit, Classical ratings increases by 1.06 among respondents who identified themselves as non-musician, whereas Classical ratings decreases by (1.06-2.62) points (with 1 point higher on average) among respondents identified themselves as musicians. Note that only interaction term is statistically significant. Also among the experimental variables (`Instrument, Harmony, Voice`), `Voice` variable changed to be insignificant comparing to older model without covariates. And one of the covariates, `APTheory` remains to be significant.

Question 4

Classical vs. Popular. Please re-examine the data in terms of the "Popular" ratings, instead of the "Classical" ratings, using similar hierarchical linear models. Provide brief answers to the following questions:

4-(a)

Comment on the influence of Instrument, Harmony & Voice on Popular ratings, providing suitable brief evidence for each result.

```
# To compare the influences of each experimental variable and covariate on the different
    dependent variable, Popular, I will use the same model as before
# For Q4-(a), I use the same model as Q1-(c)-iii except dependent variable.

lmer.Q4.a<- lmer(Popular ~ Instrument + Harmony + Voice + (0 + Instrument + Harmony +
    Voice | Subject))
summary(lmer.Q4.a)
```

these are not the randomeffects I asked you to explore. you have not made an argument that they are better than the random effects that I asked youto explore.

```
Linear mixed model fit by REML ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice + (0 + Instrument + Harmony + Voice |
    Subject)

REML criterion at convergence: 9981.723

Random effects:
 Groups   Name             Variance Std.Dev. Corr
 Subject  Instrumentguitar 1.48551  1.2188
          Instrumentpiano  3.71764  1.9281    0.50
          Instrumentstring 4.16469  2.0408    0.18  0.78
          HarmonyI-V-IV    0.19702  0.4439    0.58  0.47  0.33
          HarmonyI-V-VI    0.82043  0.9058    0.12 -0.01 -0.09 -0.30
          HarmonyIV-I-V    0.41518  0.6443   -0.29 -0.31 -0.19 -0.07  0.04
          Voicepar3rd      0.09791  0.3129   -0.14  0.24  0.39  0.23 -0.47  0.39
          Voicepar5th      0.13786  0.3713   -0.05  0.00  0.43  0.05 -0.16  0.54  0.42
 Residual                  2.44136  1.5625
Number of obs: 2493, groups: Subject, 70

Fixed effects:
                 Estimate Std. Error t value
(Intercept)       6.57855    0.17041   38.60
Instrumentpiano  -0.94848    0.21605   -4.39
Instrumentstring -2.60561    0.27166   -9.59
HarmonyI-V-IV    -0.02646    0.10323   -0.26
HarmonyI-V-VI    -0.27214    0.13988   -1.95
HarmonyIV-I-V    -0.18555    0.11730   -1.58
Voicepar3rd       0.16630    0.08535    1.95
Voicepar5th       0.16406    0.08859    1.85
```

As we can see above, the effect directions of each experimental variable are the opposite direction comparing to original model using Classical variable as dependent variable. Effect size changed slightly

4-(b)

Question 2c, for Popular ratings.

```
lmer.Q4.b<-lmer(Popular ~ Instrument + Harmony + Voice + factor(APTheory) + X1stInstr +
    NoClass + (0 + Instrument + Harmony + Voice | Subject))
summary(lmer.Q4.b)
Linear mixed model fit by REML ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice + factor(APTheory) + X1stInstr +
    NoClass + (0 + Instrument + Harmony + Voice | Subject)

REML criterion at convergence: 3545.713

Random effects:
 Groups   Name             Variance Std.Dev. Corr
 Subject  Instrumentguitar 1.4629   1.2095
          Instrumentpiano  1.6252   1.2748    0.57
          Instrumentstring 3.8290   1.9568    0.22  0.73
          HarmonyI-V-IV    0.1139   0.3375    0.74  0.34  0.43
          HarmonyI-V-VI    1.1482   1.0715    0.00  0.13  0.14 -0.14
```

```
         HarmonyIV-I-V    0.9283   0.9635    0.07 -0.25 -0.19  0.38 -0.25
         Voicepar3rd      0.1507   0.3882    0.03  0.42  0.37  0.11  0.22  0.28
         Voicepar5th      0.6683   0.8175    0.20  0.46  0.47  0.35  0.29  0.52  0.81
 Residual                 2.3529   1.5339
Number of obs: 892, groups: Subject, 25

Fixed effects:
                  Estimate Std. Error t value
(Intercept)        5.69438    0.61379   9.277
Instrumentpiano   -1.26541    0.26289  -4.813
Instrumentstring  -2.85500    0.43182  -6.612
HarmonyI-V-IV     -0.04139    0.16019  -0.258
HarmonyI-V-VI     -0.10722    0.25895  -0.414
HarmonyIV-I-V     -0.11753    0.24134  -0.487
Voicepar3rd        0.24794    0.14793   1.676
Voicepar5th        0.22323    0.20632   1.082
factor(APTheory)1  0.46931    0.55836   0.841
X1stInstr          0.11468    0.14733   0.778
NoClass            0.30547    0.38335   0.797
```

When we used the same model as before,we can see that those three covariates are insignificant. Thus we need to identify meaningful covariates for new model with Popular variable.

```
tmp<-lm(Popular~Instrument + Harmony + Voice + OMSI + X16.minus.17 + X1stInstr +
    Selfdeclare + ConsInstr + ConsNotes + Instr.minus.Notes + PachListen + ClsListen +
    KnowRob + KnowAxis + X1990s2000s + CollegeMusic + NoClass + APTheory + Composing +
    PianoPlay + GuitarPlay + X1stInstr)
summary(tmp)


lmer.covariate.1a <- update(lmer.Q4.a, . ~ . + OMSI)
lmer.covariate.1b <- update(lmer.Q4.a, . ~ . + Selfdeclare)
lmer.covariate.1c <- update(lmer.Q4.a, . ~ . + ClsListen)
lmer.covariate.1d <- update(lmer.Q4.a, . ~ . + KnowAxis)
lmer.covariate.1e <- update(lmer.Q4.a, . ~ . + X1990s2000s)
lmer.covariate.1f <- update(lmer.Q4.a, . ~ . + factor(APTheory))
lmer.covariate.1g <- update(lmer.Q4.a, . ~ . + PianoPlay)
lmer.covariate.1h <- update(lmer.Q4.a, . ~ . + GuitarPlay)
lmer.covariate.1i <- update(lmer.Q4.a, . ~ . + OMSI + Selfdeclare + ClsListen + KnowAxis
    + X1990s2000s + factor(APTheory) + PianoPlay + GuitarPlay)

a<-cbind(
  AIC=sapply(list(lmer.covariate.1a=lmer.covariate.1a, lmer.covariate.1b=lmer.covariate.1
    b, lmer.covariate.1c=lmer.covariate.1c, lmer.covariate.1d=lmer.covariate.1d, lmer.
    covariate.1e=lmer.covariate.1e, lmer.covariate.1f=lmer.covariate.1f, lmer.covariate.1
    g=lmer.covariate.1g, lmer.covariate.1h=lmer.covariate.1h, lmer.covariate.1i=lmer.
    covariate.1i),AIC)
  ,
  DIC=sapply(list(lmer.covariate.1a=lmer.covariate.1a, lmer.covariate.1b=lmer.covariate.1
    b, lmer.covariate.1c=lmer.covariate.1c, lmer.covariate.1d=lmer.covariate.1d, lmer.
    covariate.1e=lmer.covariate.1e, lmer.covariate.1f=lmer.covariate.1f, lmer.covariate.1
    g=lmer.covariate.1g, lmer.covariate.1h=lmer.covariate.1h, lmer.covariate.1i=lmer.
    covariate.1i), invisible(function(x) display(x)$DIC))
```

```
  ,
  BIC=sapply(list(lmer.covariate.1a=lmer.covariate.1a, lmer.covariate.1b=lmer.covariate.1
    b, lmer.covariate.1c=lmer.covariate.1c, lmer.covariate.1d=lmer.covariate.1d, lmer.
    covariate.1e=lmer.covariate.1e, lmer.covariate.1f=lmer.covariate.1f, lmer.covariate.1
    g=lmer.covariate.1g, lmer.covariate.1h=lmer.covariate.1h, lmer.covariate.1i=lmer.
    covariate.1i),BIC)
)

                         AIC      DIC       BIC
lmer.covariate.1a 10108.416 9844.875 10376.193
lmer.covariate.1b 10092.366 9855.979 10360.144
lmer.covariate.1c  9985.469 9768.683 10252.801
lmer.covariate.1d  8977.449 8737.118  9239.580
lmer.covariate.1e  9587.022 9340.917  9852.297
lmer.covariate.1f  9274.855 9036.675  9538.705
lmer.covariate.1g 10123.639 9829.498 10391.416
lmer.covariate.1h 10133.777 9817.805 10401.554
lmer.covariate.1i  8064.670 7751.701  8360.582
```

does not look like you are making sure the data set is the same across all models.

```
# AIC, BIC, and DIC prefer model lmer.covariate.1i.

summary(lmer.covariate.1i)
Fixed effects:
                    Estimate Std. Error t value
(Intercept)        6.0799445  0.8053826   7.549
Instrumentpiano   -1.1421080  0.2264322  -5.044
Instrumentstring  -2.9975473  0.3231840  -9.275
HarmonyI-V-IV      0.0093357  0.1097650   0.085
HarmonyI-V-VI     -0.3306633  0.1980077  -1.670
HarmonyIV-I-V     -0.2277782  0.1364675  -1.669
Voicepar3rd        0.2044649  0.1457046   1.403
Voicepar5th        0.2528644  0.1296060   1.951
OMSI              -0.0004755  0.0013471  -0.353
Selfdeclare        0.2897580  0.2588317   1.119
ClsListen          0.0438824  0.1302149   0.337
KnowAxis           0.0593246  0.0823374   0.721
X1990s2000s       -0.0280559  0.1188218  -0.236
factor(APTheory)1 -0.0638202  0.4562731  -0.140
PianoPlay          0.0065089  0.1170186   0.056
GuitarPlay        -0.0527905  0.1604702  -0.329
```

you are not getting any significant fixed effects.

there are indeed significant fixed effects here

Although model comparison results indicate that model 1i is better than any other models, all the covariates included in model 1i turned out to be insignificant. In addition, other combinations of additional covariates show that covariates are insignificant. Thus, in this case, we can think that there are no meaningful, significant covariates that we should include in the model.

4-(c)

Question 3, for Popular ratings.

```
# Since in question 4-(b), we could not find any covariates that need to be included in
    the model, we can approach Q4-(c) as we have done in earlier. This model includes
    three covariates and one dummy variable. But in this case, we can drop three
```

```
    covariates, and test the interaction effect of "musician" variable and three
    experimental variable.

lmer.Q4.c<-lmer(Popular ~ Instrument + Harmony + Voice + factor(APTheory) + X1stInstr +
    NoClass * musician + (0 + Instrument + Harmony + Voice | Subject) + (0 + Instrument +
     Harmony + Voice | Subject:musician) )
summary(lmer.Q4.c)
Linear mixed model fit by REML ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice + factor(APTheory) + X1stInstr +
    NoClass * musician + (0 + Instrument + Harmony + Voice |      Subject) + (0 +
    Instrument + Harmony + Voice | Subject:musician)

REML criterion at convergence: 3576.707

Random effects:
 Groups           Name             Variance Std.Dev. Corr
 Subject          Instrumentguitar 2.43149  1.5593
                  Instrumentpiano  2.18670  1.4788    0.91
                  Instrumentstring 2.73144  1.6527    0.56  0.79
                  HarmonyI-V-IV    0.06044  0.2459   -0.26 -0.43 -0.34
                  HarmonyI-V-VI    0.67041  0.8188   -0.24 -0.35 -0.44  0.16
                  HarmonyIV-I-V    0.39162  0.6258    0.16 -0.03 -0.04  0.12  0.51
                  Voicepar3rd      2.31903  1.5228    0.21  0.19  0.21 -0.06  0.13  0.35
                  Voicepar5th      1.83521  1.3547    0.37  0.39  0.37 -0.21  0.19  0.42
     0.95
 Subject:musician Instrumentguitar 0.42703  0.6535
                  Instrumentpiano  1.26076  1.1228    0.28
                  Instrumentstring 2.21273  1.4875    0.38  0.81
                  HarmonyI-V-IV    0.03913  0.1978    0.39 -0.61 -0.16
                  HarmonyI-V-VI    1.06737  1.0331   -0.14  0.11  0.38 -0.07
                  HarmonyIV-I-V    0.66098  0.8130    0.01 -0.21 -0.21  0.26  0.00
                  Voicepar3rd      0.45380  0.6736   -0.07  0.17  0.19 -0.16  0.41  0.30
                  Voicepar5th      0.45316  0.6732    0.15  0.35  0.33 -0.19  0.37  0.47
     0.90
 Residual                          2.25392  1.5013
Number of obs: 892, groups: Subject, 25; Subject:musician, 25

Fixed effects:
                  Estimate Std. Error t value
(Intercept)        5.79198    0.84584   6.848
Instrumentpiano   -1.25966    0.29027  -4.340
Instrumentstring  -2.85500    0.42607  -6.701
HarmonyI-V-IV     -0.04393    0.15559  -0.282
HarmonyI-V-VI     -0.11205    0.29961  -0.374
HarmonyIV-I-V     -0.11963    0.24969  -0.479
Voicepar3rd        0.23921    0.35516   0.674
Voicepar5th        0.21736    0.32668   0.665
factor(APTheory)1  0.89985    0.73231   1.229
X1stInstr          0.08265    0.20873   0.396
NoClass            0.33894    0.66893   0.507
musician          -0.53976    0.90129  -0.599
```

```
NoClass:musician   0.40944    0.95758   0.428

new.c.1<-lmer(Popular ~ Instrument*musician + Harmony + Voice  + (0 + Instrument +
    Harmony + Voice | Subject) + (0 + Instrument + Harmony + Voice | Subject:musician) )
new.c.2<-lmer(Popular ~ Instrument + Harmony*musician + Voice  + (0 + Instrument +
    Harmony + Voice | Subject) + (0 + Instrument + Harmony + Voice | Subject:musician) )
new.c.3<-lmer(Popular ~ Instrument + Harmony + Voice*musician  + (0 + Instrument +
    Harmony + Voice | Subject) + (0 + Instrument + Harmony + Voice | Subject:musician) )
summary(lmer.Q4.c)


lmer.interaction.1<-lmer(Popular ~ Instrument*musician + Harmony + Voice + (0 +
    Instrument + Harmony + Voice | Subject))
lmer.interaction.2<-lmer(Popular ~ Instrument + Harmony*musician + Voice + (0 +
    Instrument + Harmony + Voice | Subject))
lmer.interaction.3<-lmer(Popular ~ Instrument + Harmony + Voice*musician + (0 +
    Instrument + Harmony + Voice | Subject))
lmer.interaction.4<-lmer(Popular ~ Instrument*musician + Harmony + Voice + (0 +
    Instrument + Harmony + Voice | Subject:musician))
lmer.interaction.5<-lmer(Popular ~ Instrument + Harmony*musician + Voice + (0 +
    Instrument + Harmony + Voice | Subject:musician))
lmer.interaction.6<-lmer(Popular ~ Instrument + Harmony + Voice*musician + (0 +
    Instrument + Harmony + Voice | Subject:musician))

a<-cbind(
  AIC=sapply(list(lmer.interaction.1=lmer.interaction.1, lmer.interaction.2=lmer.
    interaction.2, lmer.interaction.3=lmer.interaction.3, lmer.interaction.4=lmer.
    interaction.4, lmer.interaction.5=lmer.interaction.5, lmer.interaction.6=lmer.
    interaction.6),AIC)
  ,
  DIC=sapply(list(lmer.interaction.1=lmer.interaction.1, lmer.interaction.2=lmer.
    interaction.2, lmer.interaction.3=lmer.interaction.3, lmer.interaction.4=lmer.
    interaction.4, lmer.interaction.5=lmer.interaction.5, lmer.interaction.6=lmer.
    interaction.6), invisible(function(x) display(x)$DIC))
  ,
  BIC=sapply(list(lmer.interaction.1=lmer.interaction.1, lmer.interaction.2=lmer.
    interaction.2, lmer.interaction.3=lmer.interaction.3, lmer.interaction.4=lmer.
    interaction.4, lmer.interaction.5=lmer.interaction.5, lmer.interaction.6=lmer.
    interaction.6),BIC)
)

                        AIC      DIC      BIC
lmer.interaction.1 10084.72 9866.748 10364.14
lmer.interaction.2 10097.22 9842.635 10382.46
lmer.interaction.3 10078.09 9873.392 10357.51
lmer.interaction.4 10084.72 9866.748 10364.14
lmer.interaction.5 10097.22 9842.635 10382.46
lmer.interaction.6 10078.09 9873.392 10357.51
```

these models do not make a lot of sense...

Basically interaction model 4 – 6 are the same as model 1 – 3. Based on AIC, BIC, and DIC above, interaction model 3 is better than others.

```
summary(lmer.interaction.3)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice * musician + (0 + Instrument +
    Harmony + Voice | Subject)

REML criterion at convergence: 9982.091

Random effects:
 Groups   Name             Variance Std.Dev. Corr
 Subject  Instrumentguitar 2.6465   1.6268
          Instrumentpiano  5.3325   2.3092    0.78
          Instrumentstring 6.0628   2.4623    0.62  0.85
          HarmonyI-V-IV    0.1634   0.4043    0.48  0.50  0.34
          HarmonyI-V-VI    1.1497   1.0722   -0.14 -0.14 -0.16 -0.04
          HarmonyIV-I-V    0.3625   0.6021   -0.45 -0.51 -0.43 -0.62 -0.03
          Voicepar3rd      0.2194   0.4684   -0.20  0.02 -0.01  0.30 -0.15  0.15
          Voicepar5th      0.2044   0.4521   -0.20 -0.19 -0.03 -0.11  0.05  0.39  0.72
 Residual                  2.3933   1.5470
Number of obs: 2493, groups: Subject, 70

Fixed effects:
                     Estimate Std. Error t value
(Intercept)           6.50778    0.26181  24.857
Instrumentpiano      -0.94907    0.18928  -5.014
Instrumentstring     -2.60502    0.24338 -10.703
HarmonyI-V-IV        -0.02748    0.10010  -0.275
HarmonyI-V-VI        -0.27359    0.15530  -1.762
HarmonyIV-I-V        -0.18574    0.11336  -1.639
Voicepar3rd           0.22285    0.12076   1.845
Voicepar5th           0.12461    0.11927   1.045
musician              0.17992    0.37961   0.474
Voicepar3rd:musician -0.14583    0.18878  -0.772
Voicepar5th:musician  0.09463    0.18654   0.507
```

what about interactions with these other fixed effects?

Note that in this interaction model, the interaction terms art not significant. We can think that all the covariates and interactions are not for Popular variable. Those are related to Classical ratings.

Question 5

Brief write up. Write a one page professional-quality summary of your findings for Classical and Popular ratings, suitable for Dr. Jimenez. Be sure to address:

- The influence of the three main experimental factors (Instrument, Harmony & Voice);
- A brief discussion of variance components – is this a standard repeated measures model, or did we need to include other variance components?
- A discussion of other individual covariates in the model. You may refer to your earlier work (e.g. "As I showed in my answer to part 1b, blah-blah-blah.."). Don't be sloppy about the statistical findings, but try to highlight things that will be of substantive interest to Dr. Jimenez. Make your summary very readable and clear.

`# Answer`

According to the result from question 3, Instrument and one of the Harmony variables have positive significant relationship with Classical variable. But Voice variable's effect disappeared once including additional covariates and interaction term. With comprehensive test, we found that `NoClass` and `musician` interaction

why does the writeup start in the middle of the analysis instead of the beginning?

has strong negative effect on Classical ratings. Variance components did not change a lot compared to previous mode.

Using Popular variable as a dependent variable, we found that no covariates are actually having association with the original model. So we dropped all the covariates. So we also examined the effect of interaction term, self-perception as a musician. In fixed effect, we found that interaction variables are not significant. However, on the other hand, model comparison using AIC, BIC, and DIC indicate that original model with interaction of musician and Voice variables is the best choice among the choice option. In addition, the effect size of each experimental variables changed a little. But more important thing is that the effect direction changed (in opposite way) from the original model. It is not surprising because rating Classical or Popular are negatively correlated. very incomplete.

4: 3 + 5 + 2 = 10
5: 9

19/40