Final Paper

Completed By Lois Wong

Initial setup: Installing and loading necessary packages and reading in data

Article https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0179145 (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0179145)

The data can be found at https://osf.io/z7xkw/ (https://osf.io/z7xkw/). There are ten columns/variables in the data frame: - 'subject' is a unique identifier for every participant - 'excerpt' is a condition label for each excerpt's valence (positive, negative, or neutral) - 'stimID is a unique identifier for every stimulus (poetic or music excerpt) - 'enjoy' is the enjoyment rating for the stimulus in the row - 'match' is the match rating for the stimulus in the row - 'happy' is the happiness rating for the stimulus in the row - 'moving' is the movingness rating for the stimulus in the row - 'sad' is the sadness rating for the stimulus in the row - 'description' is the intention description (positive, negative, neutral) for the stimulus in the row - 'stimType' is the type of stimulus (music or poem) for the stimulus in the row - Additionally, note that values of 99 are used as a placeholder for data for one of the positive stimulus excerpts that was not recorded due to a programming error.

Paper Summary

Ambiguity, defined as "the capacity to sustain multiple interpretations" (Margulis et al., 2017) is a property inherent of art. This paper, titled "Expressive intent, ambiguity, and aesthetic experiences of music and poetry", examines the role of ambiguity in the aesthetic processing of poetry and music, specifically how expressive disambiguation of the composer or author's emotional or communicative intent in writing the piece affects an audience's enjoyment and perception of that piece. To see whether this information influenced their perception of the excerpts, participants were exposed to 30 second excerpts of music and poetry that were previously categorized as expressively ambiguous (works that can support multiple interpretations, e.g. can be understood as being either positively or negatively valenced). 'Expressive disambiguation' surfaces as either neutral, positively valenced, or negatively valenced information about the author or composer's intentions in their composition of each excerpt, which is presented to the participants prior to hearing the excerpts.

Descriptions of authorial intention served as expressive disambiguation - An example of a positive intention is "*The author/composer wrote this poem to express his passion and devotion for his love*)" (Margulis et al., 2017) - An example of ambiguous intention is "*The author/composer wrote this piece to express mourning over the death of a family member*)" (Margulis et al., 2017) - An example of a negative intention is "*The author/composer wrote this poem to experiment with different writing techniques"* (Margulis et al., 2017)

Research Questions: The three aims of this paper are to find out

1. Whether extrinsic information or expressive disambiguation of the artist's intent affects aesthetic appreciation differently for musical and poetic excerpts. This is relevant because poetry, with its use of language and its carrying of semantic meaning, is considerably less vague than music.
    - If this information affects the processing of music and poetry differently, one would have to further investigate the semantic aspects of poetry and compare it with that of music
2. Whether verbal/extrinsic information about an artist's expressive intent can influence the way a piece of music or poetry is processed affectively. Specifically, whether positively valenced information leads

participants to perceive excerpts as happier, and negatively valenced information leads participants to perceive excerpts as sadder, and if the same information impacts the evaluation of an excerpt's enjoyability.

- If positively and negatively valenced music is leads participants to perceive excerpts as happier and sadder, respectively, it would indicate that verbal information provided before aesthetic experience can be integrated into the emotional processing component of art. Furthermore, if being given this information affects the rating of an excerpt's enjoyability or movingness, "it would suggest there existing a role for cultural messaging beyind the intrinsic content of a work of art" (Margulis et al., 2017)

3. Whether people prefer and are more moved by experiencing music and poetry with the ambiguity intact. This is consulted by analyzing the effects of neutral intent information. The authors control for the content of the excerpts by changing the pairings of descriptions and excerpts.

The paper further aims to identify the role of an audience's empathy with a perceived human artist in the generation of aesthetic experience of music and poetry, and the possible factor of empathy is traced in comparing the 'match' scores, i.e. positively valenced information + happy excerpts.

Experimental Conditions, Participant and Stimuli Information

The music excerpts were all instrumental (no lyrics) reported by the authors to straightforwardly convey either positive or negative affect. Four positive, four negative, and 18 ambiguous excerpts were used. - Positive excerpts, for example, included pieces in a major key, faster tempos - Negative excerpts included pieces in a minor key and slower tempos - Ambiguous excerpts mixed these structural cues

The poetic excerpts included both classic and contemporary poets and eited to last approximately 30 seconds. These were read and recorded by a professional actor who spoke with a neutral, affectively uninflected tone. Earlier in the study, another group of participants were asked to rate positivity, negativity, and ambiguity, familiarity, and enjoyment for each, and from the excerpts reported as most unfamiliar, the 4 most positive, 4 most negative, and 18 most ambiguous excerpts were selected

After hearing each excerpt, participants rated on a scale of 1-7 - How happy each excerpt seemed - How sad each excerpt seemed - How much they enjoyed the excerpt - How moving the excerpt seemed - How well the excerpt matched the composer or author's intention

In the experiment, the authors paired ambiguous excerpts with positive, negative, or neutral intent descriptions, while only pairing positive excerpts with positive or neutral intent descriptions, and negative excerpts only with negative or neutral descriptions so as to enhance believability of the pairings. The same descriptions were used for both poem and music excerpts, and no participant received the same description twice. 1/3 of the excerpts were paired with negatively valenced intent information, 1/3 was paired with positively valenced information, and 1/3 paired with neutrally valenced information. Each participant was exposed to positive, negative, and neutral music and poetic excerpts paired with positive, negative, and neutral descriptions, and participants were randomly assigned to one of the six lists of stimulus pairings, and for the ambiguous excerpts. The design was a 2 x 3 (stimulus type: music vs poetry) × (intention description: positive, negative, neutral) repeated-measures study.

Model Selection

The authors used linear mixed modelling of dependent measures, and participants and stimuli were chosen as the random-effects variables. Models were fitted "with maximal random effects structure and random slopes for each of the fixed factors within each participant and stimulus. If the maximal model failed to converge, the random-effects structure was simplified incrementally by removing one random slope at a time, the one that explained the least variance in the model that did not converge." (Margulis et al., 2017)

Because the authors used a repeated measures design in their experiment, the independence assumption would not be met if they did not take into account the dependencies within the data collected. It follows that the authors chose to use linear mixed modelling, for this would allow them to account for the multiple observations taken from a single participant as well as the multiple times the same poetic or musical excerpt was used. Because the observations recorded do not vary independently of each other, the aggregation of participants and stimulus type will better account for the patterns within the collected data.

Every plot in the paper displays the mean ratings of a column within the data frame (enjoyment, happiness, sadness, and movingness ratings) as a function of intention description and stimulus type. The authors' decision to use only the mean ratings to compare the six possible combinations of intention description and stimulus type is not the most informative of the underlying distribution of ratings. Additionally, their inclusion of the standard error of the group means as the plot's error bars instead of the standard deviation or confidence interval was of particular interest to me. A possible factor for including the SE can be found in the authors' motivations for this study. They commented on the novelty of the domain of their research questions in that not much inquiry has been made into notions that aesthetic experience and perception of art could be dependent upon the concept of empathy [with a human artist]. Noting that every research question is an aim to discover whether something happens/exists or not, it follows that what the authors are investigating is very much just the beginning of a series of possible trajectories for more specific questions. As such, the authors' decision to report the SE appears indicative of their desire to make an inference about the population. The SD only indicates how variable each individual in their sample data is from their mean, but in their desire to make a broad inference about how humans perceive and experience art, the SE seems more relevant as it measures the uncertainty of their calculated means encompassing the true population means. In what seems to be an initial glance into the nature of perception in the arts, the authors are testing general claims and inquiries which are important for laying the groundwork of future studies within the delicate domain of something very fundamental, abstract, and human.

Loading in Packages

Data Setup

```
df <- read_csv('Quant Lecture and Section Data/df.csv')
```

```
## New names:
## * `` -> ...1
```

```
## Rows: 6136 Columns: 11
```

```
## ── Column specification ────────────────────────────────────────────────────────
## Delimiter: ","
## chr (4): excerpt, stimID, description, stimType
## dbl (7): ...1, subject, enjoy, match, happy, moving, sad
```

```
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
#I am filtering out the rows containing the value 99 because the authors said these valu
es are a placeholder for data that was not collected in the following columns due to a p
rogramming error

df <- subset(df, enjoy != 99 & match != 99 & happy != 99 & moving != 99 & sad != 99)
```

The first group of calculations the authors made were to make sure that participants were on the same page as the authors on the differentiation of the stimuli/excerpts. They confirmed this shared agreement by comparing happiness and sadness ratings of previously defined positive vs negative excerpts. This checked out in their reported mean and SE values, which are reported below.

The authors reported the means and standard errors for positive vs negative excerpts for happiness ratings which are stated below: - Positive excerpts (M = 5.37, SE = 0.13) - Negative excerpts (M = 1.79, SE = 0.11), t(17.9) = 15.64, p < .001.

For sadness ratings, the authors reported that they negative excerpts were rated sadder than positive excerpts, and that these effects did not interact significantly with excerpt type: - Negative excerpts (M = 5.04, SE = 0.12) - Positive excerpts (M = 1.73, SE = 0.10), t(18.2) = 16.5, p < .001

Additionally, the authors noted that music was overall rated happier than poetry, - Poetry (M = 3.11, SE = 0.11), t(15.6) = 3.24, p = .005 - Music (M = 3.87, SE = 0.12).

And music was was also rated less sad than poetry. - Music (M = 3.21, SE = 0.11) - Poetry (M = 3.75, SE = 0.11), t(17.6) = 1.60, p = .13

The authors concluded that happiness ratings were higher for positive than negative excerpts, sadness ratings were higher for negative than positive excerpts, and that these effects did not have significant interaction with the type of the excerpt (music vs poetry).

When I calculated these values using the data provided, my results showed a similar trend in that mean happiness ratings were higher for positive than negative excerpts, and mean sadness ratings were higher for negative than positive excerpts. I also saw that happiness ratings in music was higher than happiness ratings for poetry, and sadness ratings for music was lower than sadness ratings for poetry. My mean values were a few decimal places off from the authors', and a trend I identified is that the dichotomy between the two groups being compared in every area of measurement is slightly more pronounced. For example, my mean happiness rating for positive excerpts is higher, and my mean happiness rating for negative excerpts is lower than the authors' reported values. Additionally, my SE values throughout were consistently and considerably lower than the reported SE values. This is likely the result of some unreported pre-processing of data on the authors' part that potentially included the removal of certain observations and/or outliers from the data they published alongside the paper.

Mean and SE tables of happy and sad ratings for positive and negative excerpts

```
#selecting the data for only positive and negative excerpts
excerpt_pn <- df %>% filter(df$excerpt== c('positive', 'negative'))

#summary of happiness ratings (mean and SE) for positive and negative excerpts
as_tibble(excerpt_pn %>% group_by(excerpt) %>% summarize(happy_mean = mean(happy), happy
_se = sd(happy)/sqrt(n())))
```

```
## # A tibble: 2 x 3
##   excerpt  happy_mean happy_se
##   <chr>         <dbl>    <dbl>
## 1 negative       1.80   0.0524
## 2 positive       5.31   0.0665
```

```
#t test
t.test(happy ~ excerpt, data=excerpt_pn)
```

```
##
##  Welch Two Sample t-test
##
## data:  happy by excerpt
## t = -41.499, df = 809.54, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group negative and group pos
itive is not equal to 0
## 95 percent confidence interval:
##  -3.679498 -3.347137
## sample estimates:
## mean in group negative mean in group positive
##               1.796610               5.309927
```

```
#summary of sadness ratings (mean and SE) for positive and negative excerpts
excerpt_pn %>% group_by(excerpt) %>% summarize(sad_mean = mean(sad), sad_se = sd(sad)/sq
rt(n()))
```

```
## # A tibble: 2 x 3
##   excerpt  sad_mean sad_se
##   <chr>       <dbl>  <dbl>
## 1 negative     5.01 0.0690
## 2 positive     1.81 0.0536
```

```
t.test(sad ~ excerpt, data=excerpt_pn)
```

```
##
##  Welch Two Sample t-test
##
## data:  sad by excerpt
## t = 36.687, df = 854.72, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group negative and group pos
itive is not equal to 0
## 95 percent confidence interval:
##  3.034874 3.377959
## sample estimates:
## mean in group negative mean in group positive
##               5.012712               1.806295
```

```
#summary of happiness ratings for music vs poetry: Mean and SE
df %>% group_by(stimType) %>% summarize(happy_mean = mean(happy), happy_se = sd(happy)/s
qrt(n()))
```

```
## # A tibble: 2 x 3
##   stimType happy_mean happy_se
##   <chr>         <dbl>    <dbl>
## 1 music          4.11   0.0325
## 2 poem           2.90   0.0312
```

```
t.test(happy ~ stimType, data=excerpt_pn)
```

```
##
##  Welch Two Sample t-test
##
## data:  happy by stimType
## t = 3.3983, df = 862.98, p-value = 0.000709
## alternative hypothesis: true difference in means between group music and group poem i
s not equal to 0
## 95 percent confidence interval:
##  0.2068758 0.7725431
## sample estimates:
## mean in group music  mean in group poem
##            3.697337            3.207627
```

```
#summary of sadness ratings for music vs poetry: Mean and SE
df %>% group_by(stimType) %>% summarize(sad_mean = mean(sad), sad_se = sd(sad)/sqrt(n
()))
```

```
## # A tibble: 2 x 3
##   stimType sad_mean sad_se
##   <chr>       <dbl>  <dbl>
## 1 music        2.74 0.0315
## 2 poem         3.70 0.0326
```

```
t.test(sad ~ stimType, data=excerpt_pn)
```

```
##
##  Welch Two Sample t-test
##
## data:  sad by stimType
## t = -2.6411, df = 861.38, p-value = 0.008414
## alternative hypothesis: true difference in means between group music and group poem i
s not equal to 0
## 95 percent confidence interval:
##  -0.64313020 -0.09476326
## sample estimates:
## mean in group music  mean in group poem
##            3.319613            3.688559
```

The authors then calculated intent description match ratings for positive and negative excerpts. The match ratings for the combination of positive excerpts and positive intentions were compared to the match ratings for ambiguous excerpts and 'any kind of intention', and the match ratings for negative excerpts and negative intentions were compared to the match scores for ambiguous excerpts and any kind of intention. They concluded that "the positive and negative excerpts respectively matched positive and negative intention descriptions better than ambiguous excepts matched any kind of intention description" because positive excerpts and positive intentions "were much higher than for ambiguous excerpts and any kind of intention description" (Margulis et al., 2017).

- Match rating between positive excerpts and positive intentions (M = 5.21, SE = 0.14)
- Match rating between negative excerpts and negative intentions (M = 5.46, SE = 0.15)
- Match rating for ambiguous excerpts and any kind of intention description(M = 3.95, SE = 0.09)

Additionally, the authors reported that intention description significantly interacted with excerpt types (poetry vs music) in predicting match scores $F(2, 83.7) = 3.37$, $p = .04$

Intent Description and Match ratings for positive, negative, and ambiguous excerpts

```
#summary of match ratings for the 8 possible combinations of excerpts (positive, negativ
e, ambiguous) and descriptions (neutral, positive, negative)

df %>% group_by(excerpt, description) %>% summarize(match_mean = mean(match), match_se =
sd(match)/sqrt(n()), match_mean = mean(match), match_se = sd(match)/sqrt(n()))
```

```
## `summarise()` has grouped output by 'excerpt'. You can override using the `.groups` a
rgument.
```

```
## # A tibble: 8 x 4
## # Groups:   excerpt [3]
##   excerpt   description match_mean match_se
##   <chr>     <chr>            <dbl>    <dbl>
## 1 ambiguous negative          3.94   0.0505
## 2 ambiguous neutral           3.89   0.0444
## 3 ambiguous positive          4.05   0.0496
## 4 negative  negative          5.53   0.0650
## 5 negative  neutral           3.90   0.0748
## 6 negative  positive          5.28   0.205
## 7 positive  neutral           4.23   0.0863
## 8 positive  positive          5.16   0.0764
```

```
ambiguous <- df %>% filter(df$excerpt == 'ambiguous')

#match rating for ambiguous excerpts and any description
ambiguous %>% summarize(match_mean = mean(match), match_se = sd(match)/sqrt(n()), match_
mean = mean(match), match_se = sd(match)/sqrt(n()))
```

```
## # A tibble: 1 x 2
##   match_mean match_se
##        <dbl>    <dbl>
## 1       3.96   0.0278
```

```
#model predicting match scores from excerpt type and intention description
match_lm <- lm(match ~ stimType + description, data=df)
summary(match_lm)
```

```
##
## Call:
## lm(formula = match ~ stimType + description, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4433 -1.4081  0.1502  1.5567  3.1502
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          4.40812    0.04780  92.217  < 2e-16 ***
## stimTypepoem        -0.21903    0.04654  -4.706 2.58e-06 ***
## descriptionneutral  -0.33932    0.05649  -6.007 2.00e-09 ***
## descriptionpositive  0.03521    0.05923   0.594    0.552
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.804 on 6014 degrees of freedom
## Multiple R-squared:  0.01286,    Adjusted R-squared:  0.01237
## F-statistic: 26.12 on 3 and 6014 DF,  p-value: < 2.2e-16
```

```
#new model predicting match scores from the interaction of excerpt type and intention de
scription
match_lm1 <- lm(match ~ stimType * description, data=df)
summary(match_lm1)
```

```
##
## Call:
## lm(formula = match ~ stimType * description, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7701 -1.6697  0.2299  1.2669  3.2669
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       3.95657    0.05779  68.460  < 2e-16 ***
## stimTypepoem                      0.71311    0.08304   8.588  < 2e-16 ***
## descriptionneutral                0.23488    0.07841   2.996  0.00275 **
## descriptionpositive               0.81353    0.08313   9.786  < 2e-16 ***
## stimTypepoem:descriptionneutral  -1.17151    0.11124 -10.531  < 2e-16 ***
## stimTypepoem:descriptionpositive -1.54714    0.11669 -13.259  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.776 on 6012 degrees of freedom
## Multiple R-squared:  0.04353,    Adjusted R-squared:  0.04273
## F-statistic: 54.72 on 5 and 6012 DF,  p-value: < 2.2e-16
```

```
#anova to assess the significance of the additional interaction as a predictor that show
s a significant interaction between excerpt type and intention description
anova(match_lm, match_lm1)
```

```
## Analysis of Variance Table
##
## Model 1: match ~ stimType + description
## Model 2: match ~ stimType * description
##   Res.Df   RSS Df Sum of Sq      F    Pr(>F)
## 1   6014 19564
## 2   6012 18956  2    607.76 96.375 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Finally, the authors observed a significant difference in how well the intention descriptions were perceived to match the ambiguous excerpts, reporting higher match ratings for positive intention descriptions than negative and neutral intention descriptions. They concluded that the difference in match ratings across intention descriptions to be significant $F(2, 176.5) = 3.40$, $p = .04$

- Match ratings for positive intention descriptions ($M = 4.05$, $SE = 0.10$)
- Match ratings for negative intention descriptions ($M = 3.94$, $SE = 0.10$)
- Match ratings for neutral intention descriptions ($M = 3.89$, $SE = 0.10$)

Match scores for music were higher than match scores for poetry, but the authors reported that this difference was not significant. They noted, however, that the interaction between intention description and excerpt type was very strong, “reflecting that for music, positive and neutral intention descriptions are better matched than negative descriptions, but the opposite pattern emerges for poetry”. This interaction is what motivated the authors to consider that the possibility that “the perceived match between intention description and stimulus mediates aesthetic experience”. They further concluded that since there is no significant difference between music and poetry on overall match ratings, this confirms “that the intent descriptions were equally well-suited to the music and poetry samples.” (Margulis et al., 2017)

- Match ratings for music M = 4.10, SE = 0.12)
- Match ratings for poetry (M = 3.82, SE = 0.12), $F(1, 44.0) = 4.03$, $p = .051$

```
#match ratings for ambiguous excerpts and the three intention descriptions
ambiguous %>% group_by(excerpt, description) %>% summarize(match_mean = mean(match), mat
ch_se = sd(match)/sqrt(n()))
```

```
## `summarise()` has grouped output by 'excerpt'. You can override using the `.groups` a
rgument.
```

```
## # A tibble: 3 x 4
## # Groups:   excerpt [1]
##   excerpt   description match_mean match_se
##   <chr>     <chr>            <dbl>    <dbl>
## 1 ambiguous negative          3.94   0.0505
## 2 ambiguous neutral           3.89   0.0444
## 3 ambiguous positive          4.05   0.0496
```

```
#testing the significance of match ratings across intention descriptions
match.lm <- lm(match ~ description, data=ambiguous)
summary(match.lm)
```

```
##
## Call:
## lm(formula = match ~ description, data = ambiguous)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0473 -1.8927  0.1073  1.1073  3.1073
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          3.94209    0.04822  81.757   <2e-16 ***
## descriptionneutral  -0.04944    0.06819  -0.725    0.469
## descriptionpositive  0.10523    0.06819   1.543    0.123
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.814 on 4245 degrees of freedom
## Multiple R-squared:  0.001263,   Adjusted R-squared:  0.0007923
## F-statistic: 2.684 on 2 and 4245 DF,  p-value: 0.06842
```

```
#match ratings for music vs poetry
ambiguous %>% group_by(excerpt, stimType) %>% summarize(match_mean = mean(match), match_
se = sd(match)/sqrt(n()))
```

```
## `summarise()` has grouped output by 'excerpt'. You can override using the `.groups` a
rgument.
```

```
## # A tibble: 2 x 4
## # Groups:   excerpt [1]
##   excerpt   stimType match_mean match_se
##   <chr>     <chr>         <dbl>    <dbl>
## 1 ambiguous music          4.10   0.0397
## 2 ambiguous poem           3.82   0.0388
```

```
t.test(match ~ stimType, data=ambiguous)
```

```
##
##  Welch Two Sample t-test
##
## data:  match by stimType
## t = 4.9927, df = 4243.8, p-value = 6.192e-07
## alternative hypothesis: true difference in means between group music and group poem i
s not equal to 0
## 95 percent confidence interval:
##  0.1684151 0.3861988
## sample estimates:
## mean in group music  mean in group poem
##            4.099341            3.822034
```

```
#match ratings for ambiguous excerpts and positive descriptions grouped by stimType
ambiguous %>% filter(ambiguous$description == 'positive') %>% group_by(stimType) %>% sum
marize(positive_match_mean = mean(match), positive_match_se = sd(match)/sqrt(n()))
```

```
## # A tibble: 2 x 3
##   stimType positive_match_mean positive_match_se
##   <chr>                  <dbl>             <dbl>
## 1 music                   4.63            0.0654
## 2 poem                    3.46            0.0677
```

```
#match ratings for ambiguous excerpts and negative descriptions grouped by stimType
ambiguous %>% filter(ambiguous$description == 'negative') %>% group_by(stimType) %>% sum
marize(negative_match_mean = mean(match),negative_match_se = sd(match)/sqrt(n()))
```

```
## # A tibble: 2 x 3
##   stimType negative_match_mean negative_match_se
##   <chr>                  <dbl>             <dbl>
## 1 music                   3.48            0.0716
## 2 poem                    4.40            0.0669
```

Plot Replication

There are four visualizations included within the paper that show the means of the ambiguous data grouped by happiness, sadness, enjoyment, and match ratings, and the standard error bars are included within the plot. The authors only used ratings for ambiguous excerpts, so I first filtered df to only include those ratings, which were then “examined as a function of intention description and excerpt type, as well as their interaction.” Participants and stimuli were chosen as the random effects, and the authors first fitted models with maximal random-effects structure “that included random slopes for each of the fixed factors within each participant and stimulus. If the maximal model failed to converge, the random-effects structure was simplified incrementally by removing one random slope at a time, the one that explained the least variance in the model that did not converge.” (Margulis et al., 2017)

Because authors did not specify the random slopes used in their final model, I had to retrace their steps, which I did by fitting a maximal model and incrementally removing slopes until the model converged. I used the interaction of ‘description’ * ‘stimType’ as the fixed effects, and used that to predict the means of the following ratings: ‘enjoy’, ‘happy’, ‘moving’, and ‘sad’. For the random effects, I grouped by subject (each unique participant) and stimID (each unique excerpt) as the authors did, and included random slopes for each fixed effect within each participant and stimulus, gradually reducing the complexity of the random effects structure.

Testing different random slope structures

```
ambiguous <- df %>% filter(df$excerpt == 'ambiguous')

#maximal random slope model structure; this model failed to converge
enjoy.lm0 <- lmer(enjoy ~ description * stimType + (1+description+stimType|subject) + (1
+description+stimType|stimID), data=ambiguous)
```

```
## boundary (singular) fit: see ?isSingular
```

```
## Warning: Model failed to converge with 3 negative eigenvalues: -1.4e-02 -3.6e+00
## -1.0e+02
```

```
summary(enjoy.lm0)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: enjoy ~ description * stimType + (1 + description + stimType |
##     subject) + (1 + description + stimType | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 14943.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8154 -0.6585  0.0046  0.6562  3.2116
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.599529 0.77429
##           descriptionneutral  0.002271 0.04765  -1.00
##           descriptionpositive 0.100305 0.31671  -0.21  0.21
##           stimTypepoem        0.490583 0.70042  -0.07  0.07 -0.26
##  stimID   (Intercept)         0.318853 0.56467
##           descriptionneutral  0.000811 0.02848  -1.00
##           descriptionpositive 0.041716 0.20425   0.17 -0.17
##           stimTypepoem        0.687041 0.82888  -1.00  1.00 -0.20
##  Residual                     1.709763 1.30758
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                                  Estimate Std. Error         df t value
## (Intercept)                       3.87845    0.15878   26.05353  24.426
## descriptionneutral                0.23157    0.07000 1092.28089   3.308
## descriptionpositive               0.38298    0.08947   46.32733   4.281
## stimTypepoem                     -0.42285    0.17495   39.24927  -2.417
## descriptionneutral:stimTypepoem  -0.50567    0.09880 1135.36436  -5.118
## descriptionpositive:stimTypepoem -0.54866    0.11961   39.48531  -4.587
##                                  Pr(>|t|)
## (Intercept)                       < 2e-16 ***
## descriptionneutral                0.00097 ***
## descriptionpositive              9.28e-05 ***
## stimTypepoem                      0.02039 *
## descriptionneutral:stimTypepoem  3.62e-07 ***
## descriptionpositive:stimTypepoem 4.47e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn dscrptnp stmTyp dscrptnn:T
## dscrptnntrl -0.326
## dscrptnpstv -0.122  0.381
## stimTypepom -0.736  0.272    0.053
## dscrptnnt:T  0.211 -0.706   -0.267   -0.307
## dscrptnps:T  0.069 -0.282   -0.669   -0.216  0.400
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
#removing description as the random slope in stimID because it exhibits the least varian
ce. This model failed to converge
enjoy.lm1 <- lmer(enjoy ~ description * stimType + (1+description+stimType|subject) + (1
+stimType|stimID), data=ambiguous)
```

```
## boundary (singular) fit: see ?isSingular
```

```
## Warning: Model failed to converge with 2 negative eigenvalues: -1.1e-02 -8.9e+00
```

```
summary(enjoy.lm1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: enjoy ~ description * stimType + (1 + description + stimType |
##     subject) + (1 + stimType | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 14949.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8447 -0.6676  0.0039  0.6563  3.2965
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.600115 0.77467
##           descriptionneutral  0.002322 0.04819  -1.00
##           descriptionpositive 0.098775 0.31429  -0.20  0.20
##           stimTypepoem        0.488481 0.69891  -0.07  0.07 -0.26
##  stimID   (Intercept)         0.316088 0.56222
##           stimTypepoem        0.549324 0.74116  -0.95
##  Residual                     1.719472 1.31129
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                                   Estimate Std. Error         df t value
## (Intercept)                        3.87868    0.15836   31.04671  24.493
## descriptionneutral                 0.23145    0.06988 3338.82532   3.312
## descriptionpositive                0.38393    0.07550  362.65751   5.085
## stimTypepoem                      -0.42310    0.17476   43.37838  -2.421
## descriptionneutral:stimTypepoem   -0.50569    0.09861 3856.70220  -5.128
## descriptionpositive:stimTypepoem  -0.54986    0.09861 3856.70252  -5.576
##                                  Pr(>|t|)
## (Intercept)                       < 2e-16 ***
## descriptionneutral               0.000935 ***
## descriptionpositive              5.90e-07 ***
## stimTypepoem                     0.019729 *
## descriptionneutral:stimTypepoem  3.07e-07 ***
## descriptionpositive:stimTypepoem 2.63e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn dscrptnp stmTyp dscrptnn:T
## dscrptnntrl -0.248
## dscrptnpstv -0.238  0.466
## stimTypepom -0.733  0.201    0.148
## dscrptnnt:T  0.156 -0.706   -0.327   -0.282
## dscrptnps:T  0.156 -0.353   -0.653   -0.282  0.500
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
#removing stimType as the random slope within stimID because it exhibits the least varia
nce. This model converged
enjoy.lm2 <- lmer(enjoy ~ description * stimType + (1+description+stimType|subject) + (1
|stimID), data=ambiguous)
summary(enjoy.lm2)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: enjoy ~ description * stimType + (1 + description + stimType |
##     subject) + (1 | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 14955.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8397 -0.6652  0.0086  0.6545  3.2828
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.59705  0.7727
##           descriptionneutral  0.01618  0.1272   -0.38
##           descriptionpositive 0.09439  0.3072   -0.19  0.03
##           stimTypepoem        0.48996  0.7000   -0.04 -0.30 -0.33
##  stimID   (Intercept)         0.19373  0.4402
##  Residual                     1.71679  1.3103
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                                  Estimate Std. Error         df t value
## (Intercept)                       3.87850    0.13509   76.86858  28.710
## descriptionneutral                0.23161    0.07066  429.26390   3.278
## descriptionpositive               0.38431    0.07520  351.78184   5.110
## stimTypepoem                     -0.42285    0.17473   58.05481  -2.420
## descriptionneutral:stimTypepoem  -0.50587    0.09854 3742.39398  -5.134
## descriptionpositive:stimTypepoem -0.55042    0.09854 3742.39350  -5.586
##                                 Pr(>|t|)
## (Intercept)                      < 2e-16 ***
## descriptionneutral               0.00113 **
## descriptionpositive             5.29e-07 ***
## stimTypepoem                     0.01867 *
## descriptionneutral:stimTypepoem  2.98e-07 ***
## descriptionpositive:stimTypepoem 2.49e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn dscrptnp stmTyp dscrptnn:T
## dscrptnntrl -0.287
## dscrptnpstv -0.277  0.459
## stimTypepom -0.567  0.178    0.139
## dscrptnnt:T  0.182 -0.697   -0.328   -0.282
## dscrptnps:T  0.182 -0.349   -0.655   -0.282  0.500
```

The authors reported no significant effect of intention description type on enjoyment, $F(2, 66.9) = 1.93$, $p = .15$, but a clear interaction of intention and excerpt type, $F(2, 54.9) = 15.51$, $p < .001$. They further observed that negative intentions increased enjoyment of poetry relative to the neutral descriptions and decreased enjoyment of music relative to the neutral condition. There was also a large enjoyment advantage for music (M = 4.08, SE = 0.13) over poetry (M = 3.31, SE = 0.14), t(49.8) = 5.08, p < .001. (Margulis et al., 2017)

This figure shows that music is consistently rated with higher enjoyment scores than poetry across all three intention descriptions. Additionally, music seems to be the most enjoyable when paired with positive descriptions, and the least enjoyable when paired with negative descriptions. This is not the case for poetry, which seems to be the most enjoyable when paired with negative descriptions, and the least enjoyable when paired with neutral descriptions.

Figure 1: Dot plot with error bars showing the standard error of mean enjoyment ratings as a function of intention description and stimulus type (music or poem)

```
ambiguous <- df %>% filter(df$excerpt == 'ambiguous')

#the effect of intention description type on enjoyment
enjoy.lm0 <- lmer(enjoy ~ description + (1+description+stimType|subject) + (1|stimID), d
ata=ambiguous)

summary(enjoy.lm0)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: enjoy ~ description + (1 + description + stimType | subject) +
##     (1 | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 15004.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7598 -0.6704  0.0130  0.6595  3.1763
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.59606  0.7721
##           descriptionneutral  0.01355  0.1164   -0.40
##           descriptionpositive 0.09176  0.3029   -0.19  0.00
##           stimTypepoem        0.49993  0.7071   -0.04 -0.33 -0.33
##  stimID   (Intercept)         0.31309  0.5595
##  Residual                     1.73325  1.3165
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                     Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)          3.70377    0.12547  81.52018  29.518   <2e-16 ***
## descriptionneutral  -0.02577    0.05064 116.47742  -0.509    0.612
## descriptionpositive  0.09732    0.05674 117.21816   1.715    0.089 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn
## dscrptnntrl -0.257
## dscrptnpstv -0.264  0.426
```

```
#the effect of the interaction of intention and excerpt type on enjoyment
enjoy.lm <- lmer(enjoy ~ description * stimType + (1+description+stimType|subject) + (1|
stimID), data=ambiguous)
summary(enjoy.lm)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: enjoy ~ description * stimType + (1 + description + stimType |
##     subject) + (1 | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 14955.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8397 -0.6652  0.0086  0.6545  3.2828
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.59705  0.7727
##           descriptionneutral  0.01618  0.1272   -0.38
##           descriptionpositive 0.09439  0.3072   -0.19  0.03
##           stimTypepoem        0.48996  0.7000   -0.04 -0.30 -0.33
##  stimID   (Intercept)         0.19373  0.4402
##  Residual                     1.71679  1.3103
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                                   Estimate Std. Error         df t value
## (Intercept)                        3.87850    0.13509   76.86858  28.710
## descriptionneutral                 0.23161    0.07066  429.26390   3.278
## descriptionpositive                0.38431    0.07520  351.78184   5.110
## stimTypepoem                      -0.42285    0.17473   58.05481  -2.420
## descriptionneutral:stimTypepoem   -0.50587    0.09854 3742.39398  -5.134
## descriptionpositive:stimTypepoem  -0.55042    0.09854 3742.39350  -5.586
##                                  Pr(>|t|)
## (Intercept)                       < 2e-16 ***
## descriptionneutral                0.00113 **
## descriptionpositive              5.29e-07 ***
## stimTypepoem                      0.01867 *
## descriptionneutral:stimTypepoem  2.98e-07 ***
## descriptionpositive:stimTypepoem 2.49e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn dscrptnp stmTyp dscrptnn:T
## dscrptnntrl -0.287
## dscrptnpstv -0.277  0.459
## stimTypepom -0.567  0.178    0.139
## dscrptnnt:T  0.182 -0.697   -0.328   -0.282
## dscrptnps:T  0.182 -0.349   -0.655   -0.282  0.500
```

```
#comparing the two models
anova(enjoy.lm0, enjoy.lm)
```

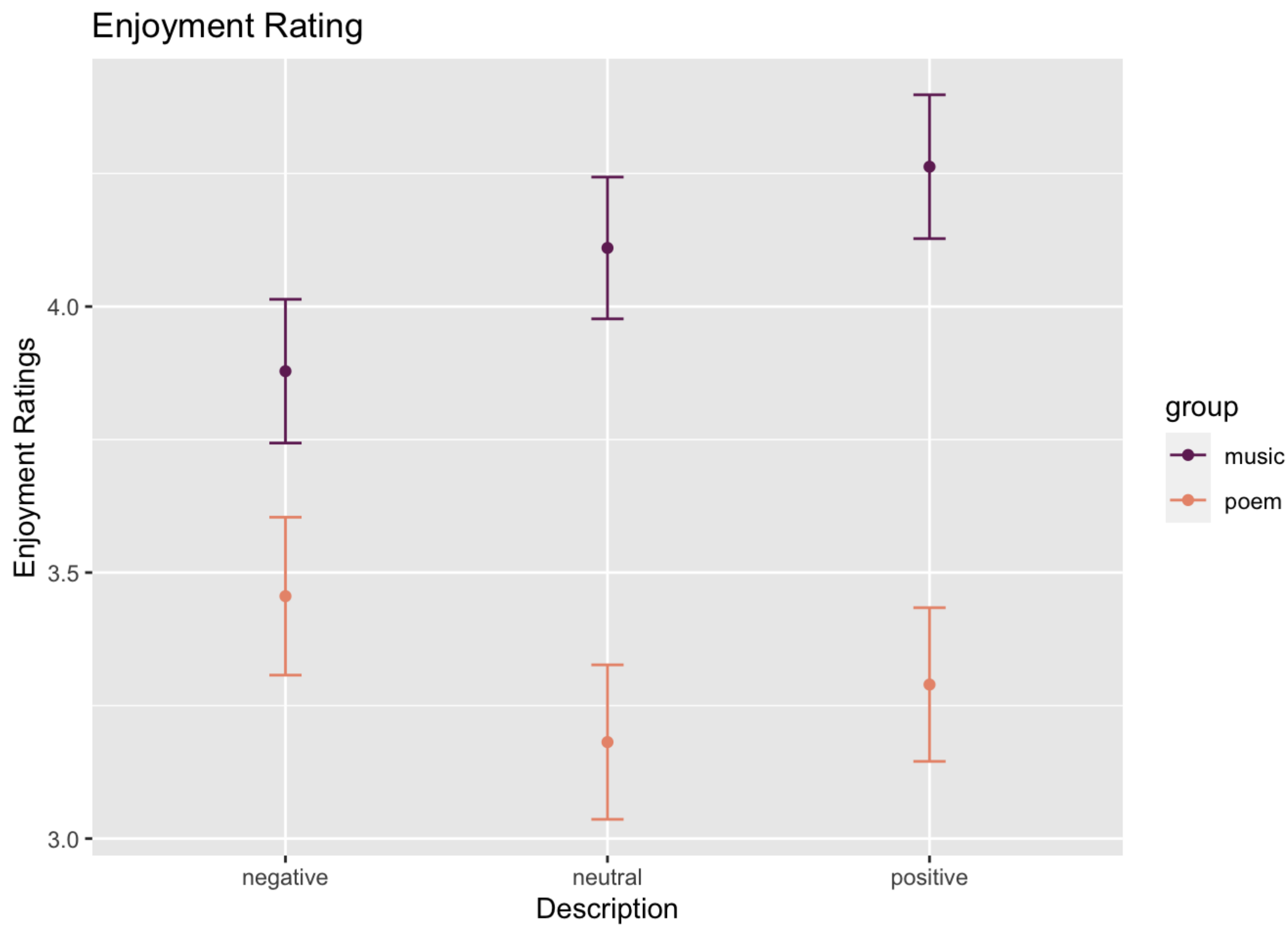
```
## refitting model(s) with ML (instead of REML)
```

```
## Data: ambiguous
## Models:
## enjoy.lm0: enjoy ~ description + (1 + description + stimType | subject) + (1 | stimI
D)
## enjoy.lm: enjoy ~ description * stimType + (1 + description + stimType | subject) +
(1 | stimID)
##           npar   AIC   BIC  logLik deviance Chisq Df Pr(>Chisq)
## enjoy.lm0   15 15024 15119 -7496.9    14994
## enjoy.lm    18 14973 15087 -7468.5    14937  56.8  3  2.835e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
enjoy_preds <- as_tibble(ggpredict(enjoy.lm, terms = c('description', 'stimType')))
enjoy_preds
```

```
## # A tibble: 6 x 6
##   x        predicted std.error conf.low conf.high group
##   <fct>        <dbl>     <dbl>    <dbl>     <dbl> <fct>
## 1 negative      3.88     0.135     3.61      4.14 music
## 2 negative      3.46     0.148     3.16      3.75 poem
## 3 neutral       4.11     0.133     3.85      4.37 music
## 4 neutral       3.18     0.145     2.90      3.47 poem
## 5 positive      4.26     0.135     4.00      4.53 music
## 6 positive      3.29     0.144     3.01      3.57 poem
```

```
ggplot(enjoy_preds, aes(x=x, y=predicted, fill=group)) +
  geom_errorbar(
    aes(ymin = predicted-std.error, ymax = predicted+std.error, color = group),
    position = position_dodge(0), width = 0.2
    ) +
  geom_point(aes(color = group), position = position_dodge(0)) +
  scale_color_manual(values = c('#702963', '#E9967A')) +
  labs(title='Enjoyment Rating', x='Description', y='Enjoyment Ratings')
```



For the next model, the authors reported a “clear, predictable effect of intentions, F(2, 154.0) = 199.46, p < .001, such that positive intentions led to an increase in happiness ratings relative to neutral descriptions, and negative intentions led to a decrease in happiness ratings relative to neutral descriptions. Music (M = 4.26, SE = 0.16) elicited higher happiness ratings than poetry overall (M = 2.75, SE = 0.16), t(43.2) = 6.47, p < .001. The two factors did not interact significantly (F < 1).” (Margulis et al., 2017)

This figure shows that music is consistently rated happier than poetry in all three description types. Additionally, the ambiguous excerpts with positive descriptions score highest in happiness, followed by those paired with neutral descriptions and finally those paired with negative descriptions. This shows a correlation between happiness ratings and intention description types.

Figure 2: Dot plot with SE bars showing mean happiness ratings as a function of intention description and stimulus type

```
happy.lm <- lmer(happy ~ description * stimType + (1+description+stimType|subject) + (1|
stimID), data=ambiguous)
```

```
## boundary (singular) fit: see ?isSingular
```

```
## Warning: Model failed to converge with 2 negative eigenvalues: -8.7e-03 -4.1e+02
```

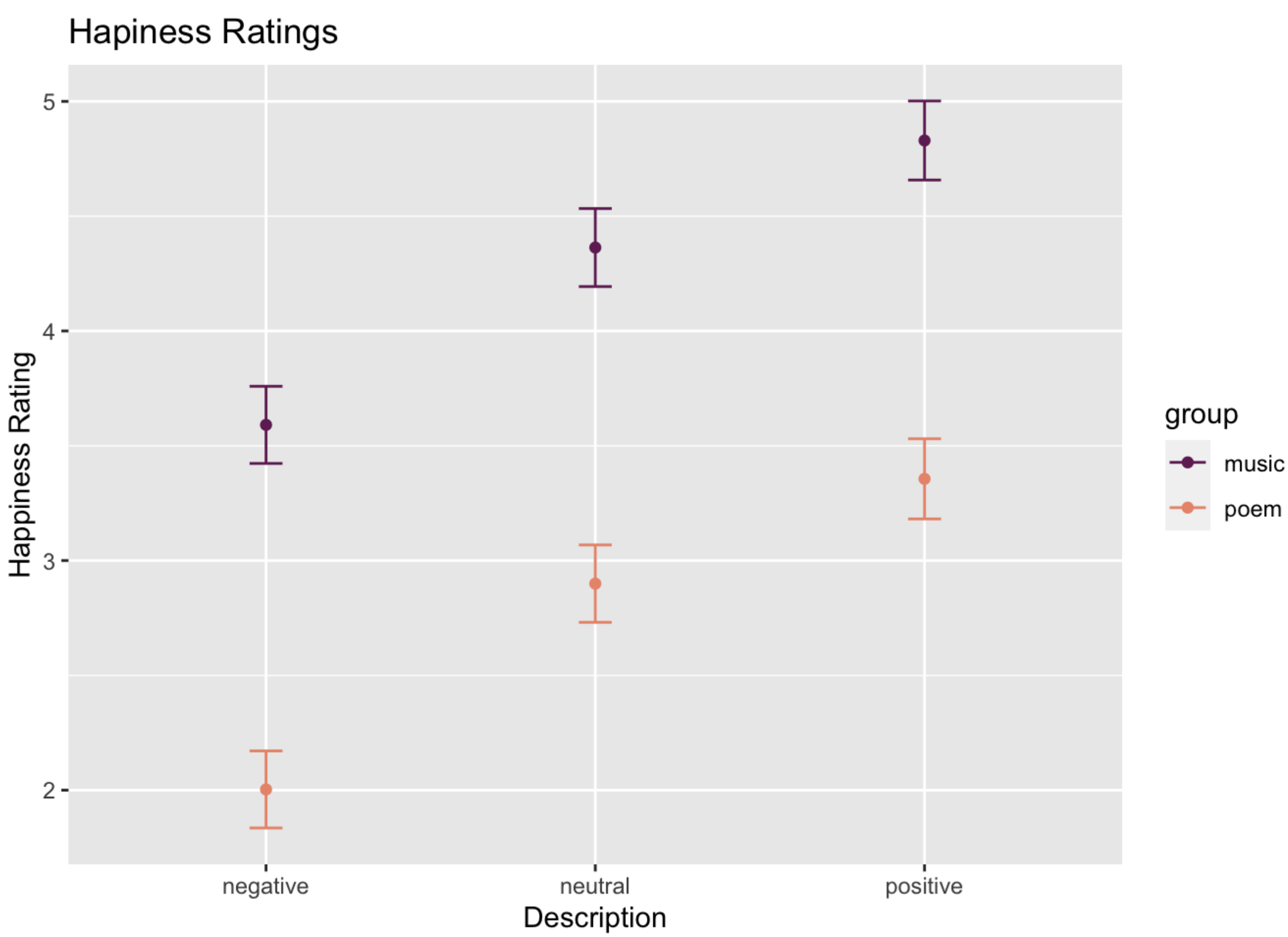
```
summary(happy.lm)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: happy ~ description * stimType + (1 + description + stimType |
##     subject) + (1 | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 13701.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1875 -0.6670 -0.0099  0.6521  3.7200
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.290128 0.53864
##           descriptionneutral  0.004457 0.06676   1.00
##           descriptionpositive 0.114379 0.33820   0.13  0.13
##           stimTypepoem        0.364797 0.60398  -0.57 -0.57  0.26
##  stimID   (Intercept)         0.431533 0.65691
##  Residual                     1.281506 1.13204
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                                  Estimate Std. Error         df t value
## (Intercept)                       3.59103    0.16806   44.54723  21.367
## descriptionneutral                0.77230    0.06052 2422.79333  12.762
## descriptionpositive               1.23851    0.06778  439.56900  18.273
## stimTypepoem                     -1.58791    0.23380   41.90187  -6.792
## descriptionneutral:stimTypepoem   0.12407    0.08514 3842.09813   1.457
## descriptionpositive:stimTypepoem  0.11422    0.08514 3842.09814   1.342
##                                 Pr(>|t|)
## (Intercept)                      < 2e-16 ***
## descriptionneutral               < 2e-16 ***
## descriptionpositive              < 2e-16 ***
## stimTypepoem                    2.92e-08 ***
## descriptionneutral:stimTypepoem    0.145
## descriptionpositive:stimTypepoem   0.180
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn dscrptnp stmTyp dscrptnn:T
## dscrptnntrl -0.148
## dscrptnpstv -0.141  0.448
## stimTypepom -0.697  0.114    0.143
## dscrptnnt:T  0.127 -0.703   -0.314   -0.182
## dscrptnps:T  0.127 -0.352   -0.628   -0.182  0.500
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
happy_preds <- as_tibble(ggpredict(happy.lm, terms = c('description', 'stimType')))
happy_preds
```

```
## # A tibble: 6 x 6
##   x        predicted std.error conf.low conf.high group
##   <fct>        <dbl>     <dbl>    <dbl>     <dbl> <fct>
## 1 negative      3.59     0.168     3.26      3.92 music
## 2 negative      2.00     0.168     1.67      2.33 poem
## 3 neutral       4.36     0.170     4.03      4.70 music
## 4 neutral       2.90     0.169     2.57      3.23 poem
## 5 positive      4.83     0.172     4.49      5.17 music
## 6 positive      3.36     0.175     3.01      3.70 poem
```

```
ggplot(happy_preds, aes(x=x, y=predicted, fill=group)) +
  geom_errorbar(
    aes(ymin = predicted-std.error, ymax = predicted+std.error, color = group),
    position = position_dodge(0), width = 0.2
    ) +
  geom_point(aes(color = group), position = position_dodge(0)) +
  scale_color_manual(values = c('#702963', '#E9967A')) +
  labs(title='Hapiness Ratings', x='Description', y='Happiness Rating')
```



According to the authors, sadness ratings are similar to those of happiness, with there being “a clear, predictable effect of intentions, $F(2, 71.4) = 106.19$, $p < .001$, such that negative intentions led to an increase in sadness ratings relative to neutral descriptions, and positive intentions led to a decrease in sadness ratings relative to neutral descriptions. Poetry ($M = 3.72$, $SE = 0.15$) elicited higher sadness ratings than music overall ($M = 2.50$, $SE = 0.15$), $t(38.5) = 5.31$, $p < .001$. The two factors did not interact significantly ($F \approx 1.5$).” (Margulis et al., 2017)

This figure shows that poetry is consistently rated sadder than music across all three intention descriptions. Additionally, for both poetry and music, excerpts paired with negative descriptions scored higher than excerpts paired with neutral descriptions which also scored higher than excerpts paired with positive descriptions. This shows a correlation between sadness ratings and intention description types, and supports the possibility that empathy with a perceived human artist does affect one’s perception of music and poetry.

Figure 3: Dot plot with SE bars showing mean sadness ratings as a function of intention description and stimulus type

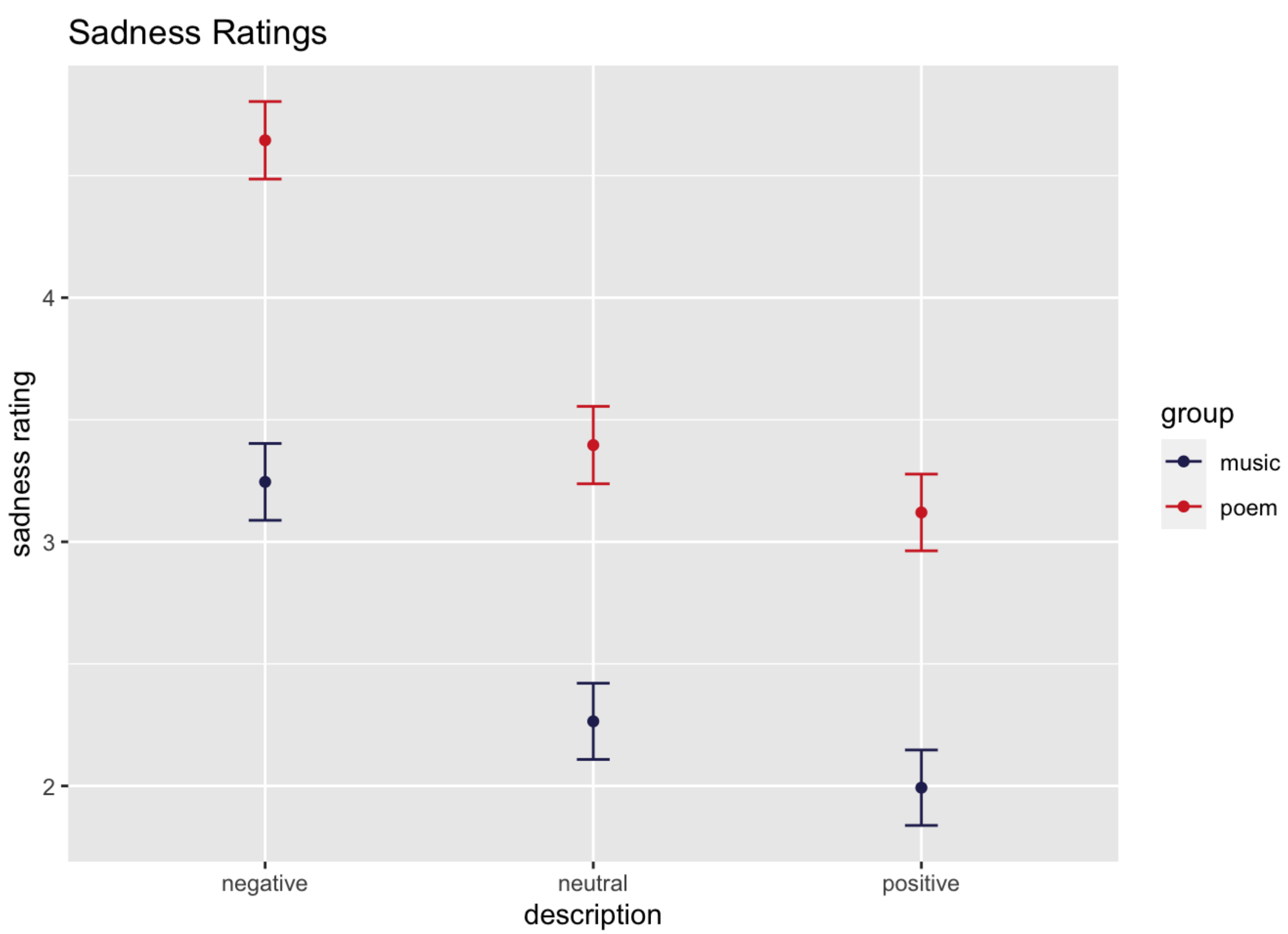
```
sad <- lmer(sad ~ description * stimType + (1+description+stimType|subject) + (1|stimI
D), data=ambiguous)

summary(sad)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: sad ~ description * stimType + (1 + description + stimType |
##     subject) + (1 | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 14161.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6635 -0.6634 -0.0676  0.6170  3.2150
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.3836   0.6194
##           descriptionneutral  0.1688   0.4109   -0.42
##           descriptionpositive 0.2989   0.5467   -0.59  0.98
##           stimTypepoem        0.2373   0.4872   -0.31  0.10  0.08
##  stimID   (Intercept)         0.3507   0.5922
##  Residual                     1.4336   1.1973
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                                   Estimate Std. Error         df t value
## (Intercept)                        3.24542    0.15735   50.42524  20.625
## descriptionneutral                -0.98079    0.07407  325.11911 -13.242
## descriptionpositive               -1.25265    0.08117  243.89369 -15.433
## stimTypepoem                       1.39919    0.21220   42.19172   6.594
## descriptionneutral:stimTypepoem   -0.26762    0.09005 3857.81030  -2.972
## descriptionpositive:stimTypepoem  -0.27177    0.09005 3857.81070  -3.018
##                                 Pr(>|t|)
## (Intercept)                      < 2e-16 ***
## descriptionneutral               < 2e-16 ***
## descriptionpositive              < 2e-16 ***
## stimTypepoem                    5.44e-08 ***
## descriptionneutral:stimTypepoem  0.00298 **
## descriptionpositive:stimTypepoem 0.00256 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn dscrptnp stmTyp dscrptnn:T
## dscrptnntrl -0.252
## dscrptnpstv -0.292  0.646
## stimTypepom -0.668  0.140    0.128
## dscrptnnt:T  0.143 -0.608   -0.277   -0.212
## dscrptnps:T  0.143 -0.304   -0.555   -0.212  0.500
```

```
sad_preds <- as_tibble(ggpredict(sad, terms = c('description', 'stimType')))
sad_preds
```

```
## # A tibble: 6 x 6
##   x          predicted std.error conf.low conf.high group
##   <fct>          <dbl>     <dbl>    <dbl>     <dbl> <fct>
## 1 negative        3.25     0.157     2.94      3.55 music
## 2 negative        4.64     0.159     4.33      4.96 poem
## 3 neutral         2.26     0.156     1.96      2.57 music
## 4 neutral         3.40     0.159     3.09      3.71 poem
## 5 positive        1.99     0.155     1.69      2.30 music
## 6 positive        3.12     0.157     2.81      3.43 poem
```

```
ggplot(sad_preds, aes(x=x, y=predicted, fill=group)) +
  geom_errorbar(
    aes(ymin = predicted-std.error, ymax = predicted+std.error, color = group),
    position = position_dodge(0), width = 0.2
    ) +
  geom_point(aes(color = group), position = position_dodge(0)) +
  scale_color_manual(values = c('#27285C', '#D22B2B')) +
  labs(title='Sadness Ratings', x='description', y='sadness rating')
```



Lastly, Movingness ratings “showed a distinct difference between the way music and poetry were experienced depending on the intention description’s valence (see Fig 4). There was a significant effect of intention description type on movingness, F(2, 57.5) = 16.82, p < .001; both negative (M = 3.73, SE = 0.11) and positive (M = 3.80, SE = 0.11) intentions led to higher movingness ratings than did neutral descriptions (M = 3.52, SE = 0.11). This pattern is qualified by an interaction of intention and excerpt type, F(2, 54.4) = 12.15, p < .001; this interaction reflects that positive intentions increased movingness for music not for poetry, and negative intentions increased movingness for poetry but not for music. Music (M = 3.89, SE = 0.13) was rated as more moving than poetry overall (M = 3.48, SE = 0.13), t(50.2) = 2.48, p = .02." (Margulis et al., 2017)

This figure shows that music seems to be overall more moving than poetry. Music is also most moving when paired with positive descriptions, which is not the case for poetry, which is the most moving when paired with negative descriptions.

Figure 4: Dot plot with SE bars showing mean movingness ratings as a function of intention description and stimulus type

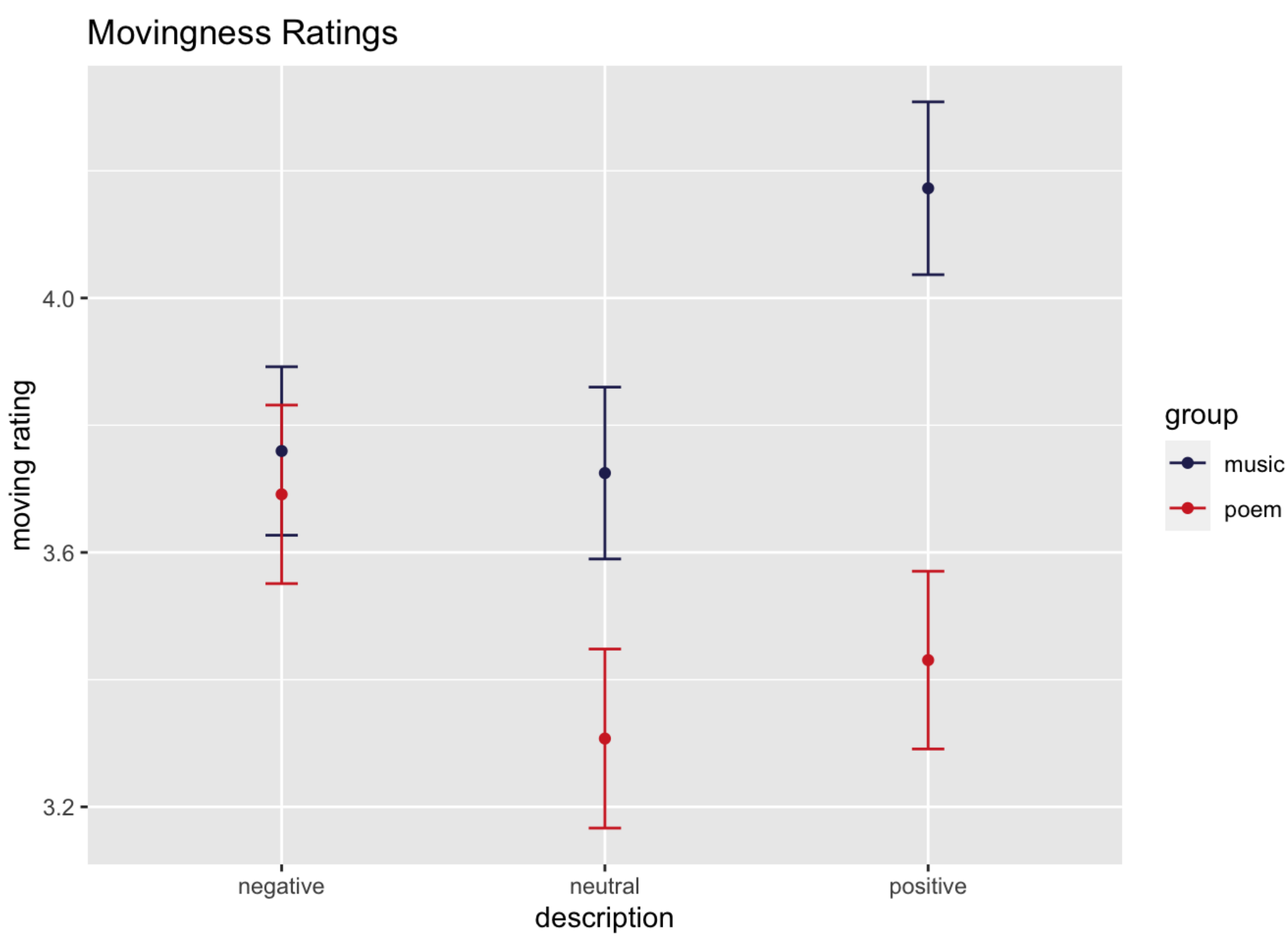
```
moving <- lmer(moving ~ description * stimType + (1+description+stimType|subject) + (1|s
timID), data=ambiguous)

summary(moving)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: moving ~ description * stimType + (1 + description + stimType |
##     subject) + (1 | stimID)
##    Data: ambiguous
##
## REML criterion at convergence: 14246.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4343 -0.6671 -0.0049  0.6577  3.1543
##
## Random effects:
##  Groups   Name                Variance Std.Dev. Corr
##  subject  (Intercept)         0.54161  0.7359
##           descriptionneutral  0.03838  0.1959    0.16
##           descriptionpositive 0.08104  0.2847    0.07  0.37
##           stimTypepoem        0.43177  0.6571   -0.18 -0.28 -0.35
##  stimID   (Intercept)         0.19640  0.4432
##  Residual                     1.44182  1.2008
## Number of obs: 4248, groups:  subject, 118; stimID, 36
##
## Fixed effects:
##                                    Estimate Std. Error         df t value
## (Intercept)                         3.75946    0.13244   72.29242  28.387
## descriptionneutral                 -0.03467    0.06636  394.80819  -0.523
## descriptionpositive                 0.41306    0.06903  350.08067   5.984
## stimTypepoem                       -0.06815    0.17192   54.52804  -0.396
## descriptionneutral:stimTypepoem    -0.34929    0.09030 3742.29601  -3.868
## descriptionpositive:stimTypepoem   -0.67366    0.09030 3742.29606  -7.460
##                                  Pr(>|t|)
## (Intercept)                       < 2e-16 ***
## descriptionneutral               0.601581
## descriptionpositive              5.38e-09 ***
## stimTypepoem                     0.693363
## descriptionneutral:stimTypepoem  0.000112 ***
## descriptionpositive:stimTypepoem 1.07e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dscrptnn dscrptnp stmTyp dscrptnn:T
## dscrptnntrl -0.210
## dscrptnpstv -0.210  0.484
## stimTypepom -0.602  0.152    0.126
## dscrptnnt:T  0.170 -0.680   -0.327   -0.263
## dscrptnps:T  0.170 -0.340   -0.654   -0.263  0.500
```

```
moving_preds <- as_tibble(ggpredict(moving, terms = c('description', 'stimType')))
moving_preds
```

```
## # A tibble: 6 x 6
##   x         predicted std.error conf.low conf.high group
##   <fct>         <dbl>     <dbl>    <dbl>     <dbl> <fct>
## 1 negative       3.76     0.132     3.50      4.02 music
## 2 negative       3.69     0.140     3.42      3.97 poem
## 3 neutral        3.72     0.135     3.46      3.99 music
## 4 neutral        3.31     0.141     3.03      3.58 poem
## 5 positive       4.17     0.136     3.91      4.44 music
## 6 positive       3.43     0.140     3.16      3.70 poem
```

```
ggplot(moving_preds, aes(x=x, y=predicted, fill=group)) +
  geom_errorbar(
    aes(ymin = predicted-std.error, ymax = predicted+std.error, color = group),
    position = position_dodge(0), width = 0.2
    ) +
  geom_point(aes(color = group), position = position_dodge(0)) +
  scale_color_manual(values = c('#27285C', '#D22B2B')) +
  labs(title='Movingness Ratings', x='description', y='moving rating')
```



The authors then tested whether “the interaction of intention description with excerpt type on enjoyment was mediated by the perceived match between ambiguous excerpts and the intention description they were paired with.” by repeating the analyses reported above for enjoyment with match as an additional predictor in the regression model. (Margulis et al., 2017)

They decomposed the interaction of intention description with excerpt type for enjoyment, observing “the unmediated effects of positive and negative intention descriptions relative to neutral descriptions separately for music and poetry.” noting that “relative to neutral intentions descriptions, negative descriptions reduced enjoyment for music but increased enjoyment for poetry”, and controlling for match, the intention description by excerpt interactions for enjoyment is no longer significant, $F(2, 183.5) = 1.82$, $p = .16$. (Margulis et al., 2017)

Conclusion

My replication of this paper returned the same patterns and trends as the authors’ results, albeit with slight numeric differences. From the results of this study, it seems to be the case that verbal information/intent disambiguation about the artist’s intent does indeed influence the way a work of music or poetry is processed. Both musical and poetic paired with positive intent information were perceived as happier, while excerpts paired with negative intent information were perceived as sadder. Additionally, music seems to be correlated more with happiness and positive descriptions, while poetry seems to be correlated with sadness and negative descriptions. This is supported in that positive intent information increased enjoyment and moving ratings for music, while negative intent information increased the same ratings for poetry, and positive intent information was judged to match musical excerpts best while negative intent information was judged to match poetry best.

People also seem to prefer disambiguation to having the ambiguity left intact provided that the description matches the type of piece (poetry with negative intent and music with positive intent), judging from the enjoyment and movingness ratings. Interestingly, ambiguity is preferred to descriptions that don’t match the type of piece (music with negative descriptions and poetry with positive descriptions). Finally, it seems as though expressive disambiguation of the artist’s intent does affect aesthetic appreciation similarly for musical and poetic excerpts for happiness and sadness, which suggest “that empathy with a perceived human artist is indeed an important shared factor across experiences of music and poetry” (Margulis et al., 2017)

References

Margulis, E. H., Levine, W. H., Simchy-Gross, R., & Kroger, C. (2017). Expressive intent, ambiguity, and aesthetic experiences of music and poetry. PLoS One, 12(7), e0179145. https://doi.org/10.1371/journal.pone.0179145 (https://doi.org/10.1371/journal.pone.0179145)