

A Model of Perceived Musical Affect Accurately Predicts Self-Report Ratings

Joshua Albrecht^{*1}

^{*}*School of Music, Ohio State University, USA*

¹albrecht.89@osu.edu

ABSTRACT

A new method of collecting self-report assessments of the perceived affective content of short musical passages is described in Albrecht & Huron (2010). This study used a procedure termed the *progressive exposure method* in which a large passage is divided into discrete five-second excerpts. These excerpts are then presented in random order, and participants evaluate the perceived affective content of these short passages. In that study, 110 participants used the progressive exposure method to analyze the second movement from Beethoven's *Pathétique* sonata. The results from this study provide a mosaic portrait of eleven affective dimensions across the movement. In this study, a model of perceived affective content is built by measuring sixteen different musical features of each excerpt and using these measurements as predictors of participant ratings. This model is used to make predictions of participant evaluations of the same eleven affective dimensions for fifteen excerpts from different Beethoven piano sonatas. To anticipate the results, the predictions for each of the fifteen excerpts along each of the eleven affective dimensions are significantly correlated with participant ratings.

I. INTRODUCTION

In Albrecht & Huron (2010), an affective analysis was carried out on the second movement of Beethoven's *Pathétique* sonata using the progressive exposure method. In that study, 110 participants listened to the entire movement of Beethoven's *Pathétique* sonata in randomly-ordered discrete five-second excerpts and rated the perceived affective content of each excerpt on eleven affective dimensions: calm/serene, carefree, contentment, dark, happy/joyful, lonely, sad/depressed/tragic, striving/yearning, suspense/anticipation, weighty, and unsettled/anxious. The result was a mosaic portrait of each affective dimension as it changes throughout the work. Moreover, direct comparisons were made between the mean ratings across participants for each excerpt and the musical content of each excerpt. This data allows an affective narrative to be constructed as a type of analysis that relates what is happening musically with ratings of perceived affect.

These results are suggestive of two deeper questions: 1) to what extent are some of the measurable musical parameters in each excerpt (e.g., mode, harmonic rhythm, pitch height, presence of dissonance, presence of cadence, dynamics, etc.) contributing to or influencing ratings of perceived affective content in this movement? and 2) how generalizable to other works is the relationship between the perceived affective ratings of the excerpts in the *Pathétique* and the musical content of those excerpts that led to an affective narrative? Both of these questions can be summarized by asking how stable the relationship is between the musical surface of excerpts and the perceived affective content of those excerpts both within the movement tested and across other pieces of music.

The way in which the progressive exposure method presents discrete five-second excerpts limits the influence of

more top-down types of processing in determining the perceived affective content of a short excerpt of music. As a result, the effect on affective ratings for deeper long-term processes like implications and realizations, large-scale harmonic structure, Schenkerian lines, modulation and key areas, or contrasts in sections, would be consequently mitigated. It seems likely that, instead of these more top-down types of features, listeners would evaluate the perceived affective content of short excerpts based on more bottom-up types of musical features, such as register, texture, density, articulations, and dissonance, for example.

It therefore seems appropriate, in building a model of perceived affective content in short excerpts of music, to take into account the types of bottom-up features that a listener may use to make judgments of the perceived affective content of the excerpts. For this study, several low-level analytical observations will be made for each of the excerpts used in the initial study. These observations will be treated as predictor variables in multiple regression analyses for each of the eleven affective dimensions used in Albrecht & Huron (2010). The result is a linear model for each affective dimension that predicts listener ratings of perceived affective content in five-second excerpts in Beethoven's *Pathétique* sonata, second movement.

After the construction of the linear model, a formal study will be described in which the purpose is to investigate the extent to which the model is generalizable to other works. In this follow-up study, fifteen excerpts sampled from the corpus of Beethoven piano sonata movements are analyzed according to the same low-level musical parameters. These parameters are fed into the linear models in order to make predictions of listener ratings of perceived affective content for each affective dimension.

II. CONSTRUCTING THE MODELS

A. Predictor variables

In constructing any linear model, the goal is to be able to account for as much variability as possible in the independent variable by using a set of predictor variables. In building a model of the data from the main study, the independent data to be predicted are the responses given by the participants that rated the perceived affective content of each excerpt. The linear model constructed will attempt to explain as much of this variance as possible using low-level musical features as predictors.

When dealing with human participants, accounting for the variability in responses is a difficult task. There are many potential reasons why participants could offer differing judgments of the affective content of the same excerpt, such as differences in previous exposure, interpretations of the terms used, the mood the participant was in at the time, etc. However, one reason participant responses may vary could be due to differences in the musical excerpts presented to them. For this study, only differences in low-level musical features

between excerpts will be considered explanatory variables to predict participant responses. All of the other sources of variance, though certainly worthy of study, are not of interest in this study and will therefore be considered part of the error term.

There are a number of musical parameters that might influence a listener's rating of the perceived affective content of an excerpt of music. For this study, sixteen different features of an excerpt that may effect listener ratings were identified, shown in Table 1. Most of these musical parameters were calculated by listening to each excerpt, identifying the corresponding boundaries of the excerpt on the score, and visually inspecting the notated segment of the music. This method was sufficient, for example, for counting the number of concurrently attacked notes, determining the direction of the melodic line, or determining whether there were more major or minor harmonies. However, for the two musical parameters that describe how loud the music sounds, the decision was made to analyze the recording of the excerpt rather than the score. Likewise, performers often regularly introduce rubato into their performance, and so the recording provides a more relevant source from which to make judgments about the tempo.

Table 1. The musical features used as predictors for the regression analyses. Most of the variables were categorical, and for the purpose of generalizability to other works, even the two continuous variable (Speed and Harmonic tempo) were discretized into categories.

Parameter	Operationalization
Dynamic	1(<i>ppp</i>) - 8(<i>fff</i>)
Crescendo	-2 (dramatic decrescendo) - 2 (dramatic crescendo)
Density	Maximum # concurrently attacked notes
Speed	# unique onsets: 1 (0-5), 2 (6-10), 3 (11-15), 4 (> 15)
Tempo	1 (Largo) - 5 (Allegro molto)
Articulation	1 (legato) - 3 (staccato)
Direction	-1 (descending melody) - 1 (ascending melody)
Surprise	-1 (only diatonic notes, and all mel. intervals either step-wise or arpeggiated), 1 (non-arpeggiated leap or non-diatonic leap), 0 (diatonic arpeggiation; if chromatic, they are step-wise and resolve to nearest)
Tendency	1 (presence of non-incident tendency tone), 0 (none)
Mode	1 (predominantly major harmonies), -1 (minor harmonies, or suggestive of minor mode)
Dissonance	0 (only major or minor triadic harmonies), 1 (dominant seventh sonority), 2 (any other dissonance)
Harmonic surprise	0 (all harmonic progressions using circle of fifths or thirds, or IV-V), -1 (any other progression)
Harm. temp	# harmonies in the segment
High pitch	Pitch height highest pitch (semita from middle <i>c</i>)
Low pitch	Pitch height of lowest pitch (semita from middle <i>c</i>)
Closure	0 (no cadence), 2 (authentic cadence), 1 (other)

B. Regression analyses

There are many methods for choosing predictor variables for a regression model, and there is little agreement in the field for the superiority of any one method over another. However, it is always a good idea to choose predictor variables thoughtfully, with careful consideration of collinearity, interaction among predictors, and theory.

The first step used in choosing a model for each affective dimension was to perform all-subsets regression. This strategy compares every possible combination of predictor variables for every possible number of predictors with every

other combination, and selects those models for each number of predictors that account for the greatest percentage of the variance in participant responses. In the case of this study, there were sixteen total predictor variables (Table 1). The five best-fitting models were chosen for each number of predictors and two plots were constructed for each of the models: each model was plotted against its adjusted R^2 in one graph and Mallow's C_p in a second graph.

Generally, the greater the number of predictors included in a model, the more variance the model will be able to account for. However, even though most models are chosen to maximize the amount of variance accounted for, there is some effect of diminishing returns; the amount of variance accounted for increases marginally for every added predictor past a certain point. It is therefore important to weigh the advantages of increased variance accounted for against the disadvantages of a model that uses too many predictors and therefore over-fits the data. The visual analyses of the graphs resulted in choosing 6 predictor variables for happy/joyful, lonely, and striving/yearning, 7 predictors for contentment and sad/depressed/tragic, 8 predictors for suspense/anticipation and weighty, and 9 predictors for calm/serene, carefree, dark, and unsettled/anxious. Finally, the best five models for the chosen number of predictor variables were compared, and one model was chosen. The predictor variables for each affective dimension, along with their effects on the affective dimensions and the amount of variance accounted for, are shown in Table 2.

Table 2 shows some interesting results. One interesting observation is the relatively high values for Adjusted R^2 , signifying the amount of variance in participant responses that can be accounted for by measuring only those musical parameters shown in the center column. Given all of the possible sources of error in this study, the high numbers here are encouraging.

In many ways, the predictor variables selected for each affective dimension, and their effect on perceived affective rating, are not surprising and confirm basic intuitions about the perception of affective content in music. Generally, positive affective dimensions are associated with the major mode and negative affective dimensions are associated with the minor mode. Higher arousal affective dimensions, such as unsettled/anxious, happy/joyful, and carefree are associated with higher pitches, whereas lower arousal affective dimensions, such as dark, calm/serene, and weighty are associated with lower pitches. More harmonic dissonance, tendency tones, and less strong or no cadential motion are generally associated with negatively valenced affective dimensions that are typically more intense, such as unsettled/anxious, striving/yearning, and suspense/anticipation.

III. TESTING THE MODELS

In order to investigate the extent to which the linear models are generalizable to other works, a study was conducted to test the predictions made by the models generated from the data from the first study. By measuring the same musical parameters used in the development of the regression models for novel excerpts, predictions can be made about new stimuli. These predictions can be tested by asking new participants to rate the perceived affective content of these new excerpts and comparing the predicted results to the participants' results.

Table 2. The results from the regression for each affective dimension. The rightmost column shows the adjusted R^2 for each affective dimension, or percentage of variance accounted for by the model. In the center column, the complete list of musical features that increase ratings of each affective dimension are given. For example, higher lonely ratings are correlated with slower rhythms, the minor mode, no cadence, fewer concurrently attacked notes, a lower highest pitch in the excerpt and a faster tempo.

Affective Dimension	Regression Parameters	Adjusted R^2
Calm/ Serene	Legato artic.; Major mode; Faster harmonic tempo; Lower high pitch; Diminuendo; Presence of cadence; Common melodic succession; Less harmonic dissonance; Expected melodic progression; No tendency tones	0.362
Carefree	Major; Higher highest pitch; Diminuendo; Faster surface rhythms; Cadence; Legato artic.; Less harm. dissonance; Higher lowest pitch; Faster harm. tempo	0.242
Contentment	Major mode; Legato articulations; Faster harmonic tempo; Cadence; Less dissonant harmonies; No tendency tones; Slower surface rhythms	0.338
Dark	Minor; Lower high pitch; Slower harmonic tempo; Staccato; No cadence; Uncommon harm. progression; More dissonance; Tend. tones; Fast tempo;	0.319
Happy/ Joyful	Major mode; Higher highest pitch; Legato articulations; Presence of cadences; Faster surface rhythm; Decrescendo	0.249
Lonely	Slower surface rhythms; Minor mode; No cadence; Less dense harmonies; Lower highest pitch; Faster tempo	0.313
Sad/ Depressed/ Tragic	Minor mode; No cadence; More common melodic progression; Tendency tones; Lower highest pitch; Faster tempo; Less dense harmonies	0.281
Striving/ Yearning	Minor mode; No cadence; Crescendo; Louder dynamic; Tendency tones; Lower lowest pitch	0.200
Suspense/ Anticipation	Minor mode; Slower harm. tempo; Faster tempo; Tendency tones; No cadence; Staccato articulation; More dissonant harmonies; Denser harmonies	0.283
Unsettled/ Anxious	Higher highest pitch; Crescendo; Minor mode; Tendency tones; No cadence; Staccato articulations; Uncommon melodic succession; More dissonant harmonies; Slower harmonic tempo;	0.353
Weighty	Lower highest pitch; Lower lowest pitch; Crescendo; No cadence; Tendency tones; Louder dynamic; Minor mode; More dense harmonies	0.256

A. Stimuli

The selection of new excerpts to test is an important issue. It seems likely that the effect of different instruments or groups of instruments would influence ratings of the perceived affective content of excerpts. Additionally, it is likely that different composers or different musical styles may utilize different musical features in the expression of the same affect. For example, what musical features are recognized as happy/joyful or lonely for J. S. Bach may be quite different from the musical features indicative of those same affects for Debussy. To reduce the influence of confounding variables, excerpts of solo piano sonatas by Beethoven were chosen.

In order to find excerpts suitable for the follow-up study, professional music theorists on the *Society for Music Theory* listserv and professional piano performers and educators on the *Piano Street* forum were polled to provide excerpts from Beethoven's piano sonatas considered particularly evocative of emotion. The following request was sent out by email to the *Society for Music Theory* "talk" listserv and was posted on the "Repertoire" section of the *Piano Street* web site forum:

"I'm interested in some of the most expressive moments in Beethoven's piano sonatas. What are some of your favorite excerpts that you feel express the most powerful emotions? Any emotion is fine (joy, sorrow, agitation, passion, jealousy, etc.), but ideally it would be an exemplar of that emotion. If you would like to share your favorite moments, please pass on the Sonata number (or Op. number), which movement it is, in what measures it happens, and what emotions you feel that it represents.

For example, I find the opening of the sonata No. 26 (Op. 81a, "*Les adieux*") to be very moving. The C minor triad seems to exude a mournful, resigned spirit after the Eb-G of the first sonority and the Bb-F open fifth almost establish Eb major. Or, I've always found the "E" major modulation in ms. 42-44 in the second movement of the *Pathétique* to exhibit an almost paradoxical joyful, resolved feeling that is yet colored by the agitation and unease expressed by the dense, insistent triplet chordal figures underneath the melody."

In total, there were seven responses from the *Society for Music Theory* and eight responses from *Piano Street*. All responses were accepted for the study, providing that they were limited to the piano sonatas and that there was some indication as to where in the movement the described emotion happened. Several responders on the *Piano Street* forum provided suggestions for Beethoven's piano concertos and some responders simply provided an entire movement or work but did not specify a section. Several responders provided multiple suggestions for excerpts, and these were all accepted.

The recordings were selected from the collection of recordings on *Naxos.com*. The database was searched by entering the composer's name, opus or sonata numbers, or the descriptive terms (such as *Waldstein*) associated with the work. The recording corresponding to the first search result that matched the correct work was chosen as the source for the excerpt. Some of the suggested excerpts spanned much more than five seconds of time for the recording selected. In these cases, a portion of the suggested excerpt was chosen. For suggested excerpts that mentioned specific themes (e.g., "*Moonlight sonata finale, opening theme*"), the beginning of the theme was chosen as the beginning of the excerpt and the excerpt was terminated at five seconds. For suggested excerpts that gave specific measures (e.g., "Op. 10, No. 3 Largo e mesto, ms. 21-25") but that were longer than five seconds, the author listened to the excerpt and tried to select the portion that he felt was most evocative of the affective description given by the recommender. The result was fifteen excerpts, each from a different Beethoven sonata movement, that theorists and pianists had identified as especially evocative of particular powerful emotions. Not all of the excerpts identified were accompanied by a specific description of the emotional content. For example, a respondent might simply exclaim how much they loved a moment, but not offer a specific emotional characterization.

The fifteen excerpts are listed in Table 3, including the movement from which they were taken, the measure numbers of the excerpt, and any descriptions of the emotional content of the excerpts provided. As in the main study, each excerpt was edited with a 500-ms fade-in and fade-out to avoid abrupt onsets and offsets.

Table 3. The fifteen suggested excerpts and the free-response descriptions of affective content provided by the responders to the email/forum post.

Excerpt	Description
1. Tempest sonata, finale, mm. 169-173	“Determination”
2. Moonlight sonata, finale, opening theme, mm. 1-4	“Fury”
3. Hammerklavier sonata, I, mm. 22	“Hope in a dark world, in context”
4. Hammerklavier sonata, II, mm. 1-6	
5. Waldstein sonata, I, opening of 2nd theme, mm. 35-37	“Contentment, nostalgia”
6. Waldstein sonata, III, mm. 240-245	“Heavenly”
7. Appassionata, I, beginning of 2nd theme, mm. 34-36	“Gaining in confidence”
8. Op. 109, finale, Var. VI, mm. 11b-12a	“Transcendental”
9. Op. 10, No. 3, Largo e mesto, ms. 11	
10. Op. 10, No. 3, Largo e mesto, ms. 21	“Anguished grappling with an inevitable and depressing conclusion”
11. Op. 26, I, Variation 3, mm. 7-8	
12. Op. 26, IV, mm. 119-124	
13. Les Adieux, finale, mm. 182-183	“Sobs and tears of joy, but somehow suggestive of the inscrutable quasi-masochistic fallibility of one who has made an accommodation with pain”
14. op. 111, I, m. 3a	“Overcoming the struggles of mortal life and transcending against all odds”
15. Op. 2, No. 1, finale, mm. 34-37	“Like a storm came and then left leaving a smooth, yet still unsettling, reflection upon what just happened”

B. Participants

19 participants were recruited from the Ohio State University subject pool. This was one of several experiments that could be selected by participants in order for them to receive course credit for sophomore-level Aural Skills. The mean age for the participants was 20.5 years (standard deviation = 2.2), and the mean number of years of musical training was 12.5 ($sd = 4.0$). Eight of the participants were female and eleven were male.

C. Procedure

Participants listened to the excerpts in blocks of 16 trials each – they were given one affective dimension and then rated each excerpt for that affective dimension. In addition to the fifteen excerpts, one randomly-selected excerpt was presented at the end of each block, in order to test for intrasubjective reliability. However, due to a randomization glitch, some affective dimensions used more than one repeated trial at the expense of presenting every excerpt to the participant. The order of presentation was randomized for both the affective dimensions and the excerpts within each trial block.

The participants were tested individually in an Industrial Acoustics Corporation sound attenuation room. Participants listened to the stimuli with free-field speakers, adjusted to a comfortable volume. The experimenter read the directions aloud while the participant read along with the printed instructions:

“The purpose of this study is to gather information about music and emotion. At the end of the experiment, I’ll say more about our specific goals.

In this study, you will hear short excerpts from early Romantic piano music and rate the emotional expression of those excerpts. Please do not evaluate your own emotional response to the music, but simply what emotions you think the music is trying to express or convey.

You can listen to each excerpt as many times as you like by clicking on the PLAY button.

Each emotional scale is represented by a line, with the right end of the scale representing the maximum amount of that emotion and the left end representing the minimum amount of the emotion.

The excerpts will be presented in blocks of 16 trials each. You will rate a different emotion for each block of trials.

For each emotion, each excerpt and scale will appear one at a time. You must adjust each scale before moving on to the next excerpt. Once the scale is adjusted, click NEXT EXCERPT and the next rating scale will appear.

For the last rating excerpt of the emotion block, you must click NEXT EXCERPT and then NEXT EMOTION to move on to the next sound excerpt. This experiment will take roughly 45 minutes to complete. Do you have any questions?”

After participants read the instructions, they were given some examples of the difference between perceived and felt emotion. They were reminded to rate only the emotions that they perceived the music to be expressing, rather than any emotions that they felt in response to the excerpts. After answering any questions the participant had, the experimenter observed while the participant attempted four practice trials, two each for two different affective dimensions.

After the portion of the study conducted on the computer participants were given a chart to fill out. The purpose of the chart was to translate the descriptions of the emotional expression of the excerpts given by responders to the email/forum post into the terms that were tested in the study. Specifically, participants were instructed to put a “+” in any box in which the affective term on the top (from this study) was similar to or close in meaning to the description given by the responders on the left, to put a “-” in any box in which the term on the top meant the opposite of the description on the left, and to put a “0” in any box in which the term on the top seemed unrelated to the description on the left.

D. Results

Sixteen musical features were measured in each of the excerpts from the *Pathétique* sonata used in the initial study. These features were used as predictor variables to create a regression model for each of the eleven tested affective dimensions. The same sixteen musical features were measured for each of the fifteen excerpts used in the follow-up study, and these measurements were used in the

regression equations to predict the perceived affective content of those excerpts. The predictions are shown in Table 4.

The variables used to build the model were first normalized by subject-scale, and so were measured in standard deviations away from the mean for that affective dimension for that participant. The values in Table 4 should therefore be interpreted for each affective dimension as standard deviations away from the mean of excerpts in the *Pathétique* sonata, second movement. While the actual numbers presented do not represent some sort of universal value of the amount of each affective dimension per excerpt, the rank-order of the values for the excerpts are nevertheless revealing. For each affective dimension in Table 7.4, the excerpt that was predicted to have the highest and lowest scores are in **bold** and highlighted.

Table 4. The predictions made by the models for each suggested excerpt, measured in standard deviations away from the mean of the *Pathétique*. The excerpts with the highest and lowest predictions for each affective dimension are in **bold** and highlighted.

	Calm/Serene	Carefree	Contentment	Dark	Happy/Joyful	Lonely	Sad/Depressed/Tragic	Striving/Yearning	Suspense/Anticipation	Unsettled/Anxious	Weighty
1. Tempest sonata, finale, mm. 169-173	-0.893	-0.814	-1.101	1.362	-0.551	1.398	1.753	1.022	1.233	1.106	0.577
2. Moonlight sonata, finale, mm. 1-4	-1.472	-1.338	-1.682	1.771	-0.896	0.899	1.453	1.189	2.613	1.823	0.573
3. Hammerklavier sonata, I, mm. 22	-0.360	0.174	0.358	-1.228	0.593	-0.346	-0.785	-0.160	-0.742	0.155	-0.920
4. Hammerklavier sonata, II, mm. 1-6	1.087	0.723	1.221	-0.325	0.856	-0.220	0.352	-0.212	-0.045	-1.334	0.045
5. Waldstein sonata, I, mm. 35-37	1.066	0.749	1.339	-0.095	0.856	0.852	0.529	-0.495	0.097	-1.201	-1.103
6. Waldstein sonata, III, mm. 240-245	-0.916	-0.166	-0.607	0.327	0.461	0.025	0.430	0.266	0.778	0.682	-0.029
7. Appassionata, I, mm. 34-36	0.446	0.606	0.442	0.060	0.380	-0.275	0.213	-0.297	0.294	-0.577	0.286
8. Op. 109, finale, Var. VI, mm. 11b-12a	-0.460	0.318	0.078	-0.090	0.692	-0.669	-0.345	-0.082	0.050	0.019	0.397
9. Op. 10, No. 3, Largo e mesto, ms. 11	-0.782	-0.637	-1.035	0.249	-0.555	0.458	0.551	0.527	0.318	1.109	-0.476
10. Op. 10, No. 3, Largo e mesto, ms. 21	-1.496	-1.321	-1.836	1.130	-0.909	0.209	0.367	0.969	1.134	1.764	0.564
11. Op. 26, I, Variation 3, mm. 7-8	0.132	-0.536	-0.207	0.774	-0.859	0.615	0.775	0.200	-0.269	0.111	0.590
12. Op. 26, IV, mm. 119-124	1.633	1.275	1.622	-0.759	1.165	-0.121	0.281	-0.549	-1.005	-1.864	-0.395
13. Les Adieux, finale, mm. 182-183	-0.063	0.123	0.255	-0.664	0.292	-0.143	-0.123	-0.111	-0.374	-0.032	-0.505
14. Op. 111, I, m. 3a	-2.318	-2.096	-2.183	1.168	-0.933	0.440	0.082	1.225	1.776	2.441	1.081
15. Op. 2, No. 1, finale, mm. 34-37	0.578	-0.056	0.574	0.253	0.136	0.797	1.163	0.306	0.092	-0.551	0.395

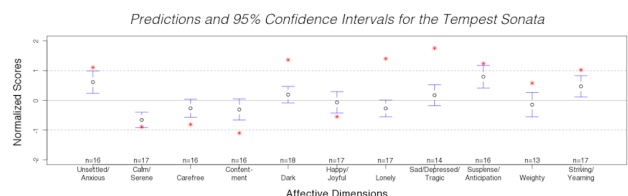
The predictions made by the models were then compared with the actual responses by participants. The participants' raw scores were converted into normalized scores by averaging across subject-scale. This important step was necessary for the comparison of the mean participant ratings in this follow-up study with the models' predictions, measured in normalized z-scores. Despite this corrective step, there is still not a one-to-one correspondence between the meaning of the model's predicted ratings and the normalized scores from the participants in the follow-up study. This is because the predictions are measured in standard deviations away from the mean of ratings of excerpts within the *Pathétique* sonata, whereas the ratings in the follow-up study are measured in standard deviations away from the mean of ratings of excerpts used in the follow-up study. Since the models are based on the *Pathétique*, one should expect their predictions to not be especially accurate when applied to passages from other works – the units of measure are different. However, if the models based on the *Pathétique* are reasonably predictive of the affective content of other passages, then this suggests either that the excerpts tested are similar to the excerpts in the *Pathétique* or that the models capture general aspects of affective musical features.

It is possible to distinguish a strong and weak version of the research hypothesis. The *strong version* takes the predictions as point-estimates of the means of ratings for the tested

excerpts; the hypothesis is that these means are statistically indistinguishable from the population means at an α -level of .95. A statistical test can be performed to determine whether the prediction falls within a 95% confidence interval of the mean ratings in the follow-up study. Despite the fact that the ratings are based on different means and standard deviations, this finding might be seen if participants somehow adjusted their ratings to the context of the non-*Pathétique* excerpts. This finding might suggest that the mean ratings for excerpts in the follow-up study maps onto the mean ratings for excerpts in the *Pathétique* sonata. However, if the predictions fail to fall within the 95% confidence interval around the data, a *weak version* of the hypothesis can be tested; this hypothesis holds that the predictions, though not exactly the population means, are scaled versions of the population means. This can be tested by using a Pearson's correlation test.

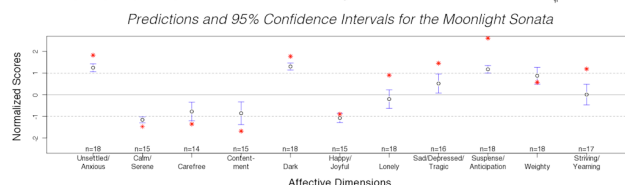
The results for each of the fifteen excerpts are given below. The score for each five-second excerpt is shown, followed by the predictions given by the models for each affective dimension (shown with red stars). The 95% confidence intervals and means are also plotted on the same graph. A sound file for each excerpt is also included. Following each example, a brief discussion of the musical features follows and the results from the correlation test.

1) *Tempest sonata*



This excerpt, taken from the music immediately preceding the retransition, uses arpeggios to outline first a diminished-seventh chord and then a German sixth chord, two harmonies that strongly imply resolution to the dominant. The passage is played loudly and is marked with a crescendo. The affective dimensions that exhibit mean ratings significantly different from zero are unsettled/anxious, suspense/anticipation, and striving/yearning with high ratings and calm/serene with low ratings. The predictions were significantly correlated at $R = +0.59$ ($p > 0.0001$).

2) *Moonlight sonata*

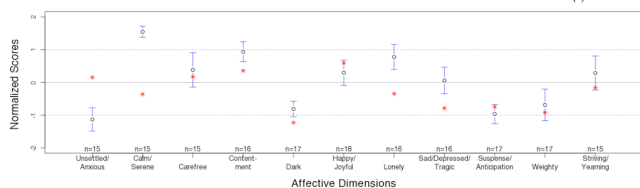


The beginning of the finale of the *Moonlight* sonata is played very quickly and dramatically. Though the harmonic tempo is very slow (only two chords), many note events (56!) happen in this short time span. Every two measures, there is also a very sudden fortissimo. The affective dimensions with the highest mean ratings for this excerpt are dark, unsettled/anxious, and suspense/anticipation, while participants rated the excerpt as low in calm/serene and happy/joyful. The predictions were significantly correlated at $R = +0.73$ ($p > 0.0001$).

3) Hammerklavier sonata (I)

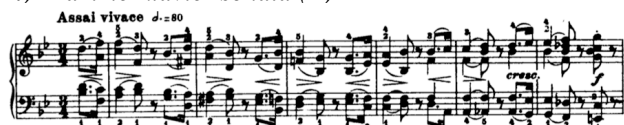


Predictions and 95% Confidence Intervals for the Hammerklavier sonata (I)

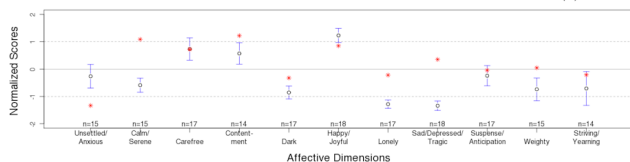


The excerpt taken from the first movement of the *Hammerklavier* sonata is very slow and very quiet. The harmonic rhythm is also very slow, with only a simple G major harmony, played first in root position and then in first inversion. The melody is marked by large leaps – first a major 10th from the grace note to the high *b*, followed by the minor 7th leap from *a* up to *g*, which doubles the lower *g* in octaves. Participant ratings were the highest for calm/serene, contentment, and lonely in this excerpt, while they were the lowest for unsettled/anxious, suspense/anticipation, dark, and weighty. The predictions were significantly correlated at $R = +0.33$ ($p > 0.0001$).

4) Hammerklavier sonata (II)



Predictions and 95% Confidence Intervals for the Hammerklavier sonata (II)

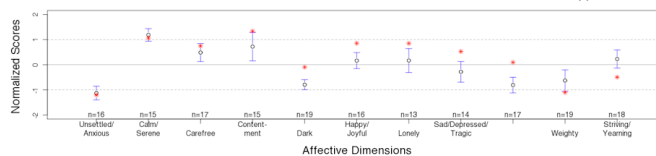


In contrast to the first movement, the second movement of the *Hammerklavier* sonata begins quickly, with rapid harmonic change, jaunty dotted rhythms, a gradual melodic descent, a rapid alternation between crescendo and diminuendo, and a strong perfect authentic cadence to close the excerpt. Participants rated this excerpt as very high in happy/joyful, carefree, and contentment, while rating it very low in sad/depressed/tragic and lonely. The predictions were significantly correlated at $R = +0.37$ ($p > 0.0001$).

5) Waldstein sonata (I)



Predictions and 95% Confidence Intervals for the Waldstein sonata (I)

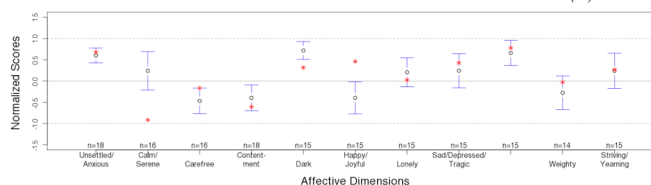


Though it exhibits a very chorale-like texture, the second theme of the *Waldstein* sonata was suggested by two separate responders for its emotionally expressive qualities. The melody and the line in thirds to the melody are doubled at the octave, set against a simple contrapuntal bass line. Like most chorale settings, there is a fast harmonic rhythm (every attack is a new harmony), coupled with a very slow surface rhythm. Perhaps the most striking thing about this short excerpt is the deceptive motion from the dominant of the submediant to the subdominant harmony. This excerpt was rated as very high in calm/serene, contentment, and carefree by participants, and very low in unsettled/anxious, dark, and suspense/anticipation. The predictions were significantly correlated at $R = +0.54$ ($p > 0.0001$).

6) Waldstein sonata (III)



Predictions and 95% Confidence Intervals for the Waldstein sonata (III)



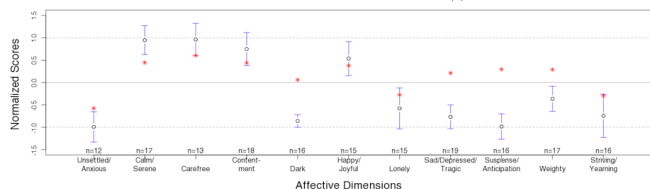
The third movement of the same sonata contains an excerpt with arpeggiated major-seventh chords. These two chords, are both over a pedal *C* in the bass, but the first chord is build on that note as root and the second chord shifts up by half-step so that this same note becomes the seventh of the chord. Participants rated this excerpt as high in dark, unsettled/anxious, and suspense/anticipation, and low in happy/joyful, carefree, and contentment. The predictions were significantly correlated at $R = +0.28$ ($p = 0.0002$).

7) Appassionata sonata

The second theme in the first movement of the *Appassionata* sonata is a low melody, doubled at the octave, and opens up the major triad with arpeggiations. The chordal support, even lower, is muddled even further by a low-level hemiola. The harmonic motion is a simple tonic-dominant- tonic motion, expanding a root position tonic to its first inversion. This excerpt was rated as high in calm/serene and



Predictions and 95% Confidence Intervals for the Appassionata sonata

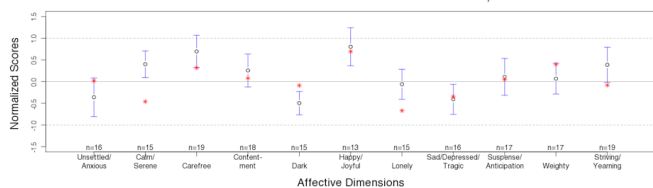


carefree and low in suspense/anticipation and unsettled/anxious. The predictions were significantly correlated at $R = +0.55$ ($p < 0.0001$).

8) Op. 109



Predictions and 95% Confidence Intervals for Op. 109

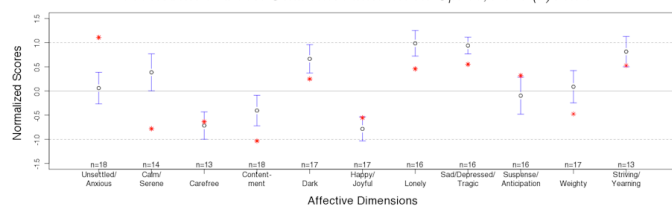


The finale for Beethoven's Op. 109 is a theme and variations, and this excerpt is taken from the sixth variation. The excerpt begins with a weak dominant-function chord in a major mode, notable for the sharp dissonance between the leading-tone and the tonic. There is also a stark rhythmic change that happens at the resolution; the rhythms move from thirty-second notes to the relatively slow triplets. However, this rhythm is contrasted against a low rumbling trill doubled in the right hand. The affective dimensions with the highest ratings for this excerpt are happy/joyful and carefree, and the lowest ratings are dark, sad/depressed/tragic, and unsettled/anxious. The predictions are significantly correlated at $R = +0.24$ ($p = 0.001$).

9) Op. 10, No. 3, Large e mesto, ms. 11



Predictions and 95% Confidence Intervals for Op. 10, No. 3 (II) ms. 11

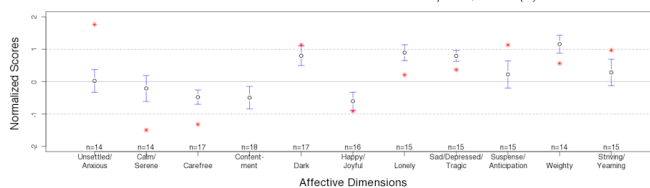


This excerpt outlines a simple tonic-dominant motion in D minor. The bass line simply arpeggiates the harmonies, while the melody is embellished with turns and grace notes. The result is a slow, lyrical melody accompanied with a graceful, thin texture. The highest ratings were for lonely, sad/depressed/tragic, striving/yearning, and dark, whereas the lowest ratings were for happy/joyful and carefree. The predictions were significantly correlated at $R = +0.42$ ($p < 0.0001$).

10) Op. 10, No. 3, Large e mesto, ms. 21



Predictions and 95% Confidence Intervals for Op. 10, No. 3 (II) ms. 21

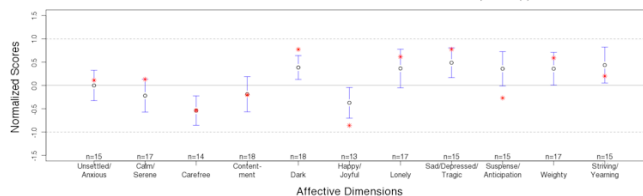


This excerpt, taken from ten measures later in the movement, tonicizes the dominant with a first-inversion dominant-seventh chord. The chord resolves deceptively, however, to a very low F major triad decorated by a half-step ornamental figure of the dominant of the chord. The dynamic of this passage is very loud, and the low register serves to make the timbre much darker. The highest ratings for this excerpt were for weighty, lonely, dark, and sad/depressed/tragic, and the lowest ratings were for happy/joyful, contentment, and carefree. The predictions were significantly correlated at $R = +0.38$ ($p < 0.0001$).

11) Op. 26 (I)



Predictions and 95% Confidence Intervals for Op. 26 (I)

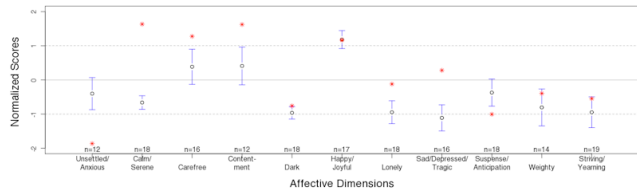


This excerpt, taken from the first movement, third variation, exhibits a transformation of the minor subdominant into the leading-tone fully-diminished seventh of the dominant. This chord resolves to the dominant with a suspension/retardation figure on the downbeat. The highest ratings for this excerpt are sad/depressed/tragic, striving/yearning, suspense/anticipation, weighty, lonely, and dark. The lowest ratings are for carefree and happy/joyful. The predictions were significantly correlated at $R = +0.38$ ($p < 0.0001$).

12) Op. 26 (IV)



Predictions and 95% Confidence Intervals for Op. 26 (IV)

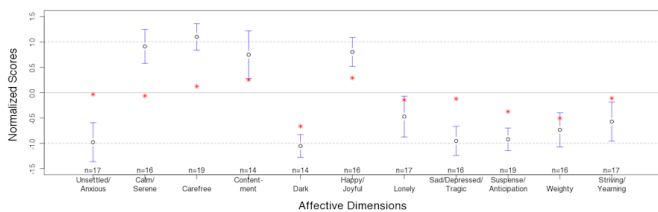


This major-mode passage is very rhythmically active with a walking bass line and a heavily figured melody. There is also a very strong perfect authentic cadence in the middle of the excerpt. The highest ratings for this excerpt were for happy/joyful, carefree, and contentment. The lowest ratings for this excerpt were sad/depressed/tragic, lonely, dark, striving/yearning, and weighty. The predictions were significantly correlated at $R = +0.36$ ($p < 0.0001$).

13) Op. 26 (IV)



Predictions and 95% Confidence Intervals for the Les Adieux sonata

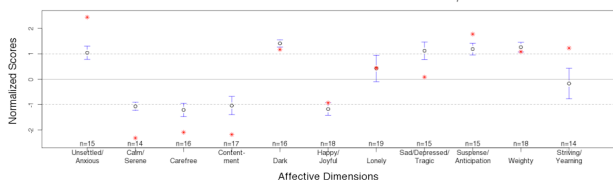


This excerpt (Example 7.14) was one of the two excerpts suggested principally because of the contrast between the excerpt and the surrounding context. After a rhythmically and chromatically dense movement, the tempo and rhythm slow down markedly and simple arpeggios outlining major triads are introduced. This excerpt simply outlines the tonic triad in a slow tempo in contrary thirds, with a crescendo at the end of the excerpt. The highest ratings for this excerpt are for carefree, calm/serene, happy/joyful, and contentment. The lowest ratings are for dark, sad/depressed/tragic, unsettled/anxious, and suspense/anticipation. The predictions were significantly correlated at $R = +0.59$ ($p < 0.0001$).

14) Op. 111



Predictions and 95% Confidence Intervals for Op. 111

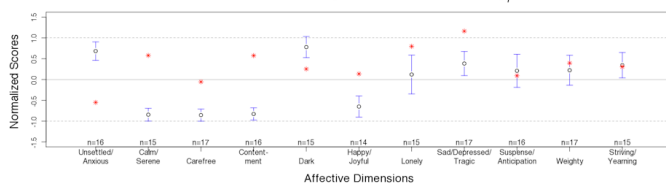


This excerpt comes from the slow introduction to Op. 111. It is marked by loud, sudden attacks, marked fortissimo, and long periods of silence. Moreover, the harmony of the excerpt clearly articulates the dominant of a minor key, outlined by the fully-diminished seventh chord and the diminished-seventh leap in the bass. Unsurprisingly, the highest ratings for this excerpt are for dark, weighty, suspense/anticipation, sad/depressed/tragic, and unsettled/anxious. The lowest ratings are for carefree, happy/joyful, contentment, and calm/serene. The predictions for this excerpt were significantly correlated at $R = +0.55$ ($p < 0.0001$).

15) Op. 111



Predictions and 95% Confidence Intervals for Op. 2



The final excerpt is taken from the final movement of Beethoven's Op. 2, No. 1. This excerpt is characterized by an insistent triplet figuration in the left hand, outlining very low, almost closed-position triads. There is a strong cadential second-inversion tonic-dominant perfect authentic cadence in the middle of this minor-mode excerpt. After this cadence, the right-hand jumps up a tenth and plays a descending line doubled at the octave. This passage had high ratings for dark and unsettled/anxious, and exhibited low ratings for calm/serene, carefree, and contentment. The predictions for this excerpt were significantly correlated at $R = +0.74$ ($p < 0.0001$).

E. Discussion

In the discussion above, each excerpt was tested individually for accuracy in predictions for all eleven affective dimensions. In testing the strong hypothesis, the predictions for some of the excerpts were much more accurate than others. For example, the third movement of the *Waldstein* sonata and the first movement from Op. 26 each had 8 predictions fall in the 95% confidence interval and seven predictions for the Op. 109 excerpt fell within the confidence interval. Additionally, five predictions fell within the confidence interval for both the first movement from the *Hammerklavier* sonata and the *Appassionata* sonata. It may be that these excerpts were more similar to the *Pathétique* sonata, and so the models built from the music in that excerpt may have been better at predicting listener responses.

Additionally, from a visual inspection of the plots above, it seems that the closer the observed responses were to the mean the more accurate the predictions were. This suggests that the more extreme the perceived affective content of an excerpt, the weaker the predictions are. A likely explanation for this effect is that the model is not powerful enough. If the variance in the excerpts tested was larger than the variance in

the excerpts used to build the model, then there would be a loss of accuracy in predictions in ratings further away from the mean.

Even though many of the predictions made by the models did not fall within the 95% confidence intervals of the participant responses, the predictions for each of the fifteen excerpts were significantly correlated with participant ratings. This is consistent with the weak hypothesis that the models are predicting something of the perceived affective content of the excerpts. The differences between the actual predictions and the ratings may be one of scale. Because the predictions were normalized according to different means and standard deviations than the participant ratings, it seems likely that the scales of the numbers were different.

IV. Conclusion

In this study, a model of perceived affective content in short musical excerpts in the *Pathétique* sonata, second movement was built by regressing sixteen musical predictors onto the participant ratings from the main study. Despite the plentiful potential sources of error, the models built were able to account for between 20.0% to 36.2% of the variance in participant responses for excerpts from the movement. A unique model was built for each of the eleven affective dimensions tested in this study.

Following the construction of the models, an experiment was described in which excerpts solicited from music theorists and pianists were tested against the models. A study was conducted to test the applicability of the model to other excerpts from different Beethoven sonata movements. The different affective dimension models were tested to determine if the predictions of the various excerpts along each affective scale were consistent with participant ratings. It was the case that the predictions for all fifteen excerpts along each of the eleven affective dimensions were significantly correlated with participant ratings of those excerpts.

While the models built seem to be generalizable to other excerpts in other movements according to the weak hypothesis, many of the predictions failed to fall within the 95% confidence intervals of the participant ratings. More work is needed to improve the accuracy of the models. A more generalizable model could be built that *begins with* excerpts from a broad range of works. The larger variability in musical features should serve to strengthen the model even further to be more accurate for excerpts with musical features that go beyond the ranges of the *Pathétique* sonata.

REFERENCES

- Albrecht, J. & Huron, D. (2010). An analysis of affective content in the second movement of Beethoven's *Pathétique* sonata using an acoustic ethological model. Seattle, WA: 11th International Conference on Music Perception and Cognition.
- Eerola, T., & Vuoskoski, J. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, 15: 159-173.
- Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, 12, 75-99.
- Francés, R. (1958/1988). *The perception of music* (W. Jay Dowling, Trans.). Hillsdale, NJ; Lawrence Erlbaum Associates.

- Gabrielsson, A., & Juslin, P. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, 24, 68-91.
- Gabrielsson, A., & Lindström Wik, S. (2010). The role of structure in the musical expression of emotions. In P. Juslin, & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications*. New York, NY: Oxford University Press.
- Gjerdingen, R., & Perrot, D. (2008). Scanning the dial: The rapid recognition of musical genres. *Journal of New Music Research*, 37 (2): 93-100.
- Hevner, K. (1935). The affective character of the major and minor modes in music. *American Journal of Psychology*, 47, 103-18.
- Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology*, 49, 621-30.
- Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- Juslin, P., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. Juslin, & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications*. New York, NY: Oxford University Press.
- Kendall, R., & Carterrette, E. (1990). The communication of musical expression. *Music Perception*, 8 (2), 129-164.
- Plazak, J., & Huron, D. (2011). The first three seconds: listener knowledge gained from brief musical excerpts. *Musicae Scientiae*, 15 (1), 29-44.
- Schubert, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology*, 51, 154-65.
- Upham, F., & McAdams, S. (2010). Towards an Audience Model of Listening Experience. Seattle, WA: 11th International Conference on Music Perception and Cognition.

Accompanying Files are found in the Proceedings CD-ROM, in directory /Multimedia/526