

Emotion Perception in Music — Contribution of Primary Music Cues: Critique and Remodelling

Echo Liu

December 5, 2018

Background Information:

Collecting music emotion data is a very subjective process, since people have different understandings of one piece of music under different contexts and associate music with their memories and life experience. In the research field of music emotion, researchers often separate extra-musical factors with factors inherent in music that cause emotions to be perceived by listeners.

T. Eerola, in order to study the latter aspect, conducted an optimized fractional factorial study on 7 primary cues and collected emotional ratings by 46 participants on a total of 200 musical samples according to 4 perceived emotional characters (happy, sad, scary and peaceful). In his paper “Emotional expression in music: contribution, linearity, and additivity of primary musical cues”, Eerola argued that many primary cues contributed to emotion in a linear and non-interactive fashion. However, my exploration of this dataset yields a different result, contradicting his linearity of cue levels argument.

Fractional factorial design: A full factorial experiment is an experiment whose design consists of two or more factors, each with discrete possible values or “levels”, and whose experimental units take on all possible combinations of these levels across all such factors (wiki). If T.Eerola is going to conduct a full factorial experiment, he needs to manipulate $6 \times 2 \times 5 \times 5 \times 4 \times 3 \times 4 = 14400$ musical samples. Instead, he did an optimized fractional factorial study with 200 musical samples that he claimed could “allow the research resources to be concentrated on particular questions, thereby minimizing redundancy and maximizing the statistical power”. However, I doubt if some interaction effects are effaced in this reduction process.

Research Question:

- How do primary cues contribute to different emotional expressions in music?
- Can we predict both happiness and sadness rating for music given single set of primary cues?
- Limitation of T. Eerola’s original model

Data set

Outcome Variables	Description
Sad, Happy, Scary and Peaceful	Emotion ratings for sadness, happiness, scariness and peacefulness

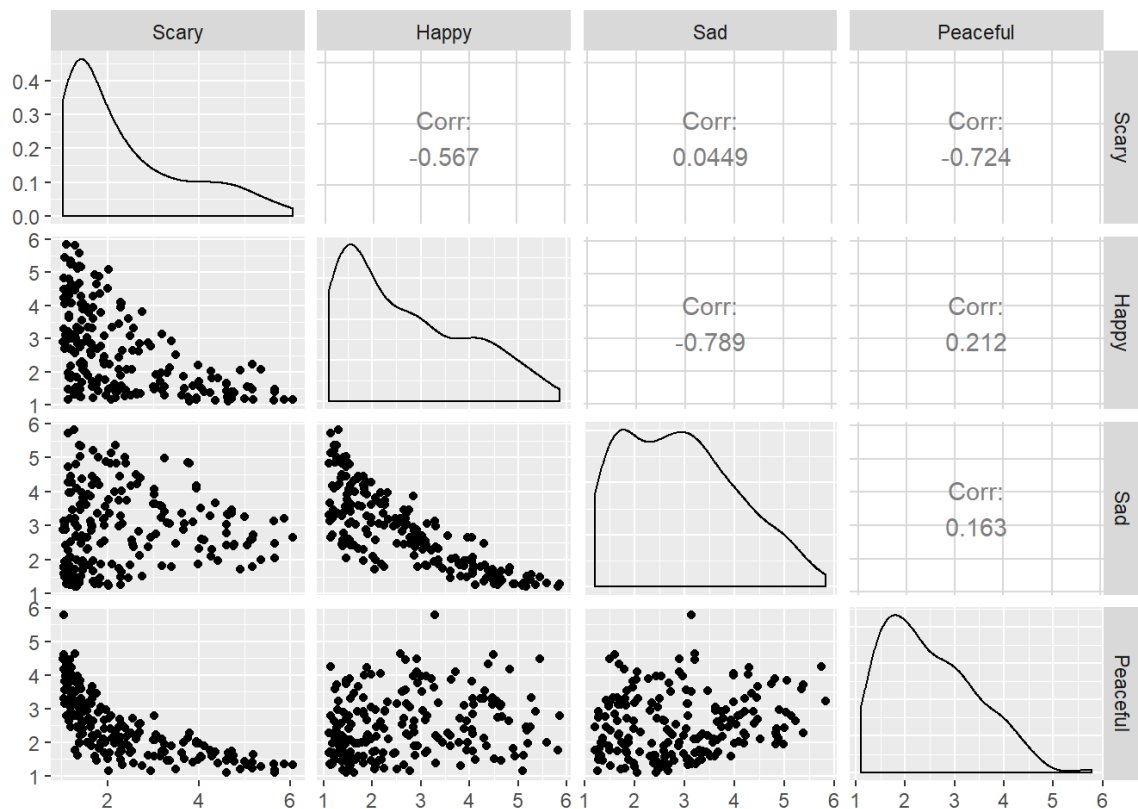
Predictor Variables	Description
Register (6 scalar levels)	Whole piece was transposed so that the average pitches of the melody were the following: F3, B3, F4, B4, F5, B5/ 53, 59, 65, 71, 77 and 83 in MIDI pitch
Mode (2 factor levels)	1 = Major, 2 = Minor

Predictor Variables	Description
Tempo (5 scalar levels)	Average number of non-simultaneous onsets per second (1.2, 2, 2.8, 4.4 and 6 NPS)
Sound Level (5 scalar levels)	-10, -5, 0, +5, +10 dB
Articulation (4 scalarlevels)	Duration of a note relative to its interonset interval
Timbre (3 scalar levels)	The estimation of brightness: 1 = flute, 2 = horn, 3 = trumpet
Music Structure (4 factor categories)	1 = Sad, 2 = Happy, 3 = Scary, 4 = Peaceful

Exploratory Data Analysis

```
##      Nro   Scary   Happy    Sad Peaceful Register Mode Tempo Soundlevel
## 1      1 1.2889 4.4667 1.7111  3.1333         4     1     4         4
## 2      2 1.0667 5.4444 1.4889  4.4889         5     1     4         1
## 3      3 2.0222 1.4889 3.7778  2.7111         2     2     5         1
## 4      4 2.2889 4.1111 1.2667  1.4889         1     1     5         4
## 5      5 1.4000 1.4667 5.0444  3.8444         3     2     1         3
## 6      6 1.6444 2.5111 3.0444  3.1556         2     2     3         2
##      Articulation Timbre Melody
## 1              2      1      4
## 2              1      1      2
## 3              1      1      1
## 4              4      3      2
## 5              2      1      1
## 6              4      1      2
```

1. Correlation between perceived emotions

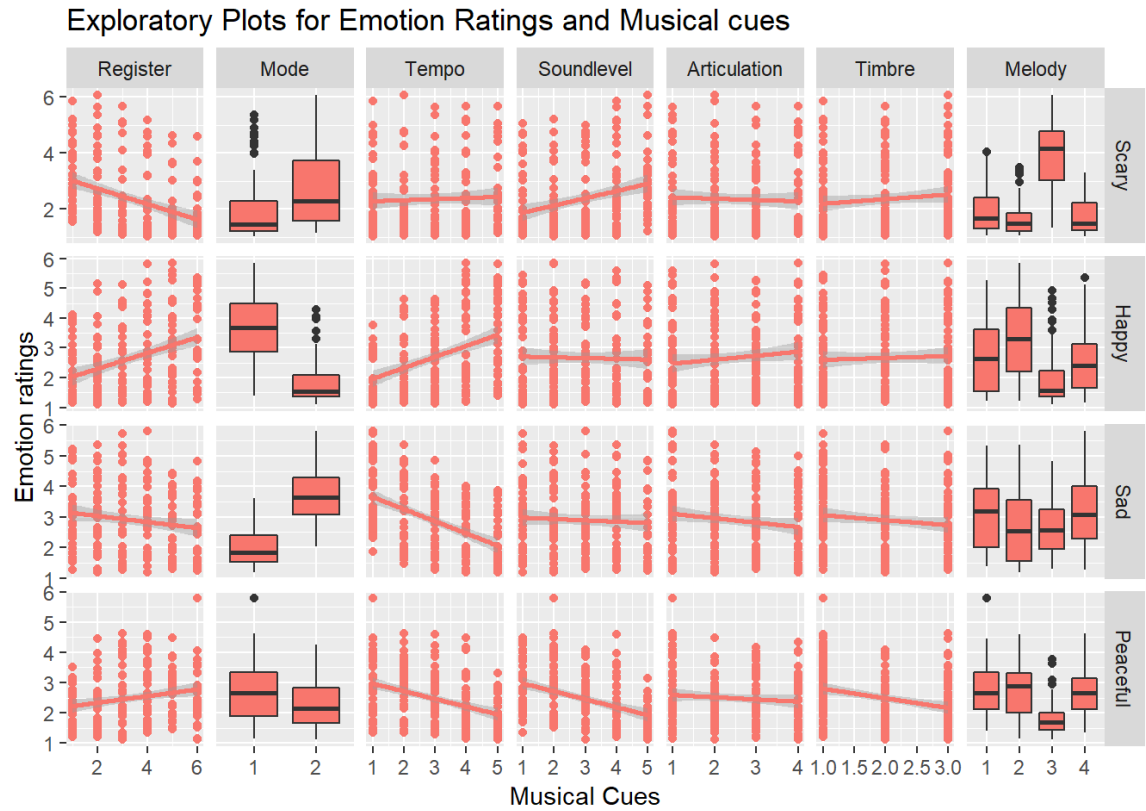


```
##
## Register      1.00000000  0.093377580 -0.002404327 -0.0092507085
## Tempo        0.093377580  1.000000000  0.059779874 -0.0306797074
## Soundlevel   -0.002404327  0.059779874  1.000000000 -0.0183955321
## Articulation -0.009250709 -0.030679707 -0.018395532  1.0000000000
## Timbre       0.024395174  0.003459393  0.029925274 -0.0006674221
##
## Timbre
## Register     0.0243951744
## Tempo       0.0034593928
## Soundlevel  0.0299252739
## Articulation -0.0006674221
## Timbre      1.0000000000
```

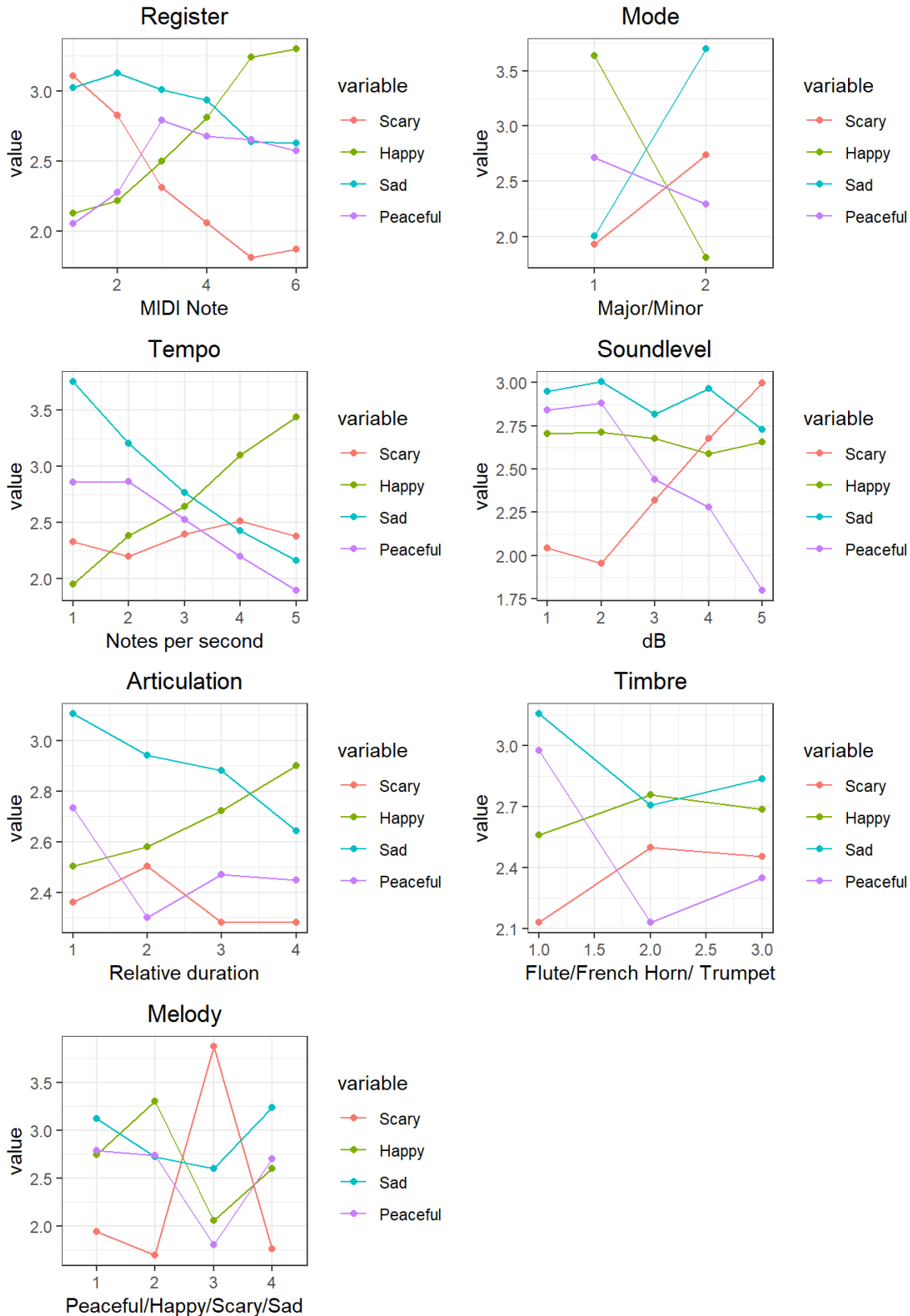
Happy is weakly correlated with peaceful (corr=0.212) and correlated negatively with Sad (corr=-0.789) as well as Scary (corr=-0.567). Sad is weakly correlated with peaceful (corr=0.163) and shows almost no correlation with scary. Peaceful is significantly correlated with scary in a negative fashion.

A quick check of correlations between predictor variables shows no multicollinearity problem.

2. Emotion ratings and Cue Levels



3. Individual Cue's Contribution for Mean Emotion Rating



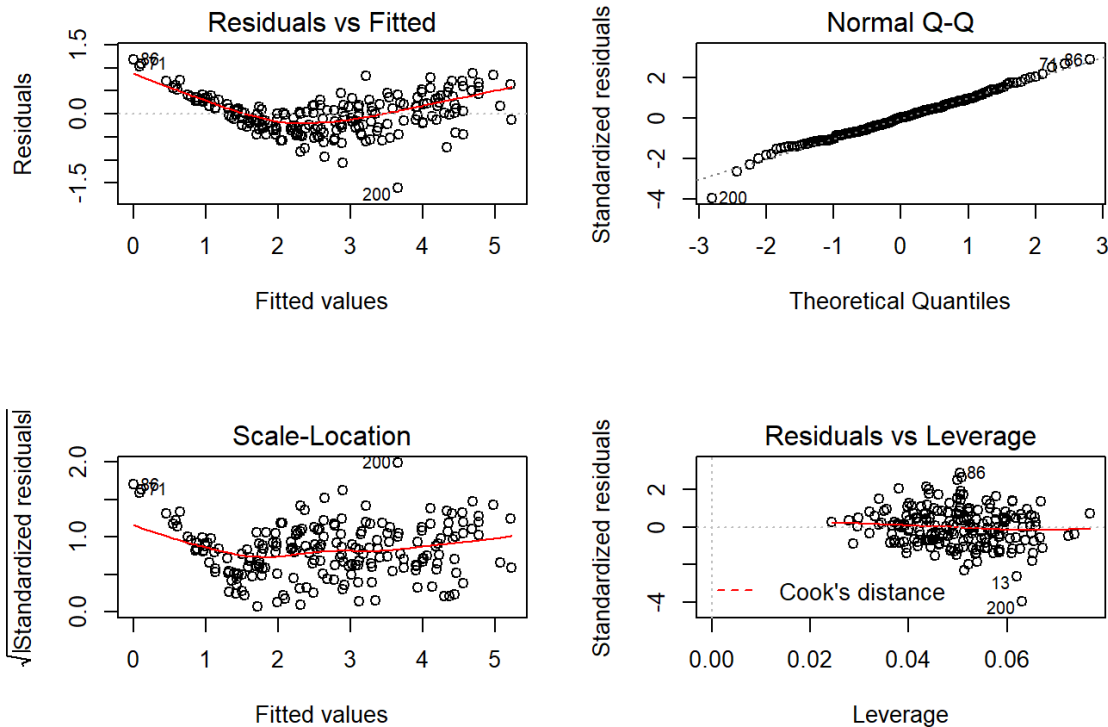
As a final step for exploratory analysis, I calculate the mean rating for four emotions grouping by each of the predictor variable. Graphs below show clear patterns across the levels for four emotions. For instance, major mode corresponds to music with higher happiness rating and minor mode corresponds to higher sadness rating. Register impacts happy and scary emotions notably. Music with happier mood leans towards a higher register while music with higher scariness rating lean towards a lower register. We can find similar patterns appeared in plots for other variables as well: tempo, soundlevel, timbre and melody, although emotions and directions of cue levels are different.

Modelling

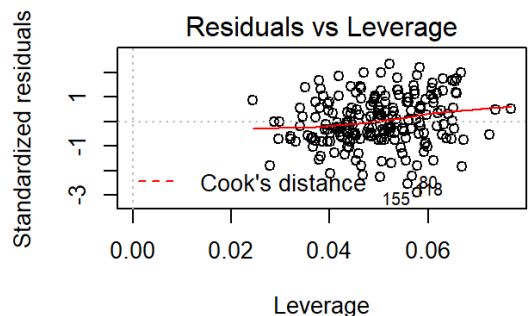
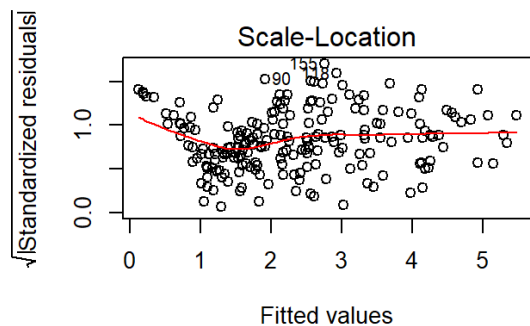
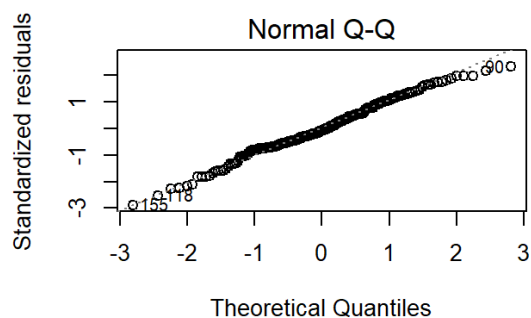
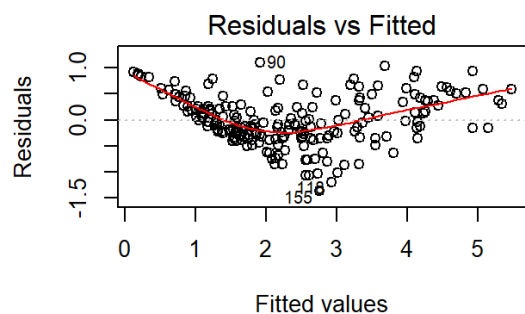
Stage 1 (No transformation): Run a simple linear regression on each of the four emotions

```
library(memisc)
music_emotion$Mode <- as.numeric(music_emotion$Mode)
music_emotion$Melody <- as.numeric(music_emotion$Melody)

reghappy <- lm(Happy ~ Register + as.factor(Mode) + Tempo + Soundlevel + Articulation + Timbre + as.factor(Melody), data = music_emotion)
par(mfrow = c(2,2))
plot(reghappy)
```

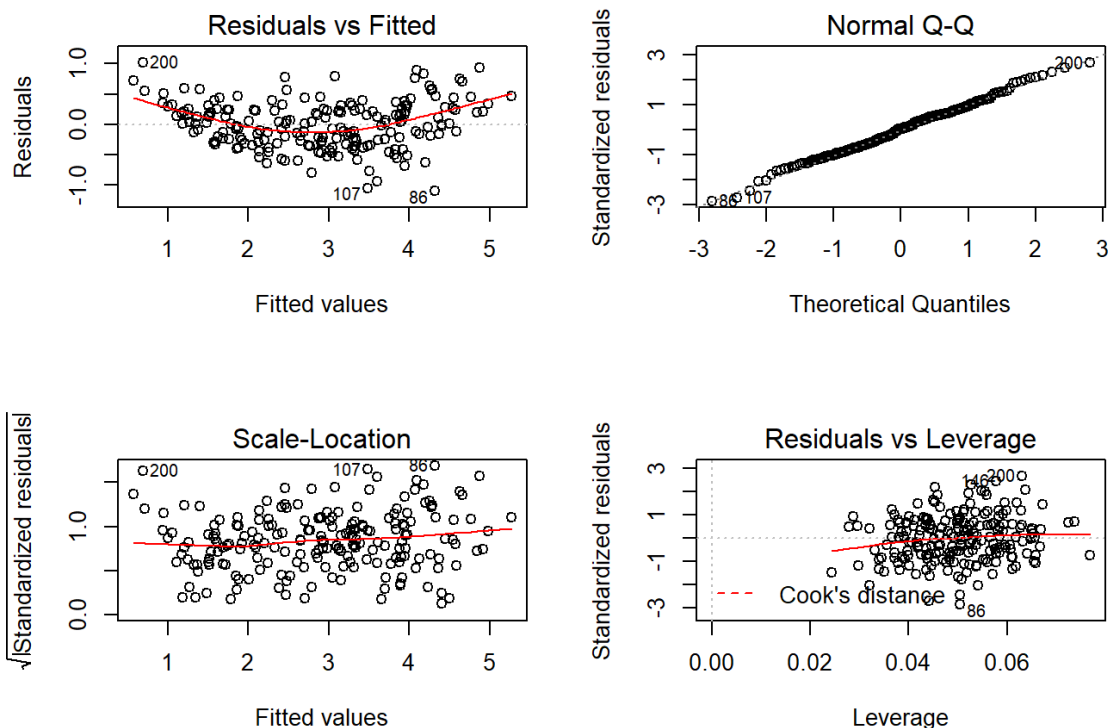


```
regscary <- lm(Scary ~ Register + as.factor(Mode) + Tempo + Soundlevel + Articulation + Timbre + as.factor(Melody), data = music_emotion)
par(mfrow = c(2,2))
plot(regscary)
```



```
regsad <- lm(Sad ~ Register + as.factor(Mode) + Tempo + Soundlevel + Articulation + Timbre + as.factor(Melody), data = music_emotion)
par(mfrow = c(2,2))
plot(regsad)

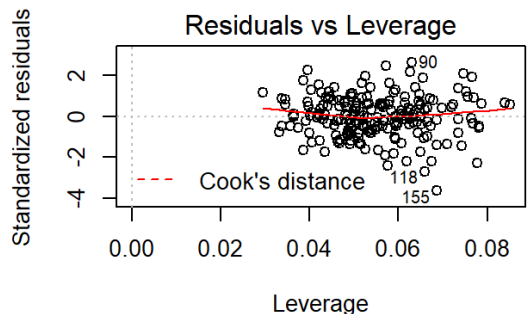
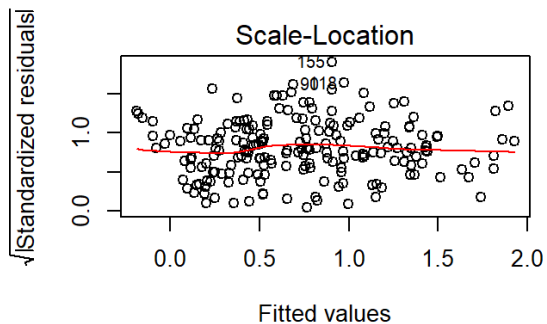
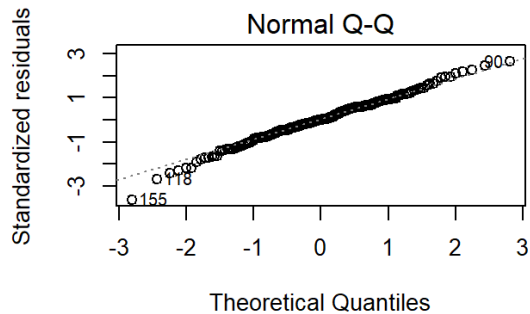
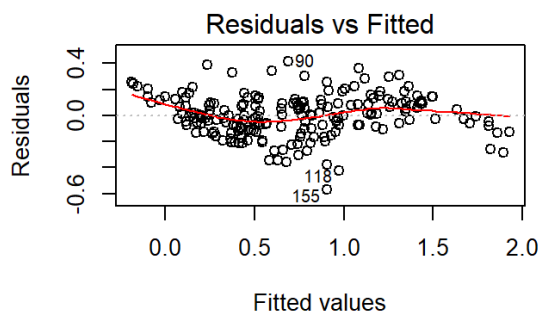
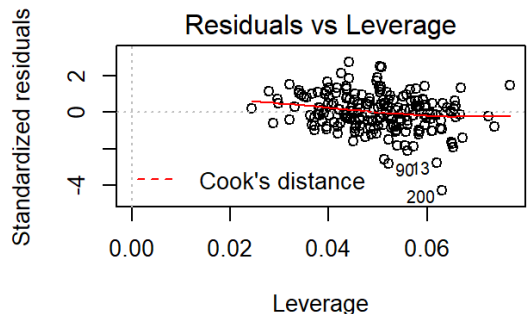
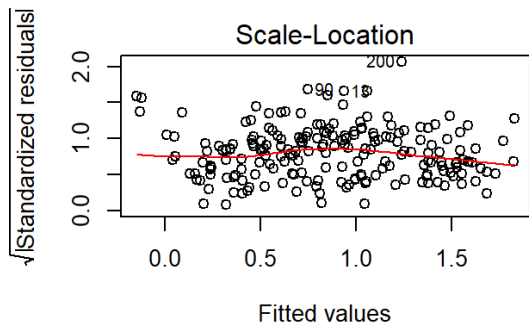
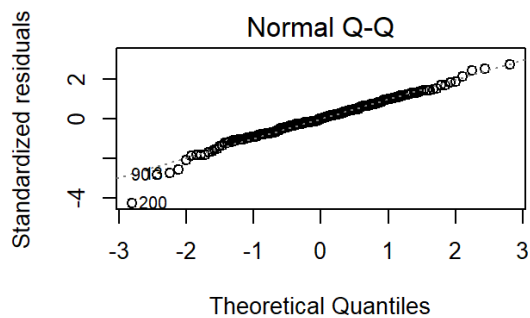
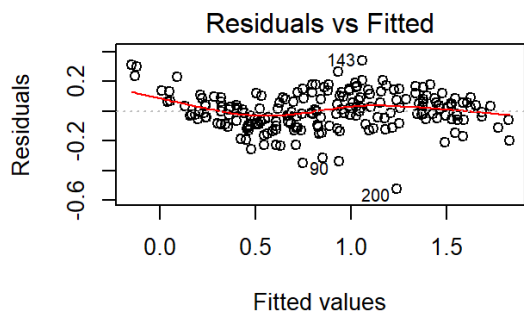
regpeaceful <- lm(Peaceful ~ Register + as.factor(Mode) + Tempo + Soundlevel + Articulation + Timbre + as.factor(Melody), data = music_emotion)
par(mfrow = c(2,2))
plot(regsad)
```



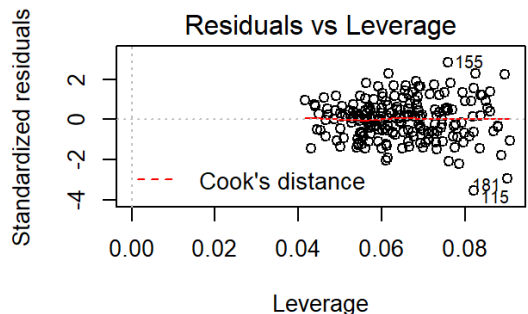
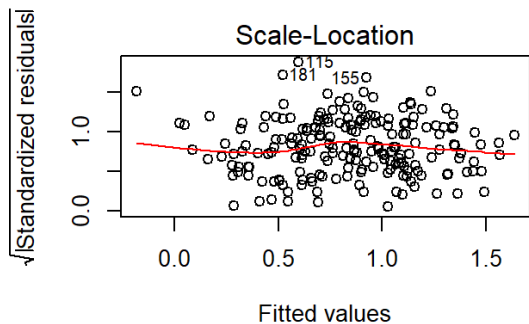
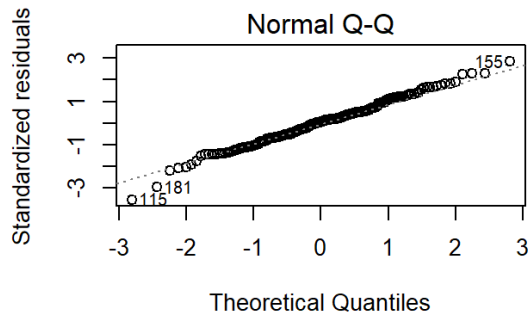
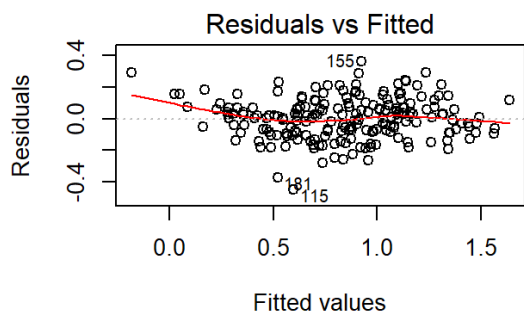
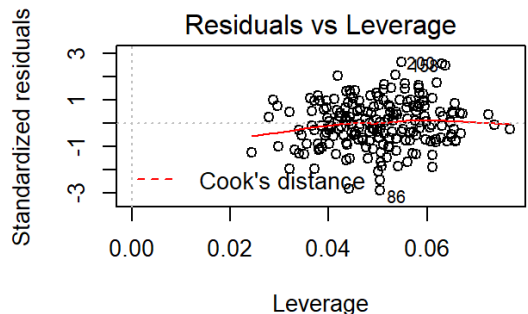
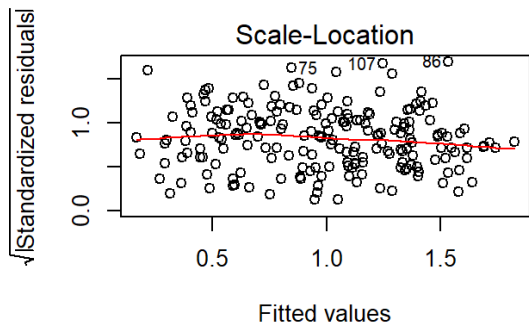
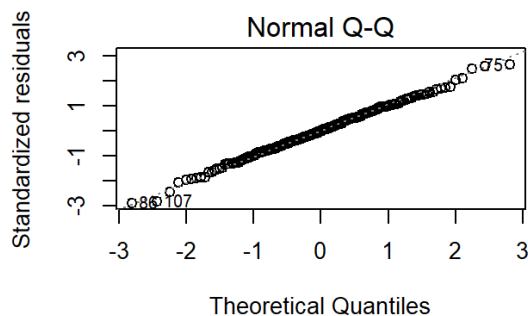
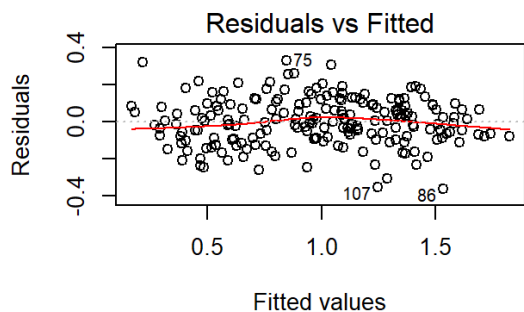
T. Eerola seemed to make a fatal mistake by not checking residual plots for all the regression models he made. The residual versus fitted value plot above shows a quadratic trend (LOESS function is not a horizontal line) and also a fanning problem for outcome variable happy. The same problem exists in all four regressions for emotions. Even though later on Eerola tried quadratic and even cubic numerical variables such as Register and Tempo, the fact that non-constant variance assumption wasn't met was the biggest problem hindering him to make correct conclusions. His linearity of cue level argument was challenged.

Stage 2 (Final model): log encoding for all emotional expressions, quadratic encoding for “Register, Tempo and Timbre” in some of the emotional expressions

```
## Analysis of Variance Table
##
## Model 1: loghappy ~ Register + as.factor(Mode) + Tempo + Soundlevel +
##   Articulation + Timbre + as.factor(Melody)
## Model 2: loghappy ~ Register + as.factor(Mode) + Tempo + Articulation +
##   Timbre + as.factor(Melody)
##   Res.Df  RSS Df Sum of Sq    F Pr(>F)
## 1     190 3.0498
## 2     191 3.0509 -1 -0.0010734 0.0669 0.7962
```

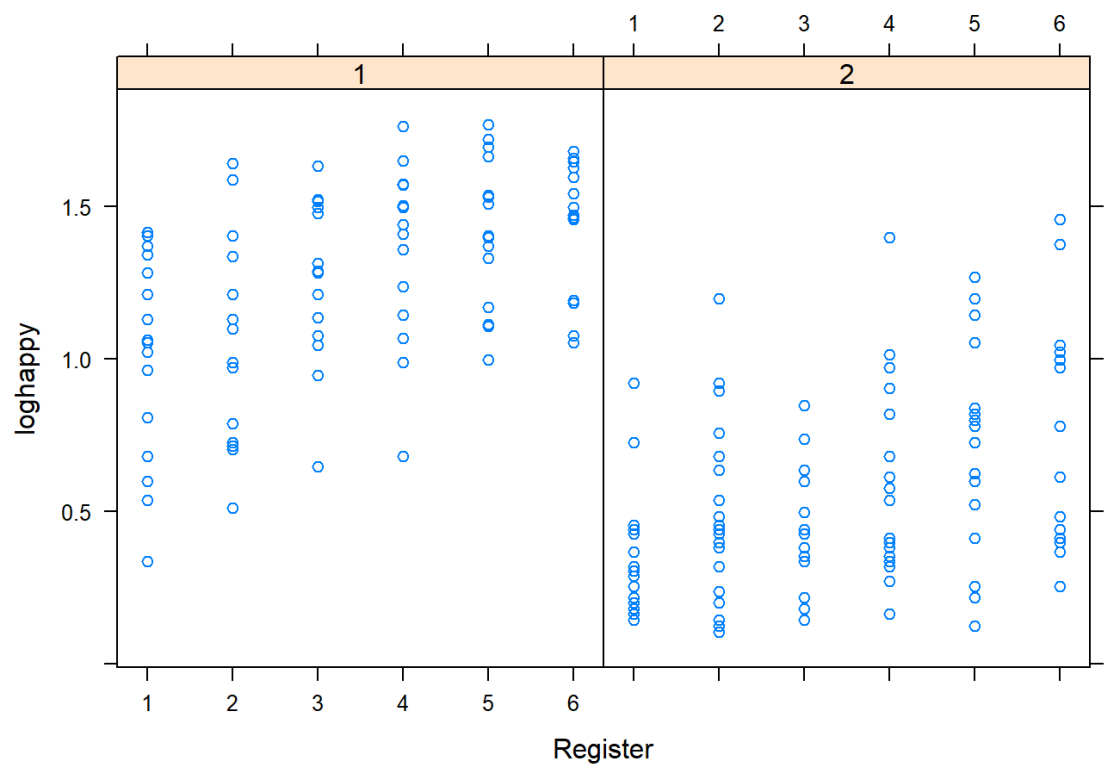
```
## Analysis of Variance Table
##
## Model 1: logscary ~ Register + Register2 + as.factor(Mode) + Tempo + Soundlevel +
##   Articulation + Timbre + as.factor(Melody)
## Model 2: logscary ~ Register + Register2 + as.factor(Mode) + Tempo + Soundlevel +
##   Timbre + as.factor(Melody)
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1      189 5.0096
## 2      190 5.0299 -1 -0.020321 0.7667 0.3824
```



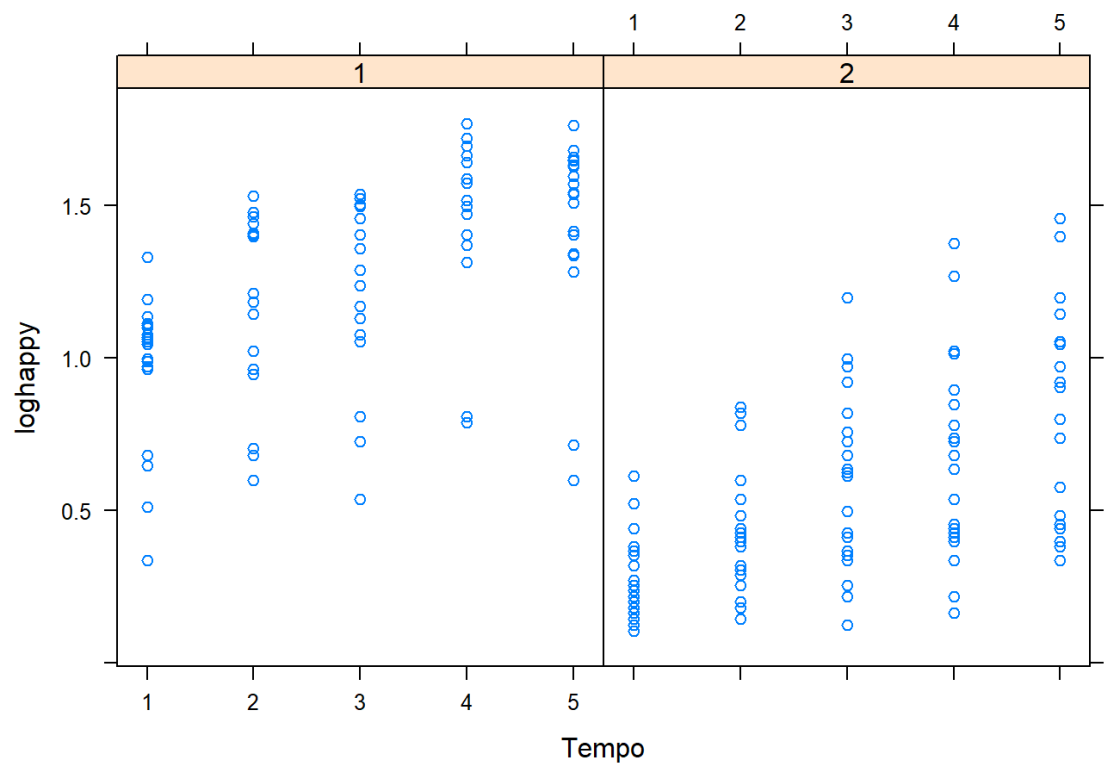
```
##
## Calls:
## Model 1: lm(formula = loghappy ~ Register + as.factor(Mode) + Tempo +
##   Articulation + Timbre + as.factor(Melody), data = music_emotion)
## Model 2: lm(formula = logscary ~ Register + Register2 + as.factor(Mode) +
##   Tempo + Soundlevel + Timbre + as.factor(Melody), data = music_emotion)
## Model 3: lm(formula = logsad ~ Register + as.factor(Mode) + Tempo + Soundlevel +
##   Articulation + Timbre + as.factor(Melody), data = music_emotion)
## Model 4: lm(formula = logpeaceful ~ Register + Register2 + as.factor(Mode) +
##   Tempo + Tempo2 + Soundlevel + Articulation + Timbre + Timbre2 +
##   as.factor(Melody), data = music_emotion)
##
## =====
##               Model 1      Model 2      Model 3      Model 4
##               -----
##               loghappy      logscary      logsad      logpeaceful
## -----
## (Intercept)      0.515***      0.432***      1.496***      2.116***
##                  (0.044)      (0.069)      (0.048)      (0.094)
## Register          0.087***      -0.316***      -0.022***      0.221***
##                  (0.005)      (0.033)      (0.005)      (0.027)
## as.factor(Mode): 2/1 -0.692***      0.339***      0.620***      -0.167***
##                  (0.018)      (0.023)      (0.018)      (0.019)
## Tempo             0.116***      0.036***      -0.135***      0.047
##                  (0.006)      (0.008)      (0.006)      (0.034)
## Articulation      0.072***                  -0.068***      -0.029**
##                  (0.008)                  (0.008)      (0.009)
## Timbre            -0.022*       0.100***      -0.034**      -0.929***
##                  (0.011)      (0.014)      (0.011)      (0.082)
## as.factor(Melody): 2/1 0.185***      -0.101**      -0.163***      -0.045
##                  (0.025)      (0.032)      (0.025)      (0.026)
## as.factor(Melody): 3/1 -0.293***      0.748***      -0.178***      -0.477***
##                  (0.025)      (0.032)      (0.025)      (0.026)
## as.factor(Melody): 4/1 -0.053*      -0.052       0.030       -0.077**
##                  (0.026)      (0.034)      (0.027)      (0.028)
## Register2                  0.028***                  -0.024***
##                  (0.005)                  (0.004)
## Soundlevel                  0.111***      -0.020**      -0.104***
##                  (0.008)      (0.007)      (0.007)
## Tempo2                  -0.026***
##                  (0.006)
## Timbre2                  0.198***
##                  (0.020)
## -----
## sigma             0.126       0.163       0.129       0.133
## R-squared          0.933       0.901       0.906       0.879
## F                 334.108      191.074      204.327      113.534
## p                  0.000       0.000       0.000       0.000
## N                  200        200        200        200
## =====
```

#Try an interaction effect

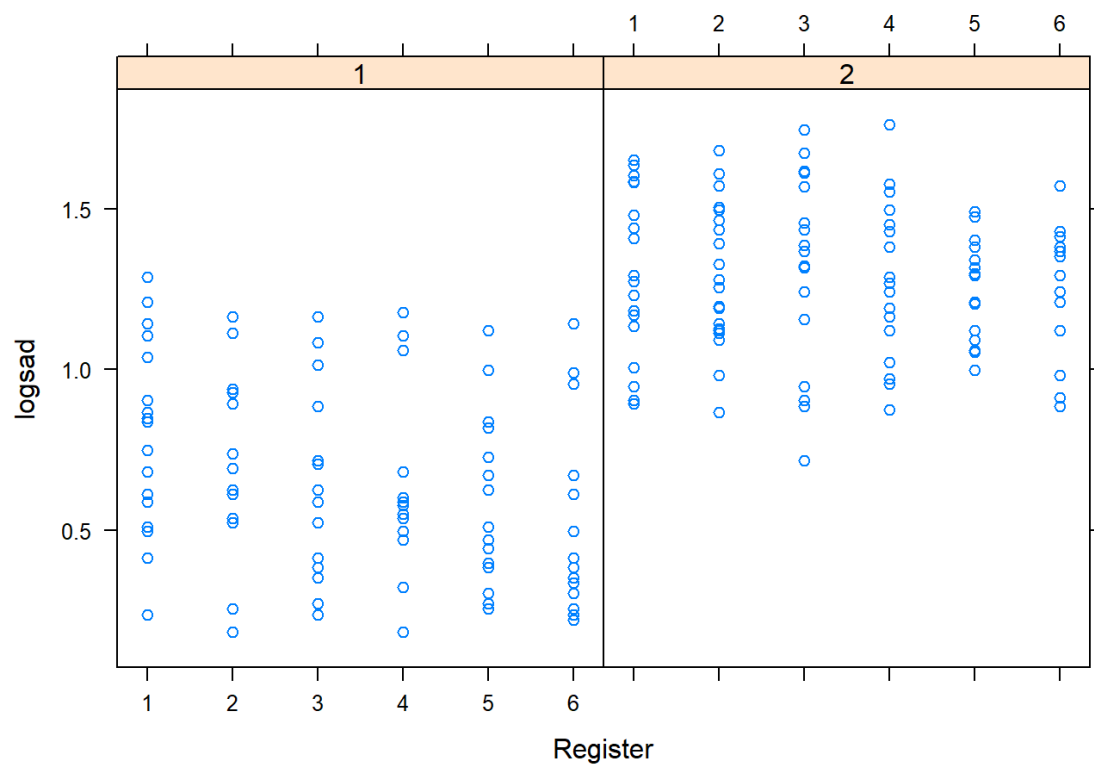
```
xyplot(loghappy ~ Register | as.factor(Mode), data = music_emotion)
```



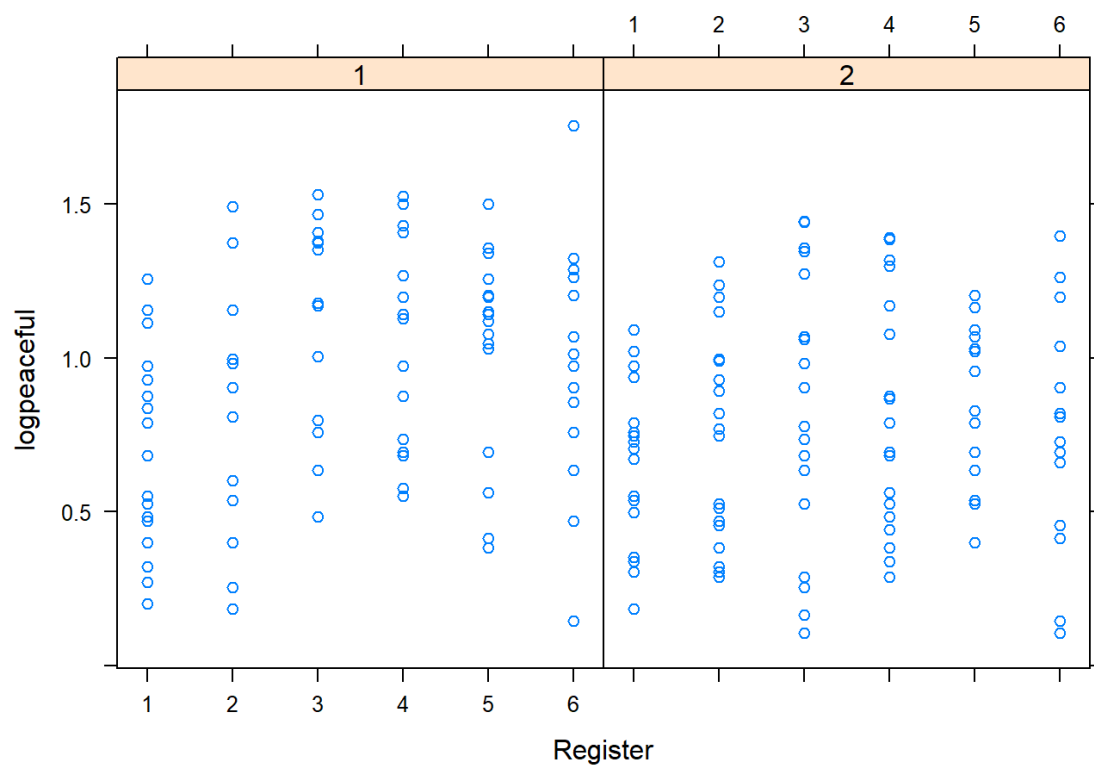
```
xyplot(loghappy ~ Tempo | as.factor(Mode), data = music_emotion)
```



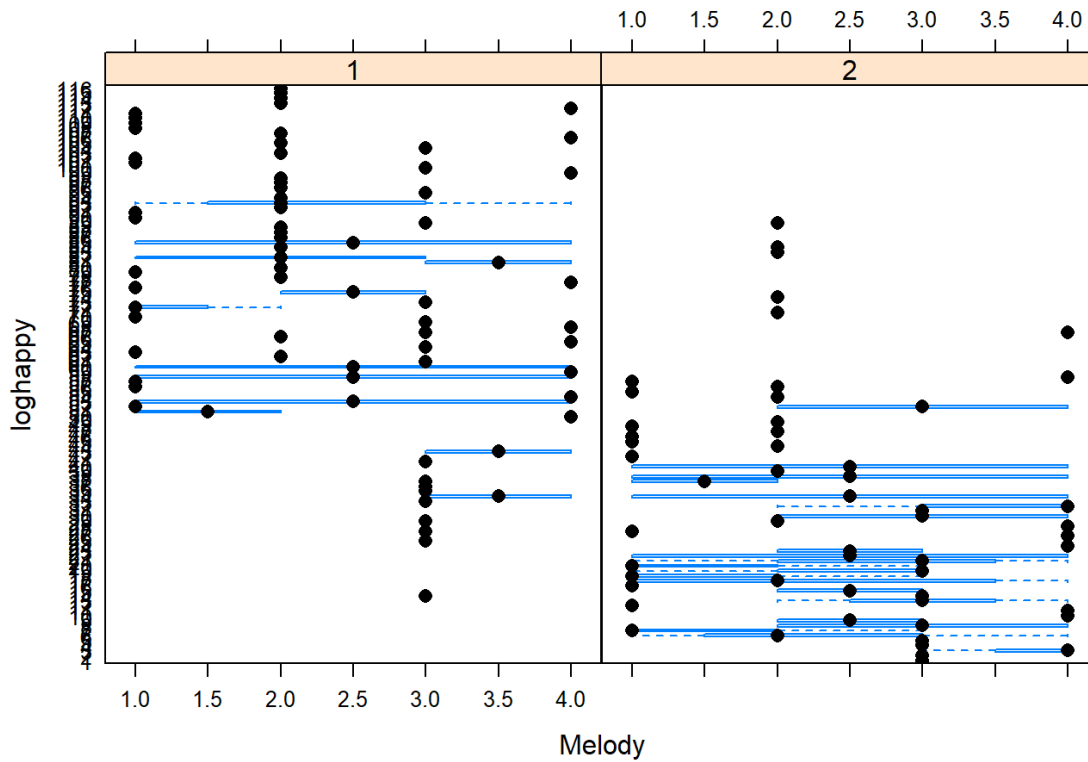
```
xyplot(logsad ~ Register | as.factor(Mode), data = music_emotion)
```



```
xyplot(logpeaceful ~ Register | as.factor(Mode), data = music_emotion)
```



```
bwplot(loghappy ~ Melody | as.factor(Mode), data = music_emotion)
```



#It doesn't seem to be an interaction effect in this case.

confidence intervals

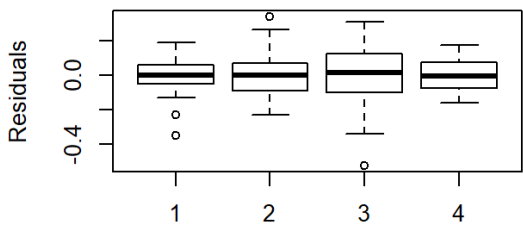
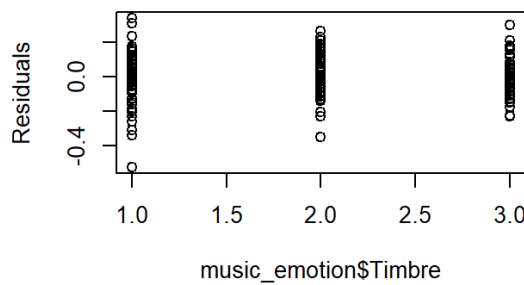
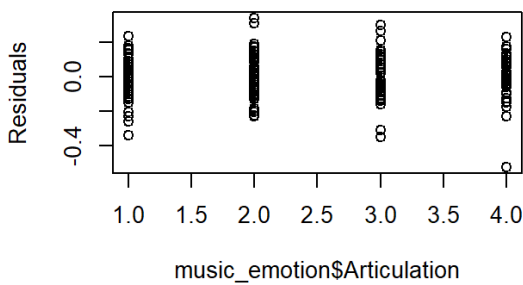
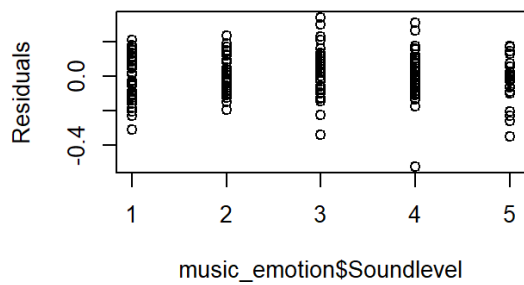
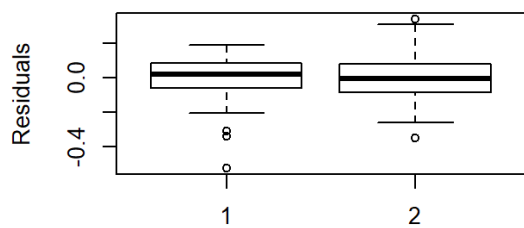
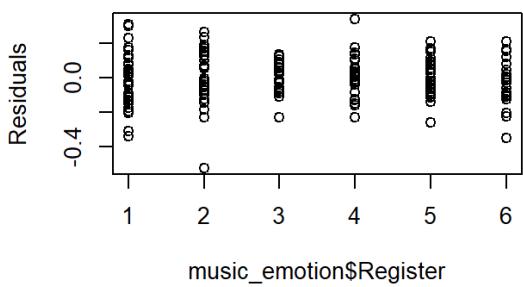
```
##                2.5 %    97.5 %
## (Intercept)    1.5357022 1.8233578
## Register       1.0800063 1.1028669
## as.factor(Mode)2 0.4829966 0.5184650
## Tempo          1.1089715 1.1368124
## Articulation    1.0571711 1.0917728
## Timbre          0.9571794 0.9995365
## as.factor(Melody)2 1.1458888 1.2637372
## as.factor(Melody)3 0.7107961 0.7829655
## as.factor(Melody)4 0.9006423 0.9986618
```

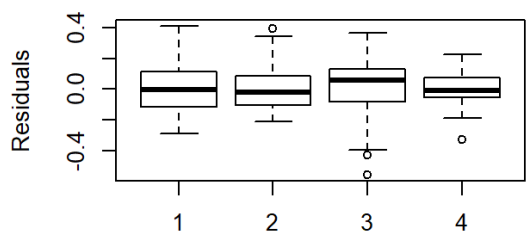
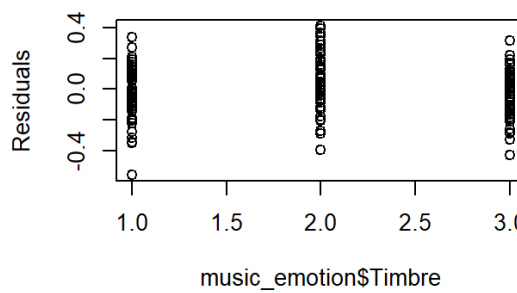
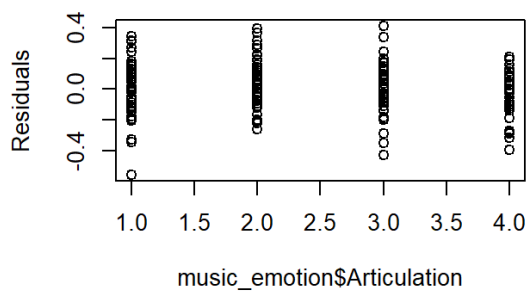
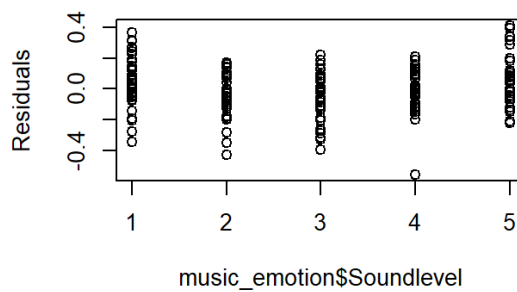
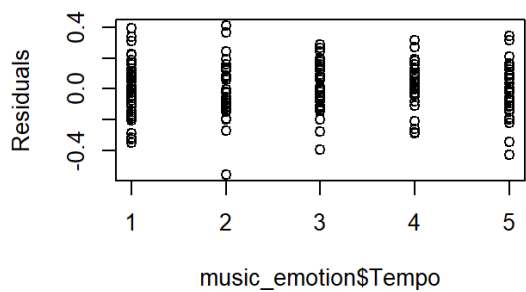
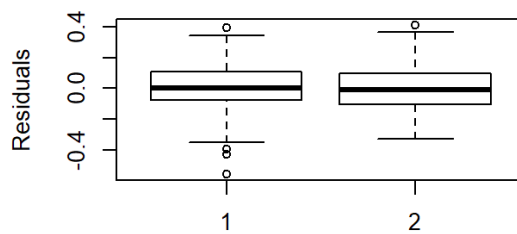
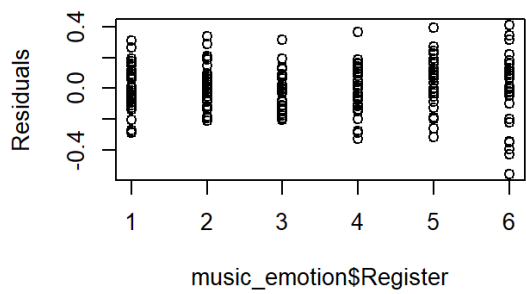
```
##                2.5 %    97.5 %
## (Intercept)    1.3442286 1.7633359
## Register       0.6831637 0.7777013
## Register2      1.0185636 1.0375690
## as.factor(Mode)2 1.3407422 1.4689893
## Tempo          1.0200244 1.0531697
## Soundlevel     1.0992886 1.1363300
## Timbre         1.0749952 1.1366820
## as.factor(Melody)2 0.8486357 0.9626266
## as.factor(Melody)3 1.9850912 2.2484920
## as.factor(Melody)4 0.8885537 1.0151422
```

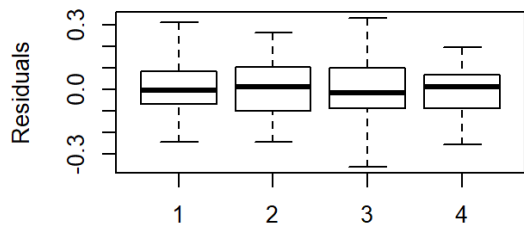
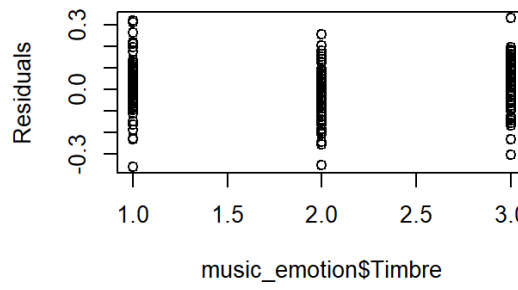
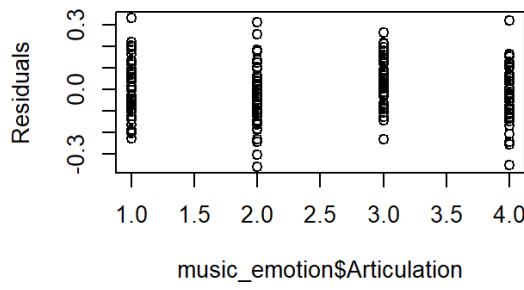
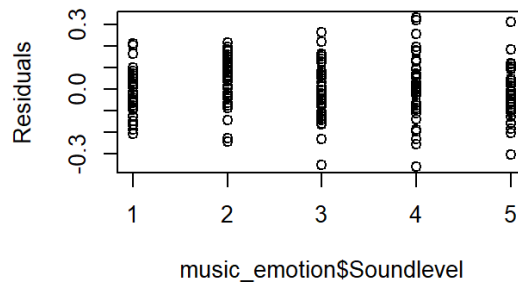
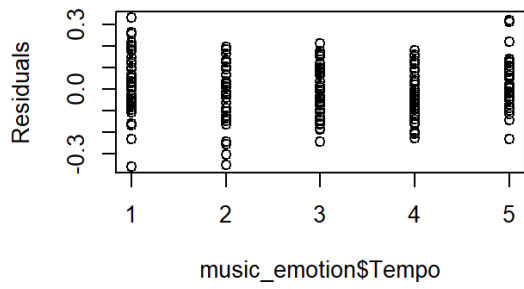
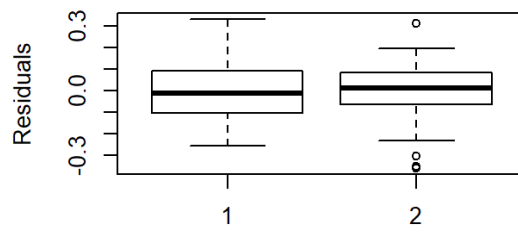
##	2.5 %	97.5 %
## (Intercept)	4.0605326	4.9030432
## Register	0.9678218	0.9887364
## as.factor(Mode)2	1.7930138	1.9276578
## Tempo	0.8630299	0.8851923
## Soundlevel	0.9675246	0.9932824
## Articulation	0.9193418	0.9500704
## Timbre	0.9451074	0.9878366
## as.factor(Melody)2	0.8081091	0.8930550
## as.factor(Melody)3	0.7969197	0.8796306
## as.factor(Melody)4	0.9777373	1.0864726

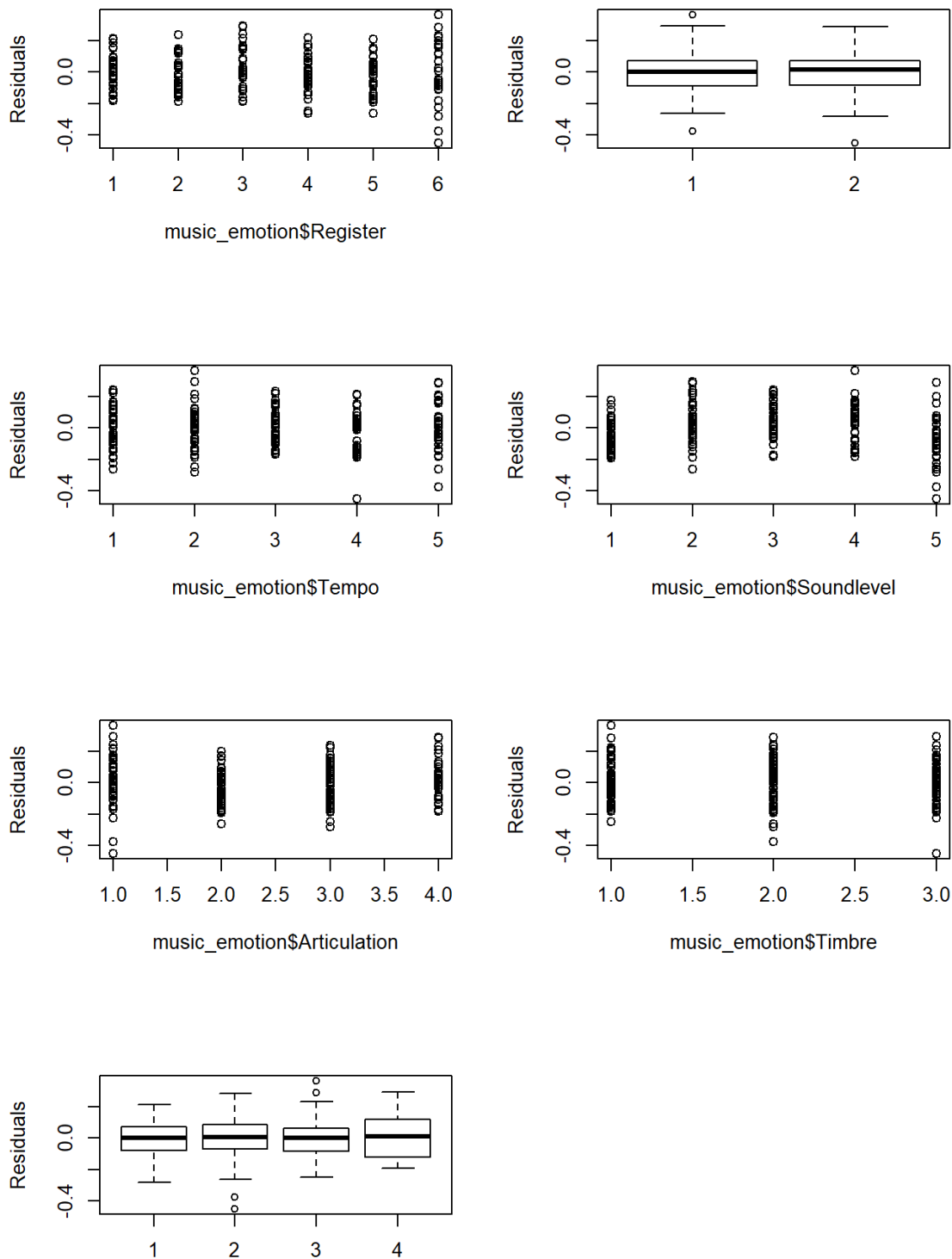
##	2.5 %	97.5 %
## (Intercept)	6.8954461	9.9812058
## Register	1.1827054	1.3149872
## Register2	0.9690989	0.9838733
## as.factor(Mode)2	0.8150578	0.8782743
## Tempo	0.9807572	1.1210275
## Tempo2	0.9639624	0.9855046
## Soundlevel	0.8894701	0.9138964
## Articulation	0.9553499	0.9883683
## Timbre	0.3361362	0.4637247
## Timbre2	1.1712084	1.2683968
## as.factor(Melody)2	0.9076390	1.0061387
## as.factor(Melody)3	0.5897172	0.6528835
## as.factor(Melody)4	0.8769380	0.9779976

Residual Plots checking









In order to satisfy the non-constant variance assumption, I logarithmized all the outcome variables. To correct quadratic trends in residual plots for emotional expression $\log(\text{scary})$ and $\log(\text{peaceful})$, I also added quadratic terms. Specifically, I added register^2 in scary model and register^2 , timbre^2 and tempo^2 in peaceful model. Careful check of interaction plots shows no interesting interaction between musical cues. I'm satisfied with these models since most predictors have really significant p-values and r-squares for these set of models are between 0.879 to 0.933.

However, in the second stage of Eerola's modelling, he essentially did regressions with quadratic/cubic encodings of all numerical variables (register, tempo, soundlevel, articulation, and timbre), which is obviously overfitting the data. Quadratic terms for a particular variable should be properly added if the residual plot of that variable shows any curve trend.

Conclusions

Happiness

An increase in **register** of 1 level is associated with a multiplicative change of $e^{0.087} \approx 1.09$ (95% CI: 1.08, 1.10) in happiness rating. Changing **mode** from major to minor is associated with a multiplicative change of $e^{-0.69} \approx 0.50$ (95% CI: 0.48, 0.51) in happiness rating.

An increase in **tempo** of 1 level is associated with a multiplicative change of $e^{0.12} \approx 1.13$ (95% CI: 1.11, 1.14) in happiness rating.

An increase in duration of a note (**articulation**) of 1 level is associated with a multiplicative change of $e^{0.07} \approx 1.07$ (95% CI: 1.06, 1.09) in happiness rating.

An increase in **brightness of timbre** of 1 level is associated with a multiplicative change of $e^{-0.02} \approx 0.98$ (95% CI: 0.96, 1.00) in happiness rating.

Changing **melody** from **sad to happy** is associated with a multiplicative change of $e^{0.19} \approx 1.2$ (95% CI: 1.15, 1.26) in happiness rating.

Changing melody from **sad to scary** is associated with a multiplicative change of $e^{-0.29} \approx 0.75$ (95% CI: 0.71, 0.78) in happiness rating.

Changing melody from **sad to peaceful** is associated with a multiplicative change of $e^{-0.05} \approx 0.95$ (95% CI: 0.90, 1.00) in happiness rating.

The fact that F statistic is pretty big shows that soundlevel has little association with happy mood.

Mode seems extremely important for happy music. And if we are to rank importance for other predictors, then the order would be melody, tempo, register, articulation and timbre.

Scariness

Changing **mode** from major to minor is associated with a multiplicative change of $e^{0.34} \approx 1.4$ (95% CI: 1.34, 1.46) in scariness rating.

An increase in **tempo** of 1 level is associated with a multiplicative change of $e^{0.036} \approx 1.04$ (95% CI: 1.02, 1.05) in scariness rating.

An increase in **soundlevel** of 5 dB is associated with a multiplicative change of $e^{0.11} \approx 1.12$ (95% CI: 1.10, 1.14) in scariness rating.

An increase in **brightness of timbre** of 1 level is associated with a multiplicative change of $e^{0.1} \approx 1.1$ (95% CI: 1.07, 1.13) in scariness rating.

Changing **melody** from **sad to happy** is associated with a multiplicative change of $e^{-0.1} \approx 0.90$ (95% CI: 0.85, 0.96) in scariness rating.

Changing melody from **sad to scary** is associated with a multiplicative change of $e^{0.7} \approx 2$ (95% CI: 1.99, 2.25) in scariness rating.

Changing melody from **sad to peaceful** is associated with a multiplicative change of $e^{-0.05} \approx 0.95$ (95% CI: 0.89, 1.02) in scariness rating.

The fact that F statistic is 0.382 shows that articulation has little association with log(scary).

Quadratic effect of register is shown in the plot.

Mode and melody are extremely important for fearful music. And if we are to rank importance for other predictors, then the order would be soundlevel, timbre and tempo.

Sadness

An increase in **register** of 1 level is associated with a multiplicative change of $e^{-0.022} \approx 0.98$ (95% CI: 0.97, 0.99) in sadness rating.

Changing **mode** from major to minor is associated with a multiplicative change of $e^{0.62} \approx 1.86$ (95% CI: 1.79, 1.93) in sadness rating.

An increase in **tempo** of 1 level is associated with a multiplicative change of $e^{-0.13} \approx 0.87$ (95% CI: 0.86, 0.89) in sadness rating. An

increase in **soundlevel** of 5 dB is associated with a multiplicative change of $e^{-0.02} \approx 0.98$ (95% CI: 0.97, 0.99) in scariness rating.

An increase in duration of a note (**articulation**) of 1 level is associated with a multiplicative change of $e^{-0.07} \approx 0.93$ (95% CI: 0.92, 0.95) in sadness rating.

An increase in **brightness of timbre** of 1 level is associated with a multiplicative change of $e^{-0.03} \approx 0.9$ (95% CI: 0.95, 0.99) in sadness rating.

Changing **melody** from **sad to happy** is associated with a multiplicative change of $e^{-0.16} \approx 0.85$ (95% CI: 0.81, 0.89) in sadness rating.

Changing melody from **sad to scary** is associated with a multiplicative change of $e^{-0.17} \approx 0.84$ (95% CI: 0.80, 0.88) in sadness rating.

Changing melody from **sad to peaceful** is associated with a multiplicative change of $e^{0.03} \approx 1.03$ (95% CI: 0.98, 1.09) in sadness rating.

Mode is extremely important in sad music. And if we are to rank importance for other predictors, then the order would be melody, tempo, timbre, articulation, soundlevel and register.

Peacefulness

Changing **mode** from major to minor is associated with a multiplicative change of $e^{-0.17} \approx 0.84$ (95% CI: 0.82,0.88) in peacefulness rating.

An increase in **soundlevel** of 5 dB is associated with a multiplicative change of $e^{0.11} \approx 1.12$ (95% CI: 0.90,0.91) in peacefulness rating.

An increase in duration of a note (**articulation**) of 1 level is associated with a multiplicative change of $e^{-0.03} \approx 0.97$ (95% CI: 0.96,0.99) in sadness rating.

Changing **melody** from **sad to happy** is associated with a multiplicative change of $e^{-0.05} \approx 0.95$ (95% CI: 0.91,1.00) in peacefulness rating.

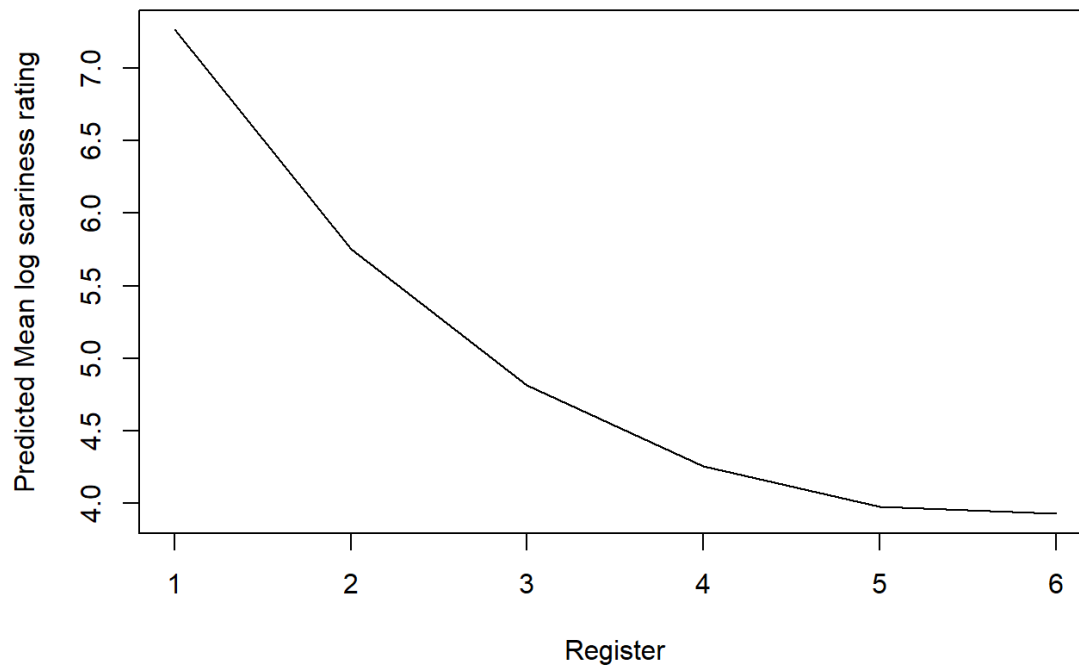
Changing melody from **sad to scary** is associated with a multiplicative change of $e^{-0.5} \approx 0.61$ (95% CI: 0.59,0.65) in peacefulness rating.

Changing melody from **sad to peaceful** is associated with a multiplicative change of $e^{-0.07} \approx 0.93$ (95% CI: 0.88,0.98) in peacefulness rating.

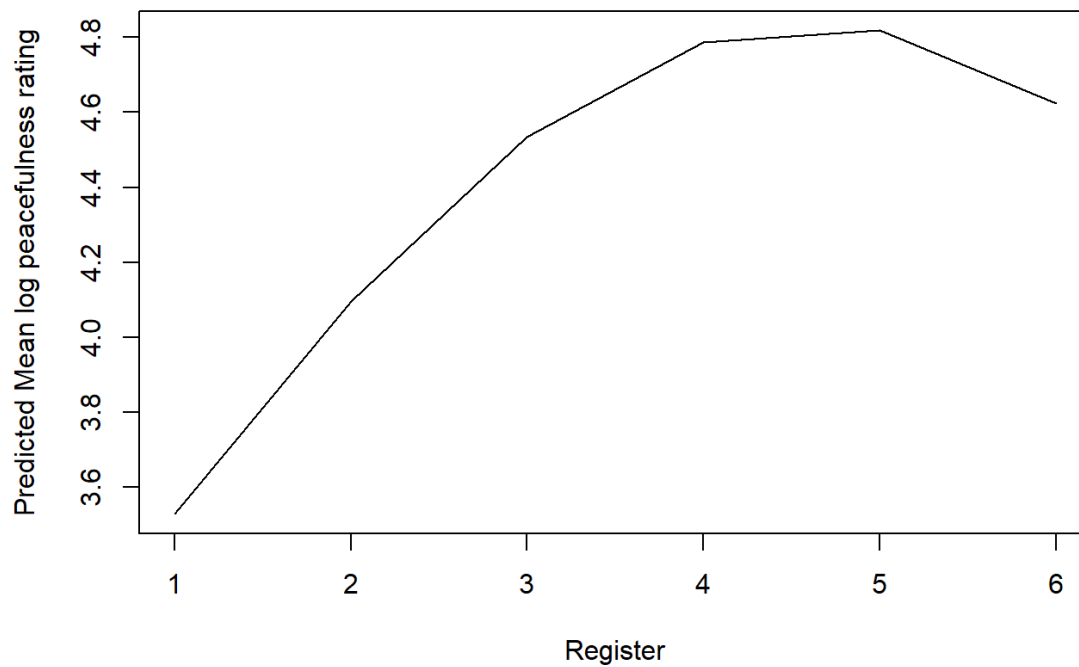
Quadratic effect of register, tempo and timbre are shown in the plot.

Mode and melody are extremely important in peaceful music. And if we are to rank importance for other predictors, than the order would be soundlevel and articulation.

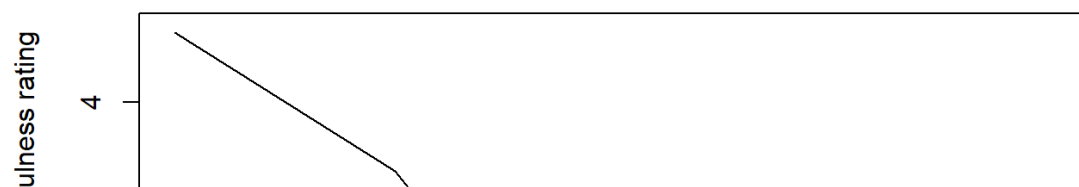
Expected Change in Log Scariness Rating with Register

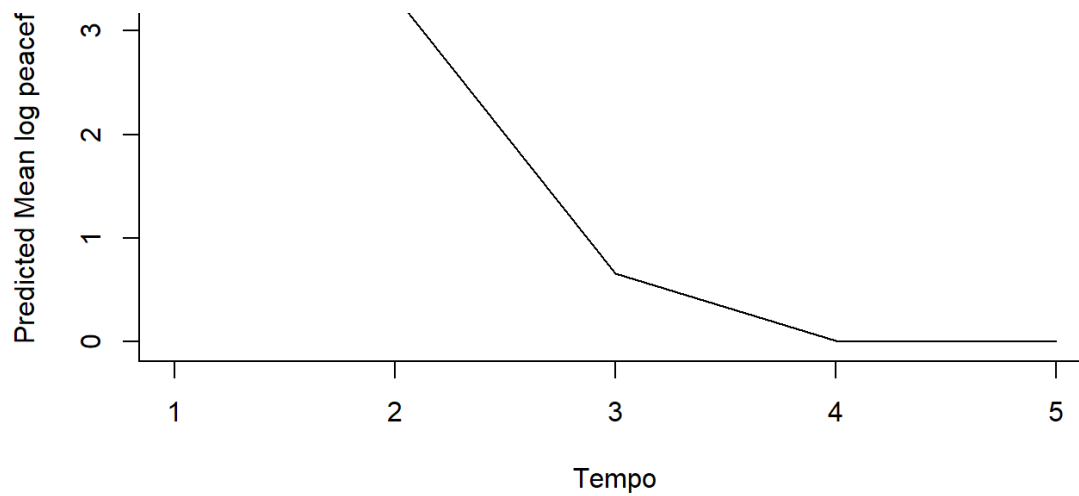


Expected Change in Log Peacefulness Rating with Register



Expected Change in Log Peacefulness Rating with Tempo





Stage 3: Multivariate linear regression between happiness and sadness

Model

Multivariate multiple regression is the method of modeling multiple response variables with a single set of predictor variables. I would like to model both happiness and sadness level as a function of register, mode, tempo, soundlevel and etc. Multivariate Multiple Regression is very similar with running separate regressions with each outcome variables. The only difference lies in hypothesis tests for regression parameters and confidence interval for prediction.

Let's run a multivariate linear regression for loghappy and logsad.


```
## Response loghappy :
##
## Call:
## lm(formula = loghappy ~ Register + as.factor(Mode) + Tempo +
##     Soundlevel + Articulation + Timbre + as.factor(Melody), data = music_emotion)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.52336 -0.08338 -0.00099  0.08192  0.33956
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.519312   0.046940   11.063 < 2e-16 ***
## Register       0.087430   0.005323   16.426 < 2e-16 ***
## as.factor(Mode)2 -0.692096   0.018027  -38.392 < 2e-16 ***
## Tempo          0.115931   0.006313   18.365 < 2e-16 ***
## Soundlevel     -0.001692   0.006541   -0.259  0.7962
## Articulation    0.071659   0.008185    8.754 1.11e-15 ***
## Timbre         -0.022023   0.011009   -2.001  0.0469 *
## as.factor(Melody)2 0.184959   0.024884    7.433 3.53e-12 ***
## as.factor(Melody)3 -0.293211   0.024585  -11.927 < 2e-16 ***
## as.factor(Melody)4 -0.052918   0.026253   -2.016  0.0452 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1267 on 190 degrees of freedom
## Multiple R-squared:  0.9333, Adjusted R-squared:  0.9302
## F-statistic: 295.5 on 9 and 190 DF, p-value: < 2.2e-16
##
##
## Response logsad :
##
## Call:
## lm(formula = logsad ~ Register + as.factor(Mode) + Tempo + Soundlevel +
##     Articulation + Timbre + as.factor(Melody), data = music_emotion)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36218 -0.08688 -0.00008  0.09019  0.33188
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.495585   0.047792   31.294 < 2e-16 ***
## Register      -0.022017   0.005419   -4.063 7.09e-05 ***
## as.factor(Mode)2 0.620102   0.018354   33.786 < 2e-16 ***
## Tempo         -0.134628   0.006427  -20.947 < 2e-16 ***
## Soundlevel     -0.019877   0.006660   -2.985 0.00321 **
## Articulation   -0.067658   0.008334   -8.118 5.81e-14 ***
## Timbre        -0.034347   0.011209   -3.064 0.00250 **
## as.factor(Melody)2 -0.163083   0.025336   -6.437 9.70e-10 ***
## as.factor(Melody)3 -0.177627   0.025031   -7.096 2.47e-11 ***
## as.factor(Melody)4 0.030211   0.026730    1.130 0.25980
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.129 on 190 degrees of freedom
## Multiple R-squared:  0.9064, Adjusted R-squared:  0.9019
## F-statistic: 204.3 on 9 and 190 DF, p-value: < 2.2e-16
```

Summary results are the same as running separate regressions (refer to my appendix). But coefficients from two models covary when we check the variance-covariance matrix and we should consider covariance when determining if a predictor is jointly contributing to both models. The effect of soundlevel doesn't seem to be significant in loghappy. Thus, we need to formally test whether to include

soundlevel in multivariate model using Anova command.

```

##                                loghappy:(Intercept) loghappy:Register
## loghappy:(Intercept)                2.203374e-03   -8.500205e-05
## loghappy:Register                    -8.500205e-05    2.833223e-05
## loghappy:as.factor(Mode)2            -1.764187e-04    1.986004e-06
## loghappy:Tempo                      -1.054214e-04   -3.145448e-06
## loghappy:Soundlevel                  -1.132775e-04    2.396401e-07
## loghappy:Articulation                 -1.613502e-04    1.984787e-07
## loghappy:Timbre                      -2.429768e-04   -1.512846e-06
## loghappy:as.factor(Melody)2          -2.926228e-04    7.666704e-07
## loghappy:as.factor(Melody)3          -3.434020e-04   -3.994593e-06
## loghappy:as.factor(Melody)4          -3.161659e-04    2.171968e-07
## logsad:(Intercept)                  -1.035636e-03    3.995291e-05
## logsad:Register                     3.995291e-05   -1.331679e-05
## logsad:as.factor(Mode)2              8.292082e-05   -9.334672e-07
## logsad:Tempo                        4.955046e-05    1.478432e-06
## logsad:Soundlevel                   5.324299e-05   -1.126363e-07
## logsad:Articulation                  7.583830e-05   -9.328951e-08
## logsad:Timbre                       1.142047e-04    7.110721e-07
## logsad:as.factor(Melody)2            1.375394e-04   -3.603526e-07
## logsad:as.factor(Melody)3            1.614068e-04    1.877550e-06
## logsad:as.factor(Melody)4            1.486052e-04   -1.020874e-07
##                                loghappy:as.factor(Mode)2 loghappy:Tempo
## loghappy:(Intercept)                -1.764187e-04   -1.054214e-04
## loghappy:Register                    1.986004e-06   -3.145448e-06
## loghappy:as.factor(Mode)2            3.249700e-04    3.834419e-06
## loghappy:Tempo                      3.834419e-06    3.984912e-05
## loghappy:Soundlevel                  -5.530754e-06   -2.529432e-06
## loghappy:Articulation                 -7.236869e-06    1.396023e-06
## loghappy:Timbre                      9.736035e-06    3.757424e-07
## loghappy:as.factor(Melody)2          -4.463955e-09   -1.216454e-06
## loghappy:as.factor(Melody)3          -1.634537e-06    3.664746e-06
## loghappy:as.factor(Melody)4          4.437543e-06    3.262985e-06
## logsad:(Intercept)                  8.292082e-05    4.955046e-05
## logsad:Register                     -9.334672e-07    1.478432e-06
## logsad:as.factor(Mode)2              -1.527433e-04   -1.802265e-06
## logsad:Tempo                        -1.802265e-06   -1.873000e-05
## logsad:Soundlevel                   2.599581e-06    1.188891e-06
## logsad:Articulation                  3.401494e-06   -6.561628e-07
## logsad:Timbre                       -4.576159e-06   -1.766075e-07
## logsad:as.factor(Melody)2            2.098161e-09    5.717613e-07
## logsad:as.factor(Melody)3            7.682696e-07   -1.722514e-06
## logsad:as.factor(Melody)4            -2.085746e-06   -1.533677e-06
##                                loghappy:Soundlevel loghappy:Articulation
## loghappy:(Intercept)                -1.132775e-04   -1.613502e-04
## loghappy:Register                    2.396401e-07    1.984787e-07
## loghappy:as.factor(Mode)2            -5.530754e-06   -7.236869e-06
## loghappy:Tempo                      -2.529432e-06    1.396023e-06
## loghappy:Soundlevel                  4.278882e-05    1.028249e-06
## loghappy:Articulation                 1.028249e-06    6.700204e-05
## loghappy:Timbre                      -2.296289e-06   -3.077658e-07
## loghappy:as.factor(Melody)2          4.247443e-06   -8.731251e-06
## loghappy:as.factor(Melody)3          4.881282e-06   -3.024989e-06
## loghappy:as.factor(Melody)4          -1.892760e-06   -1.105186e-05
## logsad:(Intercept)                  5.324299e-05    7.583830e-05
## logsad:Register                     -1.126363e-07   -9.328951e-08
## logsad:as.factor(Mode)2              2.599581e-06    3.401494e-06
## logsad:Tempo                        1.188891e-06   -6.561628e-07
## logsad:Soundlevel                   -2.011172e-05   -4.833004e-07
## logsad:Articulation                 -4.833004e-07   -3.149249e-05
## logsad:Timbre                       1.079308e-06    1.446569e-07
## logsad:as.factor(Melody)2            -1.996395e-06    4.103888e-06
## logsad:as.factor(Melody)3            -2.294314e-06    1.421814e-06

```

## logsad:as.factor(Melody)4	8.896403e-07	5.194625e-06
##	loghappy:Timbre	loghappy:as.factor(Melody)2
## loghappy:(Intercept)	-2.429768e-04	-2.926228e-04
## loghappy:Register	-1.512846e-06	7.666704e-07
## loghappy:as.factor(Mode)2	9.736035e-06	-4.463955e-09
## loghappy:Tempo	3.757424e-07	-1.216454e-06
## loghappy:Soundlevel	-2.296289e-06	4.247443e-06
## loghappy:Articulation	-3.077658e-07	-8.731251e-06
## loghappy:Timbre	1.211948e-04	-5.158093e-07
## loghappy:as.factor(Melody)2	-5.158093e-07	6.192224e-04
## loghappy:as.factor(Melody)3	1.793783e-05	3.035517e-04
## loghappy:as.factor(Melody)4	1.566978e-05	3.040359e-04
## logsad:(Intercept)	1.142047e-04	1.375394e-04
## logsad:Register	7.110721e-07	-3.603526e-07
## logsad:as.factor(Mode)2	-4.576159e-06	2.098161e-09
## logsad:Tempo	-1.766075e-07	5.717613e-07
## logsad:Soundlevel	1.079308e-06	-1.996395e-06
## logsad:Articulation	1.446569e-07	4.103888e-06
## logsad:Timbre	-5.696431e-05	2.424421e-07
## logsad:as.factor(Melody)2	2.424421e-07	-2.910487e-04
## logsad:as.factor(Melody)3	-8.431191e-06	-1.426762e-04
## logsad:as.factor(Melody)4	-7.365152e-06	-1.429038e-04
##	loghappy:as.factor(Melody)3	
## loghappy:(Intercept)	-3.434020e-04	
## loghappy:Register	-3.994593e-06	
## loghappy:as.factor(Mode)2	-1.634537e-06	
## loghappy:Tempo	3.664746e-06	
## loghappy:Soundlevel	4.881282e-06	
## loghappy:Articulation	-3.024989e-06	
## loghappy:Timbre	1.793783e-05	
## loghappy:as.factor(Melody)2	3.035517e-04	
## loghappy:as.factor(Melody)3	6.044076e-04	
## loghappy:as.factor(Melody)4	3.057315e-04	
## logsad:(Intercept)	1.614068e-04	
## logsad:Register	1.877550e-06	
## logsad:as.factor(Mode)2	7.682696e-07	
## logsad:Tempo	-1.722514e-06	
## logsad:Soundlevel	-2.294314e-06	
## logsad:Articulation	1.421814e-06	
## logsad:Timbre	-8.431191e-06	
## logsad:as.factor(Melody)2	-1.426762e-04	
## logsad:as.factor(Melody)3	-2.840854e-04	
## logsad:as.factor(Melody)4	-1.437008e-04	
##	loghappy:as.factor(Melody)4	logsad:(Intercept)
## loghappy:(Intercept)	-3.161659e-04	-1.035636e-03
## loghappy:Register	2.171968e-07	3.995291e-05
## loghappy:as.factor(Mode)2	4.437543e-06	8.292082e-05
## loghappy:Tempo	3.262985e-06	4.955046e-05
## loghappy:Soundlevel	-1.892760e-06	5.324299e-05
## loghappy:Articulation	-1.105186e-05	7.583830e-05
## loghappy:Timbre	1.566978e-05	1.142047e-04
## loghappy:as.factor(Melody)2	3.040359e-04	1.375394e-04
## loghappy:as.factor(Melody)3	3.057315e-04	1.614068e-04
## loghappy:as.factor(Melody)4	6.892389e-04	1.486052e-04
## logsad:(Intercept)	1.486052e-04	2.284070e-03
## logsad:Register	-1.020874e-07	-8.811518e-05
## logsad:as.factor(Mode)2	-2.085746e-06	-1.828799e-04
## logsad:Tempo	-1.533677e-06	-1.092824e-04
## logsad:Soundlevel	8.896403e-07	-1.174262e-04
## logsad:Articulation	5.194625e-06	-1.672596e-04
## logsad:Timbre	-7.365152e-06	-2.518757e-04
## logsad:as.factor(Melody)2	-1.429038e-04	-3.033399e-04
## logsad:as.factor(Melody)3	-1.437008e-04	-3.559788e-04

```

## logsad:as.factor(Melody)4          -3.239580e-04      -3.277452e-04
## logsad:Register logsad:as.factor(Mode)2
## loghappy:(Intercept)          3.995291e-05      8.292082e-05
## loghappy:Register          -1.331679e-05      -9.334672e-07
## loghappy:as.factor(Mode)2      -9.334672e-07      -1.527433e-04
## loghappy:Tempo          1.478432e-06      -1.802265e-06
## loghappy:Soundlevel          -1.126363e-07      2.599581e-06
## loghappy:Articulation          -9.328951e-08      3.401494e-06
## loghappy:Timbre          7.110721e-07      -4.576159e-06
## loghappy:as.factor(Melody)2      -3.603526e-07      2.098161e-09
## loghappy:as.factor(Melody)3      1.877550e-06      7.682696e-07
## loghappy:as.factor(Melody)4      -1.020874e-07      -2.085746e-06
## logsad:(Intercept)          -8.811518e-05      -1.828799e-04
## logsad:Register          2.936987e-05      2.058740e-06
## logsad:as.factor(Mode)2          2.058740e-06      3.368718e-04
## logsad:Tempo          -3.260647e-06      3.974852e-06
## logsad:Soundlevel          2.484167e-07      -5.733313e-06
## logsad:Articulation          2.057478e-07      -7.501914e-06
## logsad:Timbre          -1.568253e-06      1.009261e-05
## logsad:as.factor(Melody)2          7.947490e-07      -4.627443e-09
## logsad:as.factor(Melody)3          -4.140892e-06      -1.694400e-06
## logsad:as.factor(Melody)4          2.251514e-07      4.600064e-06
## logsad:Tempo logsad:Soundlevel
## loghappy:(Intercept)          4.955046e-05      5.324299e-05
## loghappy:Register          1.478432e-06      -1.126363e-07
## loghappy:as.factor(Mode)2      -1.802265e-06      2.599581e-06
## loghappy:Tempo          -1.873000e-05      1.188891e-06
## loghappy:Soundlevel          1.188891e-06      -2.011172e-05
## loghappy:Articulation          -6.561628e-07      -4.833004e-07
## loghappy:Timbre          -1.766075e-07      1.079308e-06
## loghappy:as.factor(Melody)2          5.717613e-07      -1.996395e-06
## loghappy:as.factor(Melody)3      -1.722514e-06      -2.294314e-06
## loghappy:as.factor(Melody)4      -1.533677e-06      8.896403e-07
## logsad:(Intercept)          -1.092824e-04      -1.174262e-04
## logsad:Register          -3.260647e-06      2.484167e-07
## logsad:as.factor(Mode)2          3.974852e-06      -5.733313e-06
## logsad:Tempo          4.130856e-05      -2.622070e-06
## logsad:Soundlevel          -2.622070e-06      4.435592e-05
## logsad:Articulation          1.447151e-06      1.065908e-06
## logsad:Timbre          3.895037e-07      -2.380389e-06
## logsad:as.factor(Melody)2          -1.261006e-06      4.403002e-06
## logsad:as.factor(Melody)3          3.798964e-06      5.060055e-06
## logsad:as.factor(Melody)4          3.382489e-06      -1.962081e-06
## logsad:Articulation logsad:Timbre
## loghappy:(Intercept)          7.583830e-05      1.142047e-04
## loghappy:Register          -9.328951e-08      7.110721e-07
## loghappy:as.factor(Mode)2          3.401494e-06      -4.576159e-06
## loghappy:Tempo          -6.561628e-07      -1.766075e-07
## loghappy:Soundlevel          -4.833004e-07      1.079308e-06
## loghappy:Articulation          -3.149249e-05      1.446569e-07
## loghappy:Timbre          1.446569e-07      -5.696431e-05
## loghappy:as.factor(Melody)2          4.103888e-06      2.424421e-07
## loghappy:as.factor(Melody)3          1.421814e-06      -8.431191e-06
## loghappy:as.factor(Melody)4          5.194625e-06      -7.365152e-06
## logsad:(Intercept)          -1.672596e-04      -2.518757e-04
## logsad:Register          2.057478e-07      -1.568253e-06
## logsad:as.factor(Mode)2          -7.501914e-06      1.009261e-05
## logsad:Tempo          1.447151e-06      3.895037e-07
## logsad:Soundlevel          1.065908e-06      -2.380389e-06
## logsad:Articulation          6.945593e-05      -3.190375e-07
## logsad:Timbre          -3.190375e-07      1.256334e-04
## logsad:as.factor(Melody)2          -9.051026e-06      -5.347004e-07
## logsad:as.factor(Melody)3          -3.135777e-06      1.859479e-05

```

```

## logsad:as.factor(Melody)4      -1.145662e-05  1.624367e-05
##                               logsad:as.factor(Melody)2
## loghappy:(Intercept)          1.375394e-04
## loghappy:Register             -3.603526e-07
## loghappy:as.factor(Mode)2      2.098161e-09
## loghappy:Tempo                5.717613e-07
## loghappy:Soundlevel           -1.996395e-06
## loghappy:Articulation          4.103888e-06
## loghappy:Timbre               2.424421e-07
## loghappy:as.factor(Melody)2   -2.910487e-04
## loghappy:as.factor(Melody)3   -1.426762e-04
## loghappy:as.factor(Melody)4   -1.429038e-04
## logsad:(Intercept)           -3.033399e-04
## logsad:Register              7.947490e-07
## logsad:as.factor(Mode)2       -4.627443e-09
## logsad:Tempo                 -1.261006e-06
## logsad:Soundlevel            4.403002e-06
## logsad:Articulation          -9.051026e-06
## logsad:Timbre                -5.347004e-07
## logsad:as.factor(Melody)2     6.419009e-04
## logsad:as.factor(Melody)3     3.146690e-04
## logsad:as.factor(Melody)4     3.151710e-04
##                               logsad:as.factor(Melody)3
## loghappy:(Intercept)          1.614068e-04
## loghappy:Register             1.877550e-06
## loghappy:as.factor(Mode)2      7.682696e-07
## loghappy:Tempo                -1.722514e-06
## loghappy:Soundlevel           -2.294314e-06
## loghappy:Articulation          1.421814e-06
## loghappy:Timbre              -8.431191e-06
## loghappy:as.factor(Melody)2   -1.426762e-04
## loghappy:as.factor(Melody)3   -2.840854e-04
## loghappy:as.factor(Melody)4   -1.437008e-04
## logsad:(Intercept)           -3.559788e-04
## logsad:Register              -4.140892e-06
## logsad:as.factor(Mode)2       -1.694400e-06
## logsad:Tempo                 3.798964e-06
## logsad:Soundlevel            5.060055e-06
## logsad:Articulation          -3.135777e-06
## logsad:Timbre                1.859479e-05
## logsad:as.factor(Melody)2     3.146690e-04
## logsad:as.factor(Melody)3     6.265435e-04
## logsad:as.factor(Melody)4     3.169286e-04
##                               logsad:as.factor(Melody)4
## loghappy:(Intercept)          1.486052e-04
## loghappy:Register             -1.020874e-07
## loghappy:as.factor(Mode)2      -2.085746e-06
## loghappy:Tempo                -1.533677e-06
## loghappy:Soundlevel           8.896403e-07
## loghappy:Articulation          5.194625e-06
## loghappy:Timbre              -7.365152e-06
## loghappy:as.factor(Melody)2   -1.429038e-04
## loghappy:as.factor(Melody)3   -1.437008e-04
## loghappy:as.factor(Melody)4   -3.239580e-04
## logsad:(Intercept)           -3.277452e-04
## logsad:Register              2.251514e-07
## logsad:as.factor(Mode)2       4.600064e-06
## logsad:Tempo                 3.382489e-06
## logsad:Soundlevel           -1.962081e-06
## logsad:Articulation          -1.145662e-05
## logsad:Timbre                1.624367e-05
## logsad:as.factor(Melody)2     3.151710e-04

```

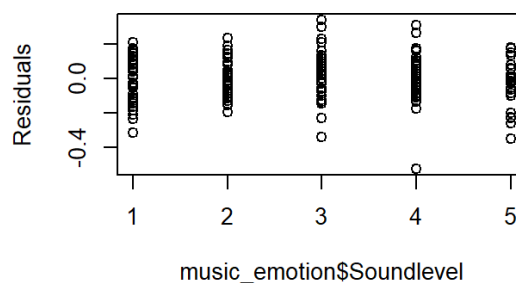
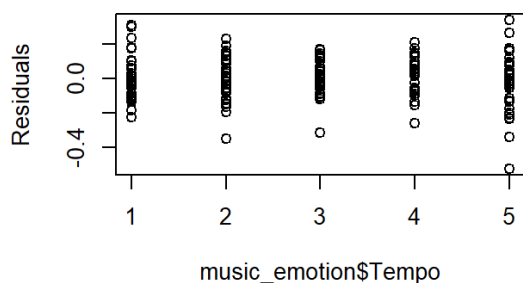
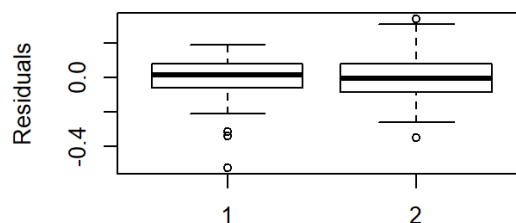
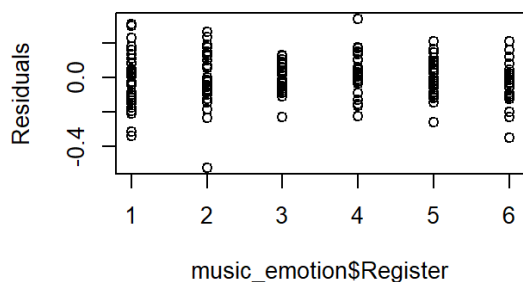
```
## logsad:as.factor(Melody)3          3.169286e-04
## logsad:as.factor(Melody)4          7.144817e-04
```

```
##
## Type II MANOVA Tests: Pillai test statistic
##
##          Df test stat approx F num Df den Df    Pr(>F)
## Register      1   0.60046   142.02      2   189 < 2.2e-16 ***
## as.factor(Mode) 1   0.90461   896.17      2   189 < 2.2e-16 ***
## Tempo          1   0.73787   266.01      2   189 < 2.2e-16 ***
## Soundlevel     1   0.06085     6.12      2   189  0.002651 **
## Articulation   1   0.33973    48.62      2   189 < 2.2e-16 ***
## Timbre         1   0.11303    12.04      2   189 1.195e-05 ***
## as.factor(Melody) 3   1.05912    71.29      6   380 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

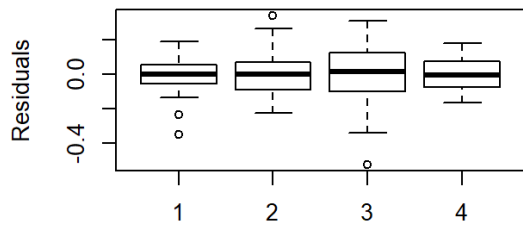
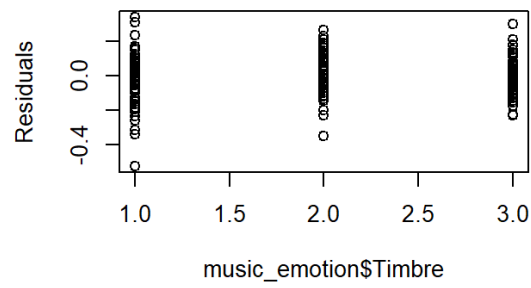
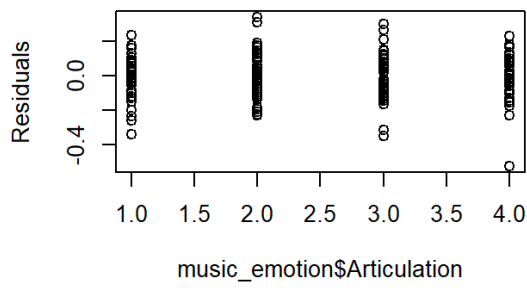
Soundlevel appears to be jointly significant for the multivariate models though in individual loghappy model it seems not as important as other predictors. P-value for soundlevel is 0.002651.

Residual Plots Checking (logsad was checked above)

```
par(mfrow = c(1,1))
par(mfrow = c(2,2))
plot(y = regloghappy$residuals, x = music_emotion$Register, ylab = "Residuals")
boxplot(regloghappy$residuals ~ music_emotion$Mode, ylab = "Residuals")
plot(y = regloghappy$residuals, x = music_emotion$Tempo, ylab = "Residuals")
plot(y = regloghappy$residuals, x = music_emotion$Soundlevel, ylab = "Residuals")
```



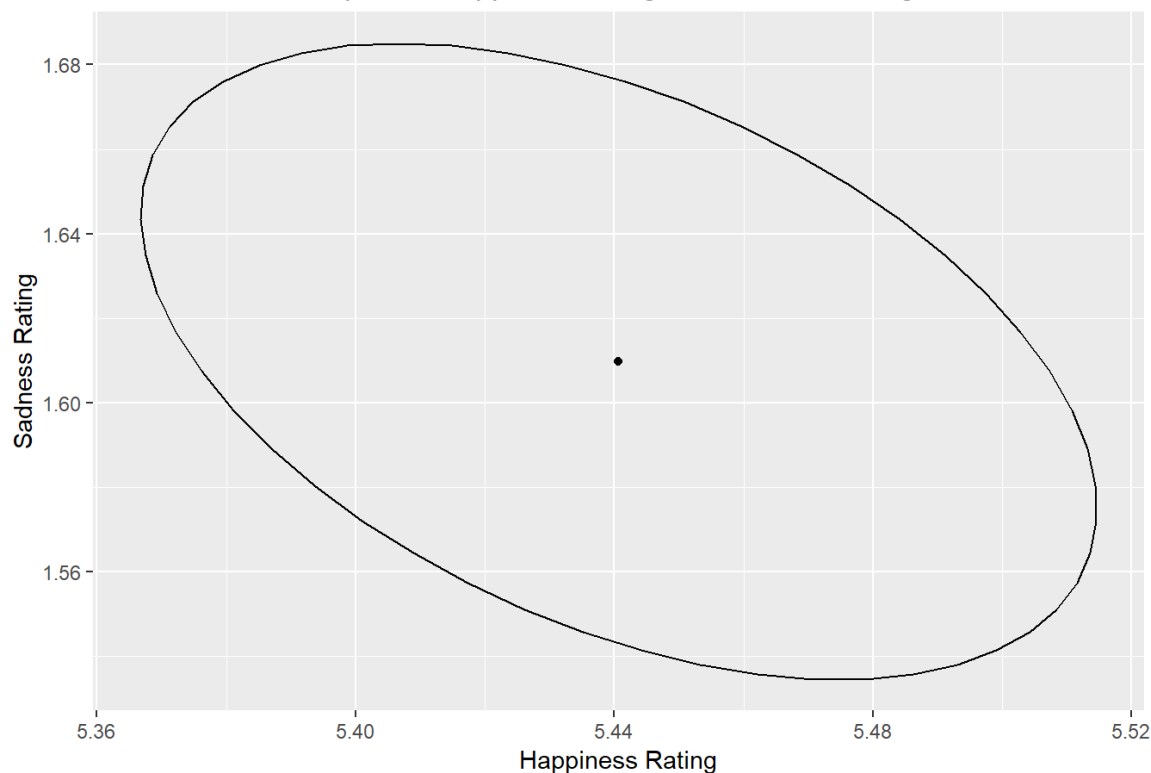
```
plot(y = regloghappy$residuals, x = music_emotion$Articulation, ylab = "Residuals")
plot(y = regloghappy$residuals, x = music_emotion$Timbre, ylab = "Residuals")
boxplot(regloghappy$residuals ~ music_emotion$Melody, ylab = "Residuals")
```



Prediction

Let's try to predict happiness and sadness rating of a piece of music with Register = 83 in MIDI pitch, Mode = major, Tempo = 4.4 NPS, Soundlevel = -5dB, Articulation = staccato, Timbre = horn and Melody = peaceful.

95% confidence ellipse for happiness rating and sadness rating



This is a 95% confidence ellipse for happiness and sadness rating, which is equivalent to confidence interval in 1d setting. Black dot in the center is our predicted value for happy and sad. We're 95% confident that the true values of happy and sad when Register = 6, Mode = 1, Tempo = 4, Soundlevel = 2, Articulation = 4, Timbre = 2 and Melody = 4 are within area of the ellipse. Happy and sad are negatively correlated. Predicting higher rating of happiness means predicting lower rating of sadness, and vice versa.

Limitation of original study by T. Eerola: * T. Eerola didn't checking residual plots for all the regression models he made. Residual plots of his final model (plain linear regression) has non-constant variance problem. I remodelled the same set of data by observing patterns in exploratory process and carefully checking residual plots. * Eerola overfit this data set in the second stage by quadratic/cubic encoding all numerical variables. * An optimized fractional factorial design might in its nature exclude interaction effect between variables.

Reference: Citation: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3726864/>
(<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3726864/>) <https://data.library.virginia.edu/getting-started-with-multivariate-multiple-regression/> (<https://data.library.virginia.edu/getting-started-with-multivariate-multiple-regression/>)
https://en.wikipedia.org/wiki/Factorial_experiment (https://en.wikipedia.org/wiki/Factorial_experiment)