1 a 9/9
b 9/9
c 9/9

2 a 9/9
b 9/9
c 9/9

3 6/9

4 a 9/9
b 9/9
c 9/9

5 10/10

Total 94/100

36-763 Homework 05

Xi Qu

1 Three main experimental factors

(a) All three factors have a significant influence on the Classical ratings. From Table 1, the partial F test tells us the effect of Harmony, Instrument, Voice are significant if we take them out from the full linear model lm.1 respectively.

Table 2 shows the fixed effects of these factors, and each 2 rows represent the resluts of F test. We can found piano and string quartet instrument indeed have the largest increase of Classical ratings for the greatest magnitude of coefficients; I-V-VI harmonic motion will have a great increase of the Classical ratings;Parrllel 3rds and 5ths tend to decrease the Classical ratings when compared to contrary motion.

The basic linear model is consistent with the researcher's main hypotheses.

I noted that there are 27 missing values of Classical and popular ratings, which come from 24,31,48 and 73 participants. For now we have no data sources to make up these missing values, amd the highest number of missing values for a participant is 12 that we still have 24 Classical ratings for this participant to examine the relationship. So I delete those 27 observations who have no Classical ratings.

```r
ratings <- read.csv("ratings.csv")
ratings <- ratings[!is.na(ratings$Classical),]

lm.1 <- lm(Classical ~ Harmony + Instrument + Voice, data=ratings)

tbl <- data.frame(matrix(ncol = 6, nrow = 6))
lm.1.harmonyout <- lm(Classical ~ Instrument + Voice, data=ratings)
tbl <- anova(lm.1.harmonyout, lm.1)

lm.1.instruout <- lm(Classical ~ Harmony + Voice, data=ratings)
tbl[3:4,] <- anova(lm.1.instruout, lm.1)
lm.1.voiceout <- lm(Classical ~ Instrument + Voice, data=ratings)
tbl[5:6,] <- anova(lm.1.voiceout, lm.1)
```

Table 1: Importance of three main factors

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)	Model
1	2488.00	13381.13					lm model without Harmony
2	2485.00	13107.48	3.00	273.65	17.29	<0.00001	lm model lm.1
3	2487.00	17235.04					lm model without Instrument
4	2485.00	13107.48	2.00	4127.56	391.26	<0.00001	lm model lm.1
5	2488.00	13381.13					lm model without Voice
6	2485.00	13107.48	3.00	273.65	17.29	<0.00001	lm model lm.1

Table 2: Effect of three main factors

		Estimate	Std. Error	t value	Pr(>|t|)
	(Intercept)	4.3402	0.1299	33.42	<0.00001
	HarmonyI-V-IV	-0.0311	0.1301	-0.24	0.81117
	HarmonyI-V-VI	0.7691	0.1301	5.91	<0.00001
9	HarmonyIV-I-V	0.0501	0.1300	0.39	0.70009
	Instrumentpiano	1.3736	0.1130	12.16	<0.00001
	Instrumentstring	3.1331	0.1123	27.90	<0.00001
	Voicepar3rd	-0.4125	0.1127	-3.66	0.00026
	Voicepar5th	-0.3706	0.1126	-3.29	0.00102

(b) (i) The hierarchical linear model:

$$Level2: \alpha_{0j} \stackrel{iid}{\sim} N(\beta_0, \tau^2)$$

$$\begin{aligned} Level1: Classical_i \stackrel{indep}{\sim} N(&\alpha_{0j[i]} + \alpha_{11} * HarmonyI\_V\_IV_i + \alpha_{12} * HarmonyI\_V\_VI_i + \\ &\alpha_{13} * HarmonyIV\_I\_V_i + \alpha_{21} * Instrument\_piano_i + \\ &\alpha_{22} * Instrument\_string_i + \alpha_{31} * Voice\_par3rd_i + \alpha_{32} * Voice\_par5th_i, \sigma^2) \end{aligned}$$

where j=(1,2,...70), denotes the individual participants (our group level in this model).

(ii) Fit the random intercept model and compare the residual variance σ^2 (1.89) which comes from each rating, and variance of individual level τ^2 (1.30), which comes from each participants, we find the variance of ratings from different participants account for a certain weight, making it reasonable to include random intercept in model.

I use overall index of fit based on the deviance and a simulation-based check to test, and they work together to ensure the need of including random intercept.

```r
library(lme4)
library(arm)
lmer.1.intercept <- lmer(Classical ~ Harmony + Instrument + Voice + (1|Subject),
                         data=ratings)
display(lmer.1.intercept)

## lmer(formula = Classical ~ Harmony + Instrument + Voice + (1 |
##     Subject), data = ratings)
##                     coef.est coef.se
## (Intercept)          4.34     0.19
```

```
## HarmonyI-V-IV    -0.03     0.11
## HarmonyI-V-VI     0.77     0.11
## HarmonyIV-I-V     0.05     0.11
## Instrumentpiano   1.38     0.09
## Instrumentstring  3.13     0.09
## 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
```

Method 1. overall index of fit based on AIC/BIC

Since two models are not nested (with different random effect terms), the comparison of deviance based on chie-square distribution is not valid. I will use the penalized deviance measures AIC and BIC to compare models. And I will use anova() function to get AIC/BIC refitted by MLEs.

```
a <- anova(lmer.1.intercept,lm.1)

## refitting model(s) with ML (instead of REML)
```

Table 3: Overall index of fit

	Df	AIC	BIC
lm.1	9	11230.45	11282.84
lmer.1.intercept	10	10468.86	10527.07

From table 3, we find AIC and BIC of the random intercept model decreases greatly. Including random intercept is necessary since it will improve the fitness of model significantly.

Method 2. simulation-based checks

I use the exact test of random effect by simulated finite sample distribution. The test is highly significant (p-value $\leq$ 2.2e-16), thus rejecting the null hypothesis $H_0 : \tau^2 = 0$. Random effect is needed based on simulated data from the fitted model.

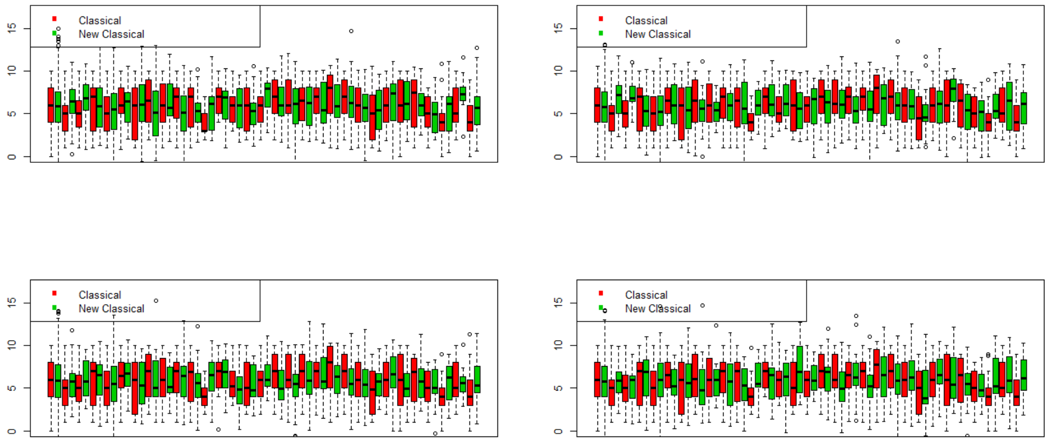
```
library(RLRsim)
exactRLRT(lmer.1.intercept)

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

```
## data:
## RLRT = 763.38, p-value < 2.2e-16
```

In addition, I use the JAGS and rube to simulate new Classical ratings based on linear model (without random intercept). Looking at the spread of the simulated data compared to the original Classical ratings between groups defined by individual participants, we find the red boxplots (actual Classical ratings) seem more variable than the green boxplots (simulated new Classical ratings based on linear model), which reflects the "shrinkage idea" that the linear model smooth out extreme observations and pull predictions back to overall mean. Since variances indeed exist between participants, it is worth to fit participants (Subject) as random effect.

nice!



(iii) All three factors have a significant influence on the Classical ratings based on repeated-measures model. After considering the difference of participants' ratings, instrument, harmonic motion and voice leading still have large influence on Classical ratings.

Table 4 tells us the effect of Harmony, Instrument, Voice are significant since they improve the fit of model significantly (Compare each 2 rows).

We also know from Table 5 that the fixed effect of these 3 variables are similar to the linear model. Again, instrument has the largest influence on rating (greatest magnitude of coefficients).

```
lmer.1.harmonyout <- lmer(Classical ~ Instrument + Voice + (1|Subject),
                          data=ratings)
lmer.1.instruout <- lmer(Classical ~ Harmony + Voice + (1|Subject),
                          data=ratings)
lmer.1.voiceout <- lmer(Classical ~ Harmony + Instrument + (1|Subject),
                          data=ratings)
```


9

Table 4: Importance of three main factors based on repeated-measures model

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)	Model
1	7.00	10538.79	10579.54	-5262.40	10524.79				without Harmony
2	10.00	10468.86	10527.07	-5224.43	10448.86	75.93	3.00	<0.0001	lmer model
3	8.00	11408.45	11455.02	-5696.22	11392.45				without Instrument
4	10.00	10468.86	10527.07	-5224.43	10448.86	943.59	2.00	<0.0001	lmer model
5	8.00	10489.10	10535.67	-5236.55	10473.10				without Voice
6	10.00	10468.86	10527.07	-5224.43	10448.86	24.24	2.00	<0.0001	lmer model

Table 5: Effect of three main factors based on repeated-measures model

	Estimate	Std. Error	t value
(Intercept)	4.34	0.19	22.97
HarmonyI-V-IV	-0.03	0.11	-0.30
HarmonyI-V-VI	0.77	0.11	7.19
HarmonyIV-I-V	0.05	0.11	0.47
Instrumentpiano	1.38	0.09	14.78
Instrumentstring	3.13	0.09	33.83
Voicepar3rd	-0.42	0.09	-4.47
Voicepar5th	-0.37	0.09	-4.03

(c) (i) Table 6 shows the model with all 3 new random effect terms is better than all previous model since the AIC/BIC decreases significantly. It improves the overall fit significantly and can better explain our data.

```
lmer.1.interceptall <- lmer(Classical ~ Harmony + Instrument + Voice +
                              (1|Subject:Instrument) + (1|Subject:Harmony) +
                              (1|Subject:Voice),data=ratings)
```

Table 6: Overall fit of all models

	Df	AIC	BIC
lm model without Harmony	6	11275.96	11310.89
lm model without Voice	6	11275.96	11310.89
lm modelwithout Instrument	7	11908.94	11949.69
lmer model without Harmony	7	10538.79	10579.54
lmer model without Instrument	8	11408.45	11455.02
lmer model without Voice	8	10489.10	10535.67
lm model	9	11230.45	11282.84
lmer model with 1 random effect	10	10468.86	10527.07
lmer model with 3 random effects	12	10057.53	10127.38

(ii) **Influence of 3 main factors**

All three factors have a significant influence on the Classical ratings based on the multilevel model with all three random effect terms. After considering the personal biases of Classical ratings may vary with the type of instrument, harmonic motion and voice leading, we still have the conclusions that instrument, harmonic motion and voice leading themselves have a large influence on Classical ratings.

Table 7 (comparing each 2 rows) tells us the effect of Harmony, Instrument, Voice are significant by comparing the model which takes them out respectively with the full model which has all 3 main factors based on multilevel model with 3 random effect terms.

Table 8 shows the fixed effect of these 3 variables are similar to the linear model and the repeated-measures model. Again, instrument has the largest influence on rating (greatest magnitude of coefficients).

```
lmer.1.intall.harmonyout <- lmer(Classical ~ Instrument + Voice +
                                 (1|Subject:Instrument) + (1|Subject:Harmony) +
                                 (1|Subject:Voice),data=ratings)
tbl <- anova(lmer.1.interceptall,lmer.1.intall.harmonyout)

lmer.1.intall.instruout <- lmer(Classical ~ Harmony + Voice +
                                 (1|Subject:Instrument) + (1|Subject:Harmony) +
                                 (1|Subject:Voice),data=ratings)
tbl[3:4,] <- anova(lmer.1.interceptall,lmer.1.intall.instruout)

lmer.1.intall.voiceout <- lmer(Classical ~ Harmony + Instrument +
                              (1|Subject:Instrument) + (1|Subject:Harmony) +
                              (1|Subject:Voice),data=ratings)

tbl[5:6,] <- anova(lmer.1.interceptall,lmer.1.intall.voiceout)
```

Table 7: Importance of three main factors based on multilevel model with 3 random effects (lmer.all in the table and lmer.1.interceptall in my code)

	Df	AIC	BIC	deviance	Chisq	Chi Df	Pr(>Chisq)	Model
1	9	10090.54	10142.93	10072.54				lmer.all without Harmony
2	12	10057.53	10127.38	10033.53	39.01	3	<0.0001	lmer.all
3	10	10160.42	10218.63	10140.42				lmer.all without Instrument
4	12	10057.53	10127.38	10033.53	106.89	2	<0.0001	lmer.all
5	10	10081.28	10139.49	10061.28				lmer.all without Voice
6	12	10057.53	10127.38	10033.53	27.75	2	<0.0001	lmer.all

Table 8: Effect of three main factors based on multilevel model with 3 random effects (lmer.all in the table and lmer.1.interceptall in my code)

	Estimate	Std. Error	t value
(Intercept)	4.34	0.21	20.25
HarmonyI-V-IV	-0.03	0.14	-0.21
HarmonyI-V-VI	0.77	0.14	5.38
HarmonyIV-I-V	0.06	0.14	0.39
Instrumentpiano	1.36	0.26	5.20
Instrumentstring	3.13	0.26	11.94
Voicepar3rd	-0.41	0.08	-4.98
Voicepar5th	-0.37	0.08	-4.54

Comparisons of estimated variance components

Table 9 tells that the variance component is biggest for Subject:Instrument, then for Subject:Harmony, and smallest for Subject:Voice. The bigger variance, the more useful of including the random effect term, since it reflects the differences of how people are inclined to rating a music as classical when encoutered that characteristic. We can see the variance component of Subject:Instrument is biggest (2.20), which is close to the residual variance σ^2 (2.44) that comes from each rating. Thus the personal biases vary greatly with the type of instrument. Instrument types not only influence the Classical ratings but also influence the difference of participants' Classical ratings. However, the variance component of Subject:Voice is very small, indicating that people vary little in the degree to which they are inclined to rate music as Classical by voice leading types. It contributes less in our model to explain the variances of ratings for different combinations of voice leading type and participants.

Table 9: Comparison of variance components

	Groups	Name	Variance	Standard Deviation
1	Subject:Harmony	(Intercept)	0.44	0.67
2	Subject:Voice	(Intercept)	0.03	0.17
3	Subject:Instrument	(Intercept)	2.20	1.48
4	Residual		2.44	1.56

(iii) The hierarchical linear model:

9

$$\begin{aligned} Level2: \alpha_{1jk} &\overset{iid}{\sim} N(0, \tau_1^2) \\ \alpha_{2jl} &\overset{iid}{\sim} N(0, \tau_2^2) \\ \alpha_{3jm} &\overset{iid}{\sim} N(0, \tau_3^2) \end{aligned}$$

$$\begin{aligned} Level1: Classical_i \overset{indep}{\sim} & N(\alpha_0 + \alpha_{1j[i]k[i]} + \alpha_{2j[i]l[i]} + \alpha_{3j[i]m[i]} + \beta_{11} * HarmonyI\_V\_IV_i + \\ & \beta_{12} * HarmonyI\_V\_VI_i + \beta_{13} * HarmonyIV\_I\_V_i + \beta_{21} * Instrument\_piano_i + \\ & \beta_{22} * Instrument\_string_i + \beta_{31} * Voice\_par3rd_i + \beta_{32} * Voice\_par5th_i, \sigma^2) \end{aligned}$$

j = (1,...70), denotes participant subjects; k= 1,2,3,4, denotes Harmonic motion; l = (1,2,3), denotes instrument types; m = (1,2,3), denotes voice leading types.

2 Individual covariates

(a) Table 11 shows the final set of fixed effect variables in the model. I will present my work process through the following description and table 10.

My best model from part 1 is the multilevel model with 3 random effects. I'm interested in variables Selfdeclare, OMSI, X16.minus.17, ConsInstr, ConsNotes,PianoPlay, GuitarPlay and ClsListen. Although Consnotes have missing values, but we can calculate it by ConsInstr - Instr.minus.Notes so there is little problem. As for ClsListen, it has 24 missing values. Since all of it come from 24th participant, I consider removing observations from the 24th participant and set a dataset "newrate". I refit my best model on the newrate dataset, it turns out there is little change of the model output. My best model from part 1 is robust after removing one participants and I can confidently continue my analysis on "newrate" dataset.

```
ratings$ConsNotes <- ratings$ConsInstr-ratings$Instr.minus.Notes
newrate <- ratings[ratings$Subject!=24,]
lmer.2.interceptall <- lmer(Classical ~ Harmony + Instrument + Voice +
                            (1|Subject:Instrument) + (1|Subject:Harmony) +
                            (1|Subject:Voice),data=newrate)
display(lmer.2.interceptall)
```

```
## lmer(formula = Classical ~ Harmony + Instrument + Voice + (1 |
##     Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice),
##     data = newrate)
##                  coef.est coef.se
## (Intercept)       4.32     0.22
## HarmonyI-V-IV    -0.03     0.14
## HarmonyI-V-VI     0.78     0.14
## HarmonyIV-I-V     0.05     0.14
## Instrumentpiano   1.37     0.27
## Instrumentstring  3.15     0.27
## Voicepar3rd      -0.40     0.08
## Voicepar5th      -0.36     0.08
##
## Error terms:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.67
##  Subject:Voice      (Intercept) 0.17
##  Subject:Instrument (Intercept) 1.49
##  Residual                       1.56
## ---
## number of obs: 2469, groups: Subject:Harmony, 276; Subject:Voice, 207; Subject:Instru
## AIC = 9967.4, DIC = 9907.7
## deviance = 9925.5
```

Table 10 shows my work of finding the best model with new individual covariates.

I first plug all of my interested individual variables into the model (lmer.2.try1), and compare the overall fit with my best model (lmer.2.interceptall). It turns out including all of them will improve the overall fit. I will replace my best model with lmer.2.try1.

I then looked at the output of each individual covariates, and find the coefficients of

OMSI, ConsInstr, ConsNotes, PianoPlay and GuitarPlay are not significant. I then remove them in the model lmer.2.try2 and compare to the lmer.2.try1. The chi-square test is not significant, indicating it's OK to removing them since they contribute less to explaining data. AIC and BIC are smaller for lmer.2.try2. I will replace my best model with lmer.2.try2.

After that, I find the coefficient of Selfdeclare is not significant. I then try removing it in the model lmer.2.try3 and compare to the lmer.2.try2. The chi-square test is also not significant, indicating it's OK to removing Selfdeclare. AIC are similar and BIC of lmer.2.try3 is better. I will keep model lmer.2.try3 as my final model.

Table 11 shows the final set of fixed effect variables in the model. And my final model contains all the 3 random effects as problem 1 does.

```
lmer.2.try1 <- lmer(Classical ~ Harmony + Instrument + Voice + Selfdeclare +
                      OMSI + X16.minus.17 + ConsInstr + ConsNotes +
                      PianoPlay + GuitarPlay + ClsListen +
                      (1|Subject:Instrument) + (1|Subject:Harmony) +
                      (1|Subject:Voice),data=newrate)
tbl <- anova(lmer.2.try1,lmer.2.interceptall)
#display(lmer.2.try1)

lmer.2.try2 <- lmer(Classical ~ Harmony + Instrument + Voice + Selfdeclare +
                      X16.minus.17 + ClsListen +
                      (1|Subject:Instrument) + (1|Subject:Harmony) +
                      (1|Subject:Voice),data=newrate)
tbl[3:4,] <- anova(lmer.2.try2,lmer.2.try1)
#display(lmer.2.try2)

lmer.2.try3 <- lmer(Classical ~ Harmony + Instrument + Voice +
                      X16.minus.17 + ClsListen +
                      (1|Subject:Instrument) + (1|Subject:Harmony) +
                      (1|Subject:Voice),data=newrate)
tbl[5:6,] <- anova(lmer.2.try3,lmer.2.try2)
```

Table 10: Process of finding important individual covariates

	Df	AIC	BIC	deviance	Chisq	Chi Df	Pr(>Chisq)	Model
1	12	9949.52	10019.26	9925.52				lmer.2.try1
2	20	9944.53	10060.77	9904.53	20.99	8	0.0072	lmer.2.interceptall
3	15	9942.04	10029.21	9912.04				lmer.2.try2
4	20	9944.53	10060.77	9904.53	7.51	5	0.1856	lmer.2.try1
5	14	9942.10	10023.46	9914.10				lmer.2.try3
6	15	9942.04	10029.21	9912.04	2.06	1	0.1513	lmer.2.try2

9

Table 11: Final set of variables of fixed effects

	Estimate	Std. Error	t value
(Intercept)	4.19	0.27	15.55
HarmonyI-V-IV	-0.03	0.14	-0.22
HarmonyI-V-VI	0.78	0.14	5.43
HarmonyIV-I-V	0.05	0.14	0.37
Instrumentpiano	1.37	0.26	5.29
Instrumentstring	3.15	0.26	12.17
Voicepar3rd	-0.40	0.08	-4.89
Voicepar5th	-0.36	0.08	-4.36
X16.minus.17	-0.11	0.04	-2.79
ClsListen	0.14	0.07	2.00

(b) **Checking process and evidence for best model**

Table 12 shows the process of checking random effects. The model with $(1|Subject : Harmony)$ and $(1|Subject : Instrument)$ random effects is the best model.

Since the variance components of $(1|Subject : Instrument)$ is biggest and the variance components of $(1|Subject : Voice)$ is smallest, and a higher variance component has more explaining ability for data variances. I need to compare our best model lmer.2.try3 (with 3 random effect terms) in last problem with the model with only $(1|Subject : Instrument)$ random effect and the model with $(1|Subject : Instrument)$ and $(1|Subject : Harmony)$ random effects. I also compare with the model wich only have $(1|Subject)$ as the random effect.

AIC and BIC is smallest for model with 2 random effects. And we've already examined
9 before that the variance components of $(1|Subject : Voice)$ is small, suggesting a less variability of different rating inclination of person/voice leading combinations. I choose the model with 2 random effects.

```
lmer.2.try3.subint <- lmer(Classical ~ Harmony + Instrument + Voice +
                           X16.minus.17 + ClsListen +(1|Subject),data=newrate)
lmer.2.try3.subinstr <- lmer(Classical ~ Harmony + Instrument + Voice +
                              X16.minus.17 + ClsListen +
                              (1|Subject:Instrument), data=newrate)
lmer.2.try3.Voiceout <- lmer(Classical ~ Harmony + Instrument + Voice +
                              X16.minus.17 + ClsListen +
                              (1|Subject:Instrument) +  (1|Subject:Harmony),
                              data=newrate)
tbl <- anova(lmer.2.try3.subint, lmer.2.try3.subinstr,lmer.2.try3,lmer.2.try3.Voiceout)[
```

Table 12: Overall fit for 3 models with different random effects

	Df	AIC	BIC
only (1|subject) random effect	12	10361.37	10431.11
only (1|Subject:Instrument) random effect	12	10039.65	10109.39
(1|Subject:Harmony)&(1|Subject:Instrument)	13	9940.85	10016.40
(1|Subject:Harmony),(1|Subject:Instrument)&(1|Subject:Voice)	14	9942.10	10023.46

(c) The coefficients of each variable is shown in Table 13.

very complete. thx!

Interpretation:

Keeping other factors constant, we have the following change of average Classical ratings:

HarmonyI_V_IV : When harmonic motion of music changes from I_VI_V to I_V_IV, the average Classical rating will decrease 0.03 points as a fixed effect.

HarmonyI_V_VI : When harmonic motion of music changes from I_VI_V to I_V_VI, the average Classical rating will increase 0.78 points as a fixed effect.

HarmonyIV_I_V : When harmonic motion changes from I_VI_V to IV_I_V, the average Classical rating will increase 0.05 points as a fixed effect.

Instrumentpiano : When instrument type changes from electric guitar to piano, the average Classical rating will increase 1.37 points as a fixed effect.

Instrumentstring : When instrument type changes from electric guitar to string quartet, the average Classical rating will increase 3.15 points as a fixed effect.

Voicepar3rd : When voice leading type changes from contrary motion to parallel 3rds, the average Classical rating will decrease 0.4 points as a fixed effect.

Voicepar5th : When voice leading type changes from contrary motion to parallel 5ths, the average Classical rating will decrease 0.36 points as a fixed effect.

X16.minus.17 : When auxiliary measure of listener's ability to distinguish classical and popular music increase by 1 degree, average Classical ratings decrease 0.11 points.

ClsListen : When the degree of listening Classical music of listeners increase by 1, the average Classical rating increase 0.14 points.

9

In overall, the 3 main hypotheses were correct. Instrument types have the greatest influence on rating; I_V_VI harmonic progression is associated with a high frequency of being rated as Classical; contrary motion of voice leading type is also frequently rated as Classical.

We also find some individual covariates which influence the Popular ratings. Controlling other variables the same, the higher ability to distinguish Classical and Popular music (higher X16.minus.17) decreases the average Popular ratings, and a higher frequency of listening classical music increases the Classical ratings.

Except the fixed effects, we also have random effects of intercept for the different combinations of person/instrument type and person/harmonic motion, which allows the difference of people's Classical ratings when encountered those instrument type and harmonic motion. That is, the different degrees for people to rate music as Classical is influenced by intrument types and harmonic motions.

```
lmer.2.final <- lmer(Classical ~ Harmony + Instrument + Voice +
                     X16.minus.17 + ClsListen +
                     (1|Subject:Instrument) +  (1|Subject:Harmony),data=newrate)
```

Table 13: Fixed effects of final model after adding individual covariates

	Estimate	Std. Error	t value
(Intercept)	4.19	0.27	15.56
HarmonyI-V-IV	-0.03	0.14	-0.22
HarmonyI-V-VI	0.78	0.14	5.43
HarmonyIV-I-V	0.05	0.14	0.37
Instrumentpiano	1.37	0.26	5.26
Instrumentstring	3.15	0.26	12.11
Voicepar3rd	-0.40	0.08	-5.18
Voicepar5th	-0.36	0.08	-4.61
X16.minus.17	-0.11	0.04	-2.80
ClsListen	0.14	0.07	2.00

3 Musicians vs. Non-musicians

I split the Selfdeclare variable at 2 and set a new variable "Musician" thus there are 1475 non-musicians and 994 musicians in my data set.

there aren't that many participants i in the data set so you must mean somethnig else...

I start from the best model in problem 2. Since we want to examine whether there is interaction between musician and other variables, I consider adding random effect terms of combination between musician and Instrument, Harmony and Voice. I use the fitLMER.fnc to automatically find the best model based on backward elimination of fixed effects and forward selection of random effects. I use AIC and BIC criteria for fixed effect backward elimination respectively.

Table 14 shows that AIC for the model picked by AIC backward elimination is smaller while the BIC are the same for both models. I will use the model picked by AIC backward elimination.

```
newrate$musician[newrate$Selfdeclare>2] <- 1
newrate$musician[newrate$Selfdeclare<=2] <- 0
newrate$musician <- factor(newrate$musician,labels=c("nonmusician","musician"))

library(LMERConvenienceFunctions)
lmer.3.big <- lmer(Classical ~ Harmony + Instrument + Voice + X16.minus.17 +
                   ClsListen + (1|Subject:Instrument)+(1|Subject:Harmony),
                    data=newrate)
lmer.3.aic.best <- fitLMER.fnc(lmer.3.big,method="AIC",
                        ran.effects = c("(1|musician:Harmony)",
                                        "(1|musician:Instrument)",
                                        "(1|musician:Voice)"))
lmer.3.bic.best <- fitLMER.fnc(lmer.3.big,method="BIC",
                        ran.effects = c("(1|musician:Harmony)",
                                        "(1|musician:Instrument)",
                                        "(1|musician:Voice)"))
```

what about interactions of musician with the fixed effects?

6

at the very least, need fixed effects corresponding to the interactions of musician with h, v, and i you are considering in the random effects

```
## refitting model(s) with ML (instead of REML)
```

Table 14: Overall fit for two picked models

	Df	AIC	BIC
best model picked by BIC	9	9943.21	9995.52
best model picked by AIC	10	9937.41	9995.52

Table 15: Final model after considering self-declared musicians

	Estimate	Std. Error	t value
(Intercept)	4.72	0.27	17.47
Instrumentpiano	1.37	0.27	5.17
Instrumentstring	3.15	0.27	11.89
Voicepar3rd	-0.40	0.08	-5.17
Voicepar5th	-0.35	0.08	-4.61
X16.minus.17	-0.11	0.04	-2.80

Table 15 shows the coefficients of fixed coefficients of the model. Except the original random effect terms, we add a new random effect term $(1|musician : Harmony)$, suggesting the Classical ratings of people who self identify as musicians may be influenced by harmonic motion types different from non-musicians.

Table 16 shows the variance components of these random effect terms. Again, the combination between subject and instrument has the greatest variances. In addition, the variance components of $(1|musician : Harmony)$ is not small. We are quite convinced by our hypothesis and conclude that the Classical ratings for people who have self identification as musicians are influenced greatly by harmonic motion types.

Table 16: Comparison of variance components

	Groups	Name	Variance	Standard Deviation
1	Subject:Harmony	(Intercept)	0.36	0.60
2	Subject:Instrument	(Intercept)	2.22	1.49
3	musician:Harmony	(Intercept)	0.23	0.48
4	Residual		2.44	1.56

4 Classical vs Popular

4.1 (a). Influence of Instrument, Harmony, Voice on Popular Ratings

The influence of Instrument on Popular ratings are significant, but the Harmony and Voice become not significant after accounting for the interaction random effect terms of their combination with participant subjects. I will show evidence in the following.

Influence of 3 main factors on Popular ratings

Table 17 shows that, after considering the personal bias of popular ratings may vary with the type of instrument, harmonic motion and voice leading, only instrument itself has a large influence on Popular ratings, while Harmony and Voice themselves have no significant influence on Popular ratings. We compare the model which takes Instrument, Harmony and Voice out respectively with the full model which has all 3 main factors.

Table 18 shows the fixed effect of these 3 variables. Again, instrument has the largest influence on Popular ratings while other variables have no significant influence.

```r
lmer.pop.all <- lmer(Popular ~ Instrument + Harmony + Voice +
                    (1|Subject:Instrument) + (1|Subject:Harmony) +
                    (1|Subject:Voice),data=ratings)
lmer.pop.harmonyout <- lmer(Popular ~ Instrument + Voice +
                           (1|Subject:Instrument) + (1|Subject:Harmony)+
                           (1|Subject:Voice),data=ratings)
tbl <- anova(lmer.pop.all,lmer.pop.harmonyout)

lmer.pop.instruout <- lmer(Popular ~ Harmony + Voice +
                          (1|Subject:Instrument) + (1|Subject:Harmony)+
                          (1|Subject:Voice),data=ratings)
tbl[3:4,] <- anova(lmer.pop.all,lmer.pop.instruout)

lmer.pop.voiceout <- lmer(Popular ~ Harmony + Instrument +
                         (1|Subject:Instrument) + (1|Subject:Harmony)+
                         (1|Subject:Voice),data=ratings)
tbl[5:6,] <- anova(lmer.pop.all,lmer.pop.voiceout)
```

Table 17: Importance of 3 main factors based on multilevel model for Popular ratings

	Df	AIC	BIC	deviance	Chisq	Chi Df	Pr(>Chisq)	Model
1	9	10078.09	10130.48	10060.09				without Harmony
2	12	10078.97	10148.82	10054.97	5.12	3	0.1634	full lmer model
3	10	10161.84	10220.05	10141.84				without Instrument
4	12	10078.97	10148.82	10054.97	86.87	2	<0.0001	full lmer model
5	10	10080.05	10138.26	10060.05				without Voice
6	12	10078.97	10148.82	10054.97	5.08	2	0.0789	full lmer model

Table 18: Effect of three main factors based on multilevel model for Popular ratings

	Estimate	Std. Error	t value
(Intercept)	6.58	0.21	31.77
Instrumentpiano	-0.95	0.25	-3.77
Instrumentstring	-2.61	0.25	-10.37
HarmonyI-V-IV	-0.03	0.14	-0.18
HarmonyI-V-VI	-0.27	0.14	-1.93
HarmonyIV-I-V	-0.19	0.14	-1.32
Voicepar3rd	0.16	0.08	1.97
Voicepar5th	0.16	0.08	1.95

Comparisons of estimated variance components

Table 19 tells that the variance component is biggest for Subject:Instrument, then for Subject:Harmony, and smallest for Subject:Voice. The bigger the variance, the more useful of including this random effect term in our model. Thus the personal biases of Popular ratings vary greatly with the type of instrument. Instrument types not only influence the Classical ratings but also influence the difference of participants' Popular ratings.

9 Again, the variance component of Subject:Voice is small, indicating that people vary little in the degree to which they are inclined to rate music as Popular by voice leading types. It contributes less in our model to explain the variances of ratings for different combinations of voice leading type and participants.

Table 19: Comparison of variance components

	Groups	Name	Variance	Standard Deviation
1	Subject:Harmony	(Intercept)	0.41	0.64
2	Subject:Voice	(Intercept)	0.03	0.18
3	Subject:Instrument	(Intercept)	2.00	1.41
4	Residual		2.49	1.58

4.2 (b). Individual Covariates for Popular Ratings

Table 21 shows the final set of fixed effect variables of the model for Popular ratings. I will present my work process through the following description and Table 20.

Except the previous interesting individual covariates Selfdeclare, OMSI, X16.minus.17, ConsInstr, ConsNotes, PianoPlay, GuitarPlay, I'm also interested in variable PachListen, since the Pachelbel's Canon in D is common chord progression in popular music. As for PachListen, all 36 missing values come from 76th participant, I consider removing those observations and set a dataset "newpop".

I first plug all of my interested individual variables into the model (lmer.4.try1), and compare the overall fit with my previous best model (lmer.4.prebest). It turns out including all of them will improve the overall fit (Chi-square test significant). I will replace my best model with lmer.4.try1.

I then looked at the output of lmer.4.try1. OMSI, ConsInstr, PianoPlay, GuitarPlay and PachListen are highly nonsignificant. Also, the variance components of Subject:Voice is small. I then consider removing them in the model lmer.4.try2 and compare AIC/BIC with the lmer.4.try1 since they're not nested. AIC and BIC are greatly smaller for lmer.4.try2. I will replace my best model with lmer.4.try2.

After that, I find the coefficient of Selfdeclare and ConsNotes is not significant. I then try removing them seperately and altogether, then compare the four models. AIC and BIC of try4 are smallest. I will keep model lmer.4.try4 as my final model.

```
#newrate[is.na(newrate£PachListen),c(2,12,13)]
newpop <- newrate[newrate$Subject!=76,]
lmer.4.prebest <- lmer(Popular ~ Instrument + Harmony + Voice +
                      (1|Subject:Instrument) + (1|Subject:Harmony) +
                      (1|Subject:Voice),data=newpop)
```

```r
lmer.4.try1 <- lmer(Popular ~ Harmony + Instrument + Voice + X16.minus.17 +
                      OMSI+ Selfdeclare + ConsInstr + ConsNotes +
                      PianoPlay + GuitarPlay + PachListen +
                      (1|Subject:Instrument) + (1|Subject:Harmony) +
                      (1|Subject:Voice),data=newpop)
#display(lmer.4.try1)
tbl <- anova(lmer.4.prebest,lmer.4.try1)

lmer.4.try2 <- lmer(Popular ~ Harmony + Instrument + Voice + X16.minus.17 +
                      Selfdeclare  + ConsNotes +
                      (1|Subject:Instrument) + (1|Subject:Harmony) ,
                      data=newpop)
#display(lmer.4.try2)
tbl[3:4,] <- anova(lmer.4.try1,lmer.4.try2)
tbl <- tbl[-4,]

lmer.4.try3 <- lmer(Popular ~ Harmony + Instrument + Voice + X16.minus.17 +
                      Selfdeclare +(1|Subject:Instrument) + (1|Subject:Harmony),
                      data=newpop)

lmer.4.try4 <- lmer(Popular ~ Harmony + Instrument + Voice + X16.minus.17 +
                      ConsNotes + (1|Subject:Instrument) + (1|Subject:Harmony) ,
                      data=newpop)

lmer.4.try5 <- lmer(Popular ~ Harmony + Instrument + Voice + X16.minus.17 +
                      (1|Subject:Instrument) + (1|Subject:Harmony) ,
                      data=newpop)

tbl[4:6,] <-anova(lmer.4.try3,lmer.4.try4,lmer.4.try5)
```

Table 20: Model selection for Popular ratings

	Df	AIC	BIC	deviance	Chisq	Chi Df	Pr(>Chisq)
lmer.4.prebest	12	9832.43	9902.00	9808.43			
lmer.4.try1	20	9832.69	9948.63	9792.69	15.742	8	0.0462
lmer.4.try2	14	9823.87	9905.03	9795.87			
lmer.4.try5	12	9828.22	9897.79	9804.22			
lmer.4.try3	13	9824.75	9900.11	9798.75			
lmer.4.try4	13	9824.30	9899.66	9798.30			

Table 21: Fixed effects of model for Popular ratings

	Estimate	Std. Error	t value
(Intercept)	6.16	0.25	24.70
HarmonyI-V-IV	-0.02	0.14	-0.15
HarmonyI-V-VI	-0.29	0.14	-2.03
HarmonyIV-I-V	-0.20	0.14	-1.36
Instrumentpiano	-0.91	0.25	-3.66
Instrumentstring	-2.59	0.25	-10.40
Voicepar3rd	0.18	0.08	2.34
Voicepar5th	0.15	0.08	1.96
X16.minus.17	0.10	0.04	2.51
ConsNotes	0.13	0.06	2.42

The coefficients of each variable is shown in Table 21.

Interpretation:

Keeping other factors constant, we have the following change of average Popular ratings:

HarmonyI_V_IV : When the harmonic motion of music people heard changes from I_VI_V to I_V_IV, the average Popular rating will decrease 0.02 points as a fixed effect.

HarmonyI_V_VI : When the harmonic motion of music people heard changes from I_VI_V to I_V_VI, the average Popular rating will decrease 0.29 points as a fixed effect.

HarmonyIV_I_V : When the harmonic motion people heard changes from I_VI_V to IV_I_V, the average Popular rating will decrease 0.2 points as a fixed effect.

Instrumentpiano : When the instrument type people heard changes from electric guitar to piano, the average Popular rating will decrease 0.91 points as a fixed effect.

Instrumentstring : When the instrument type people heard changes from electric guitar to string quartet, the average Popular rating will decrease 2.59 points as a fixed effect.

Voicepar3rd : When the voice leading type people heard changes from contrary motion to parallel 3rds, the average Popular rating will increase 0.18 points as a fixed effect.

Voicepar5th : When the voice leading type people heard changes from contrary motion to parallel 5ths, the average Popular rating will increase 0.15 points as a fixed effect.

X16.minus.17 : When the auxiliary measure of listener's ability to distinguish Classical and popular music increase by 1 degree, the average Popular rating will increase 0.10 points.

ClsListen : When the degree of people's concentration on notes while listening increase by 1, the average Popular rating will increase 0.13 points.

9

In overall, the influence of the 3 main factors for Popular ratings is opposite to that for Classical ratings, and further proved the hypotheses. Instrument types have the greatest influence on Popular rating; I_V_VI harmonic progression is associated with a low frequency of being rated as Popular; contrary motion of voice leading type is also less frequently rated as Popular.

We also find some individual covariates which influence the Popular ratings. Controlling other variables the same, the higher ability to distinguish Classical and Popular music (higher X16.minus.17) increases the average Popular ratings, and the more concentrated on the notes while listening also increases the average Popular ratings.

Table 22 shows the variance components of random effects for the different combinations of person/instrument type and person/harmonic motion, which allows the difference of people's Popular ratings when encountered those instrument type and harmonic motion. The variance component of Subject:Instrument is highest, the personal biases of Popular ratings also vary greatly with the type of instrument.

Table 22: Comparison of variance components

	Groups	Name	Variance	Standard Deviation
1	Subject:Harmony	(Intercept)	0.42	0.65
2	Subject:Instrument	(Intercept)	1.90	1.38
3	Residual		2.51	1.58

4.3 (c). Interaction Effect of Musicians on Popular ratings

I will start from best model in problem 4(b). Again, I consider adding random effect terms of combination between musician and Instrument, Harmony and Voice. I use the fitLMER.fnc to automatically find the best model. I use AIC and BIC criteria for fixed
9 effect backward elimination respectively. It turns out both model are the same that I will use this picked model as my final model.

```
lmer.4.big <- lmer(Popular ~ Harmony + Instrument + Voice + X16.minus.17 +
                        ConsNotes + (1|Subject:Instrument) +
                        (1|Subject:Harmony), data=newpop)
lmer.4.aic.best <- fitLMER.fnc(lmer.4.big,method="AIC",
                        ran.effects = c("(1|musician:Harmony)",
                                        "(1|musician:Instrument)",
                                        "(1|musician:Voice)"))
lmer.4.bic.best <- fitLMER.fnc(lmer.4.big,method="BIC",
                        ran.effects = c("(1|musician:Harmony)",
                                        "(1|musician:Instrument)",
                                        "(1|musician:Voice)"))
```

same problems as in #3...

Table 23: Final model after considering musicians for Popular ratings

	Estimate	Std. Error	t value
(Intercept)	6.61	0.20	32.56
Instrumentpiano	-0.91	0.25	-3.58
Instrumentstring	-2.59	0.25	-10.19

Table 23 shows the coefficients of the fixed effect terms of the model. Now we only have Instrument variable as fixed effects.

Table 24 shows the variance components of these random effect terms. Except the original random effect terms, we add a new random effect term $(1|musician : Harmony)$, suggesting the Popular ratings of people who self identify as musicians are influenced by harmonic motion types different from non-musicians. Again, the combination between subject and instrument has the greatest variances.

Table 24: Comparison of variance components

	Groups	Name	Variance	Standard Deviation
1	Subject:Harmony	(Intercept)	0.39	0.62
2	Subject:Instrument	(Intercept)	1.99	1.41
3	musician:Harmony	(Intercept)	0.06	0.24
4	Residual		2.51	1.59

5 Brief Summary

Introduction

We want to test how instrument, harmony motion and voice leading influence people's Classical and Popular ratings. We can expect there is difference between people's inclination of rating, thus we should consider the personal biases and interaction between personal biases with other covariates to better examine influences of the 3 factors.

Method

I use F test to test the influence of main 3 factors in hierarchical models, which allows the personal biases and interaction between individual covariates and other predictors. AIC/BIC criteria are used to compare model. The magnitude of variance component and coefficient are compared to find the most influential covariates.

Results

(1.) Instrument, harmonic motion and voice leading all have a strong influence on how people rating music as classical (Part 1.c.ii), While only instrument has a strong influence on Popular ratings after considering the personal bias may vary with the type of instrument, harmonic motion and voice leading (Part 4.a).

(2.) The effects of these factors are consistent with researchers' hypotheses: instrument have the largest influence on both Classical and Popular ratings not only as a fixed effect but also have greatest variance on personal biases (Part 1.c.ii and 4.a); holding other factors constant, string quartet and piano are associated with a higher Classical ratings and lower Popular ratings in average; I-V-VI harmony motion is frequently rated classical and less Popular; the contrary motion voice leading type will lead to a
10 higher Classical ratings and lower Popular ratings in average(Part 1.c.ii and 2.c).

(3.) Besides, instrument type and harmony motion will influence the Classical/Popular ratings on different people (Part 2.b and 4.c). That is, the personal biases of Classical/Popular ratings vary in the degree which they are inclined to call music as Classical/Popular is influenced by the instrument and harmony motion. Thus, this is not a standard repeated measures model.

(4.) In addition, I considered some individual predictors (Part 2.a and 4.b). Holding others constant, a higher ability to distinguish classical and popular music for listeners is associated with a lower Classical ratings but higher Popular ratings in average; increasing frequency of listening classical music will increase average Classical ratings; more concentrated on notes will lead to higher average Popular ratings.

(5.) I also considered interaction of self-declared musician situation and harmonic motion. The Classical/Popular ratings for people who are self-declared as musicians are influenced differently compared with non-musicians by the harmony types they heard (Part 3 and 4.c).

Discuss

We considered main 3 experimental predictors, personal bias (random intercept), individual covariates, and the interaction between individual covariates and other predictors. The results are consistent with our hypotheses. The limitation of the analysis is that we have too many missing values that we ignore some possible confounding variables and we should investigate them in later analysis.