Correlations Between Musical Descriptors and Emotions Recognized in Beethoven's *Eroica*

Erika S. Trent,[*1] Emilia Gómez[#2]

[*]*Dept. of Brain and Cognitive Sciences, Massachusetts Institute of Technology, USA*
[#]*Music Technology Group, Universitat Pompeu Fabra, Spain*
[1]etrent@mit.edu, [2]emilia.gomez@upf.edu

ABSTRACT

Investigations on music and emotion have identified broad musical elements that influence emotions recognized by listeners, such as timbre, rhythm, melody, and harmony. Not many studies have studied the correlation between quantifiable musical descriptors and their associated emotions; furthermore, only few studies have focused on how listeners' demographic and musical backgrounds influence the emotion they recognize. In this preliminary study, participants rated how strongly they recognized the six GEMS emotions (transcendence, peacefulness, power, joyful activation, tension, and sadness) while listening to excerpts from Beethoven's Eroica. Musical descriptors (loudness, brightness, noisiness, tempo/rhythm, harmony, and timbre) were also extracted from each excerpt. Results indicate significant correlations between emotional ratings and musical descriptors, notably positive correlations between key clarity and peacefulness/joyful activation ratings, and negative correlations between key clarity and tension/sadness ratings. Key clarity refers to the key strength associated to the best key candidate; as such, these results suggest that listeners recognize positive emotions in music with a straightforward key, whereas listeners recognize negative emotions in music with a less clear sense of key. The second part of the study computed correlations between demographics and emotional ratings, to determine whether people of similar demographic and musical backgrounds recognized similar emotions. The results indicate that naïve listeners (i.e. younger subjects, and subjects with less frequent exposure to classical music) experienced more similar emotions from the same musical excerpts than did other subjects. Our findings contribute to developing a quantitative understanding of how musical descriptors, and listeners' backgrounds, correlate with emotions recognized by listeners.

I. INTRODUCTION

The known relationship between music and emotion has old beginnings, but in-depth research on this relationship has been stagnant until only recently. The stale state of research regarding the influences of music on emotions can be attributed to a few reasons. The main one is perhaps that human emotions are difficult to measure and quantify and thus considered "less than scientific", and adding music—an ephemeral, languageless creation that only exists in time—to that equation makes it an even more challenging topic (Sloboda & Juslin, 2001a).

However, it is undeniable that there is a relationship between these two very human entities. There is anecdotal evidence in support of this strong connection, and recently, empirical evidence as well. In recent decades, musicologists and psychologists alike have taken interest in the relationship between music and emotion. Musical performances have been shown to communicate emotion (Siegwart & Scherer 1995). Music has shown to trigger basic physiological responses such as heart rate or blood pressure (Krumhansl, 1997, Nyklicek, Thayer, & Van Doornen, 1997), as well as more extreme physical responses such as shivers and goose bumps (Gabrielsson, 2001). Some researchers have, as consequence of such findings, proposed theories that music can either represent an emotion to a listener.

How does music convey emotion? How can a listener "recognize" an emotion from a sequence of notes? Furthering our understanding of these questions would improve current applications, such as music services that generate playlists depending a user's mood, music therapy sessions where the therapist chooses a song depending on the patient's emotional state, or automated programs that could aid musicologists in analyzing the emotional content of a piece.

However, given the intangible and relatively abstract nature of both music and human emotion, research in this field is understandably populated with some problematic issues. Problems that previous researchers have faced include how to categorize emotions, how to quantify them, which emotional descriptors to choose in the first place, the choice of music selections, and how to quantify the "elements" that make up music. We will first look to previous literature in the field of music and emotion to understand the state of the art, which will offer us guidance to how to move forward in answering the question of how music conveys emotions.

II. LITERATURE REVIEW

Previous research on the relationship between emotion and music has, first and foremost, yielded evidence that listeners can experience emotions while listening to music. Such emotions that are perceived from music are believed to have behavioural, physiological, and experiential components. Hence, the three main types of evidence behind these research studies are self-reports, expressive behaviour, and physiological behaviour

Pike (1972), for example, conducted a phenomenological analysis of music experience, in which musically untrained participants freely wrote down responses towards excerpts of music. All the collected responses could be reduced to a limited set of experiences: stable moods, transient emotions, feelings of pleasure, feeling of "oneness with the music", and feeling of movement. Following this study were several more studies that offered evidence, through self-reports, that listeners experience emotional responses to music (Sloboda & Juslin, 2001b).

The idea of "musical tension", as a subset of emotional and partially physical responses to music, also became of

interest for researchers. The occurrences of "chills" or "shivers-down-the-spine" as a result of music inducing a pleasurable experience (Sloboda, 1991; Panksepp, 1995; Blood & Zatorre, 2001) have been reported in laboratory settings. Such intense emotional responses have also been recorded in live concert settings (Grewe, Kopiez & Altenmuller, 2009). Furthermore, researchers such as Farbood (2006) have identified musical elements that contribute to listeners' perception of "musical tension".

Once the belief that music can express emotion was established, researchers experimented with various methods to categorize music-related emotions. Before diving into a deep discussion on this, it is important to recognize the distinctions between certain terms that are often considered synonyms in their everyday uses—but entail critical differences in this context.

The first distinction is between "perceived" emotion and "experienced" (or "induced") emotion. "Perceived" emotion can be understood as "emotion in music", which is the emotion the composer intends to express, and that the listener recognizes. "Induced" emotion can be understood as "emotion from music", which is the emotion the listener feels while listening to a piece of music (Juslin & Laukka, 2004). Perceived emotions are more agreed upon, whereas induced emotions tend to differ more widely between individuals. The present study focuses on perceived emotion.

The second distinction is between the terms, "affect", "mood", and "emotion". "Affect" is a general term that refers to the positive or negative valence of an emotional experience (Oatley & Jenkins, 1996, as cited in Laurier, 2011). A "mood" is a long-lasting experience that is without an identifiable stimulus event, whereas an "emotion" is a briefly lasting experience with an identifiable stimulus event. Some suggest that emotions, unlike moods, are associated with facial expressions (Ekman & Davidson, 1994, as cited in Juslin & Sloboda, 2001b). The present study focuses on brief affective experiences triggered by short musical excerpts, thus "emotions" recognized in music.

Researchers have taken one of three main approaches to conceptualize emotions in relation to music: the categorical approach, the dimensional approach, and the prototype approach (Juslin & Sloboda, 2001b). The categorical approach suggests that people experience emotions as one of several distinct and recognizable categories. Early categorical studies such as Hevner (1936) initially adopted this approach with a list of adjectives, which has been revised since then (Schubert, 2003). One of the most recent categorized models of music-related emotions is the Geneva Emotional Music Scale (Zentner, Grandjean, & Scherer, 2008). The GEMS is a 9-factorial model of music-related emotions developed as the result of four interrelated studies that compiled and categorized music-relevant emotion terms generated by hundreds of listeners. The model is domain-specific to music; Zentner et al.'s (2008) fourth study demonstrated that the GEMS accounts for music-related emotions better than other basic dimensional models. The nine main GEMS emotions are "transcendence", "wonder", "joyful activation", "power", "tension", "sadness", "tenderness", "nostalgia" and "peacefulness" (Zenter et al., 2008). A subset of these nine emotions was used for the present study.

Secondly, the dimensional approach identifies emotions according to their location on a 1-dimensional to 3-dimensional scale of dimensions such as valence, activity, and potency. While there is some debate as to what the third dimension should be in the 3-dimensional case, this approach has enabled listeners to track their changing emotional responses to music in real-time, in a continuous manner ("continuous" measurement is discussed in the next paragraph). The most widely adapted valence-arousal model has been confirmed to be a valid metric in several studies (Russel, 1980; Laurier, Serra, & Herrera, 2009). Thirdly, the prototype approach is built on the idea that emotions are categorical, but have hierarchical relationships to one another (Rosch, 1978).

In addition to the various approaches to conceptualizing emotion, researchers have developed various methods to attempt to quantify and measure these emotions perceived through music. One common distinction between methods is whether the measurement is "continuous" or "discrete". In "continuous" measurements, listeners continuously adjust their emotional response in real-time while listening to the music. Some examples are by moving a knob vertically (in the case of a 1-dimensional approach, such as rating the strength of one emotion from "weak-strong", or rating the musical tension from "weak-strong"), or dragging a cursor on a computer screen (in the case of a 2-dimensional approach, such as "valence-arousal"). A "discrete" measurement, on the other hand, would involve a listener giving their response to a musical excerpt as a whole, without the time factor. Some researchers have argued that adopting a continuous, or "dynamic", approach is important for emotion recognition in music, because music itself continuously changes through time (Kim et al., 2010).

III. OBJECTIVES/MOTIVATIONS

Previous studies have suggested that certain musical parameters especially influence the content of emotional responses—notably timbre, orchestration, acoustics, rhythm, melody, harmony, and structure (Juslin & Laukka, 2004). Several studies have created mappings between musical descriptors and emotion categories (Laurier, 2011), but these emotion categories are limited to the five emotions based on the "Big Five Inventory": "happiness", "sadness", "anger", "fear", and "tenderness" (e.g. Eerola & Vuoskoski, 2011).

Secondly, not many studies have focused on the differences in demographic and musical background in how listeners may experience different emotional interpretations to the same music. While there do exist several cross-cultural studies on music and perceived emotion (e.g. Balkwill & Thompson, Fritz et al., 2009), these studies tend to focus on greatly different cultures, rather than on more subtle differences such as age, gender, and musical experience or exposure.

Thus the present study attempts identify the correlation between quantifiable musical descriptors and the emotions recognized in listeners, and identifying the influences of demographic and musical backgrounds of listeners in how they recognize emotion. The focus of the present study is on classical music, and specifically on the same musical pieces, namely Beethoven's Third Symphony. Our rationale for focusing on classical music is we can analyze the influence of listeners' level familiarity with the piece or familiarity with

the musical style on the recognized emotion. Furthermore, by studying emotional responses to the same musical piece, we may accurately analyze the evolution of emotion along time, as features such as instrumentation, recording conditions, and musical style are all consistent. This approach is different from previous emotion studies on popular music or rock music (Laurier, 2011, p. 57).

Moreover, in the present study we integrate descriptors from both the audio and score. This approach is also novel with respect to the state of the art, such as the MIREX mood estimation (Hu, Downie, Laurier, Bay & Ehmann, 2008) which have generally focused exclusively either on the score or audio. While some studies have also studied song lyrics to detect the emotional content of songs (Hu, Chen, & Yang, 2009), this type of information is irrelevant for classical music, which unlike popular music, has no lyrics.

Gaining a further understanding of the above would have practical applications such as improved music recommendation services or playlist generators and improved automatic musical categorization, specifically for classical music. Currently, classical music is not very well covered by commercial applications, which focus on mainstream popular music. Improved applications could be combined with pre-existing mood estimators submitted to initiatives such as the Music Information Retrieval Evaluation eXchange (IMIRSEL, 2012). Furthermore, research on demographic and cultural influences on recognized emotions could also open the gateway to further research on individual differences in emotional responses to the same music.

This study was conducted in the context of the PHENICX project (Performances as Highly Enriched and Interactive Concert Experiences), whose goal is to make use of music information retrieval technologies to enrich classical music concerts, specifically in the symphonic repertoire (see http://phenicx.upf.edu for further information)

IV. METHODOLOGY

A. Materials and Online Survey

Fifteen excerpts (11 - 29 sec) were selected from Beethoven's Third Symphony, the *Eroica*. This symphony was chosen because, from virtue of being the focus of the PHENICX research project, scores, high quality wav files, and aligned MIDI files were available. The excerpts were selected by the authors, who are trained in music theory and performance, then reviewed by a musicologist in the PHENICX team. In order to select the excerpts, every section of the *Eroica* was labelled with one of the nine GEMS emotions, judged based on musical elements (tempo, rhythm, harmony, melody, and orchestration). The six emotions that most frequently appeared were transcendence, peacefulness, power, joyful activation, tension, and sadness. Criteria for selecting excerpts were that it contained a variety of musical characteristics, lasted the duration of a complete musical phrase, and strongly represented one of the above six emotions. Fifteen was decided to be an appropriate number of excerpts to ensure the subjects remained engaged throughout the study.

An online survey using Google Forms was created. The survey had two sections, a demographic information component and a listening component. The demographic information asked was alias, age, gender, country of origin, years of musical education, how often the subject listens to classical music, and how familiar the subject is with the *Eroica*.

In the listening section, subjects listened to an excerpt and rated how strongly they felt each of the six chosen GEMS emotions on a scale of 0 ("not at all") to 3 ("strongly"). These emotion categories were explained to subjects through lists of adjectives as shown in Table 1. Subjects had the option to comment on the excerpt. Subjects repeated this procedure for 15 excerpts, and had the option to give general comments at the end.

Table 1. Adjectives corresponding to each GEMS emotion, as explained to subjects while they took the online study.

Emotion	Adjectives
Transcendence:	"I feel... fascinated, overwhelmed, inspired, chills, feeling of spirituality"
Peacefulness:	"I feel... serene, calm, soothed, meditative, relaxed"
Power:	"I feel... strong, energetic, triumphant, fiery, heroic"
Joyful activation:	"I feel... animated, bouncy, joyful, dancing, amused, stimulated"
Tension:	"I feel... tense, agitated, irritated, nervous, impatient"
Sadness:	"I feel... sad, tearful, sorrowful"

To prevent ordering effects, three versions of the survey were created, each with their own order. Each order was decided using a random number arranger in MATLAB. A link given to subjects linked to one of the three versions at random.

A Spanish version of the survey was also available for subjects who felt more comfortable in their native language. Translation was assisted by a native speaker of Spanish.

B. Participants

The survey link was sent to the Universitat Pompeu Fabra Music Technology Group mailing list, a Barcelona-based gospel choir mailing list, and was posted on several social networking sites. The survey link was active for a week. 26 participants answered the survey (14 female), of average age 36.8 (standard deviation 12.8). 13 participants were of Spanish origin, 5 from Japan, 4 USA, 2 France, 1 India and 1 South Korea. Average years of musical education was 6.3 years (standard deviation 5.1). 5 participants reported they listen to classical music "almost every day", 10 reported "a few times a month", 3 reported "a few times a year", and 4 reported "almost never". 4 participants were "very familiar" with the *Eroica* (listened to it more than 3 times), 11 were "somewhat familiar" with it (listened one to three times), and 11 had never heard the *Eroica* before.

C. Musical Descriptor Extraction

Thirteen musical descriptors were extracted from each excerpt using the MIR Toolbox (Table 2).

Table 2. Musical descriptors extracted from excerpts, organized by type.

Type	Features
Loudness	mean root mean square (RMS), RMS standard deviation, low energy rate
Brightness	brightness, spectral centroid
Noisiness	zero crossing rate
Tempo/Rhythm	mean tempo, tempo standard deviation, number of note onsets/sec
Harmony	modality, key clarity
Timbre	MFCC

V. RESULTS

A. Emotional Ratings of Excerpts

To begin, the mean and standard deviations of the subjects' rankings, for each emotion, were computed for each excerpt. The results are laid out in Table 3 and illustrated in Figure 1.

B. Correlations

1) Correlations between emotions. First, correlations between emotional ratings were computed (Pearson correlation, critical value 0.388 at n=26 and df=25 and two-sided level of significance of 0.05). Statistically significant negative correlations were found between peacefulness and transcendence, power and peacefulness, joyful activation and transcendence, tension and peacefulness, tension and joyful activation, sadness and peacefulness, and sadness and joyful activation. Statistically significant positive correlations were found between power and transcendence, joyful activation and peacefulness, tension and transcendence, tension and power, and sadness and tension. Full table of correlations is illustrated in Table 4.

2) Correlations between emotions and musical descriptors. Thirteen musical descriptors (as outlined in Table 2) were extracted and computed from the fifteen extracts using MATLAB MIR Toolbox. The correlations between subject-generated emotional ratings and the values of these musical descriptors was computed (Table 5).

Statistically significant findings were as follows.

- Ratings of transcendence correlated significantly with: RMS mean, RMS standard deviation, low energy rate, MFCC (4, 6, 10, 13)
- Ratings of peacefulness correlated significantly with: Key clarity, MFCC (4, 9, 10, 11, 12)
- Ratings of power correlated significantly with: Low energy rate, MFCC (12)
- Ratings of joyful activation correlated significantly with: Key clarity, MFCC (4, 6, 10, 12, 13)
- Ratings of tension correlated significantly with: Modality, key clarity, MFCC (4, 9, 11, 12)
- Ratings of sadness correlated significantly with: Key clarity, MFCC (7, 9, 11, 12)

3) Variations between all subjects. Another question we wanted to explore in this study was how musical experience and demographic background influence how listeners emotionally react to the same piece of music. To take a look at this question, we computed the standard deviations of subjects' emotional ratings in different contexts.

4) Variations between subjects with varying musical experience. Another point of interest was whether subjects with different musical experiences or demographic backgrounds reacted differently to the musical excerpts. Thus, standard deviations between subjects with varying musical experience were compared. First, the standard deviations of emotional ratings, for each emotion and for each excerpt, were split into two or three groups, depending on the criteria (Table 6). Then, whether there was a statistically significant

Table 3. Mean and standard deviations of subjects' ratings of each excerpt, for each of the six emotions. Subjects rated how strongly they felt each emotion, from a scale of 0 ("not at all") to 3 ("strongly").

Excerpt	Transcendence		Peacefulness		Power		Joyful Activation		Tension		Sadness		Highest Ranked Emotion
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
1	0.962	0.958	1.500	1.175	1.231	1.107	1.692	1.087	0.500	0.906	0.269	0.604	Joyful A.
2	1.154	0.881	1.115	1.143	1.654	1.129	1.154	0.881	1.077	1.164	0.615	0.852	Power
3	1.231	1.210	0.808	1.132	2.038	1.038	1.000	0.748	1.385	0.941	0.308	0.679	Power
4	1.115	0.864	0.615	0.804	1.423	1.270	0.808	0.939	1.077	0.977	0.731	0.919	Power
5	1.115	0.909	0.885	1.107	1.538	1.272	1.154	0.925	1.154	0.967	0.577	0.945	Power
6	0.923	1.055	1.500	1.175	1.192	0.849	1.885	0.952	0.808	0.801	0.346	0.629	Joyful A.
7	1.192	1.167	0.269	0.533	2.038	0.774	1.269	1.002	1.615	0.941	0.500	0.707	Power
8	0.923	0.891	1.269	1.151	1.192	1.059	1.308	1.123	0.923	0.977	0.308	0.618	Joyful A.
9	0.808	1.059	0.538	0.706	1.577	0.945	1.654	1.231	1.115	1.033	0.423	0.902	Joyful A.
10	1.269	1.218	1.115	1.275	1.538	1.208	1.269	1.218	0.808	0.801	0.269	0.533	Power
11	1.115	1.033	0.462	0.948	1.308	0.970	0.385	0.496	1.808	0.895	1.000	1.020	Tension
12	0.846	0.834	1.154	1.156	1.000	1.131	0.808	0.895	1.000	1.200	0.808	1.059	Peaceful.
13	1.231	0.908	0.500	0.906	1.577	1.137	0.538	0.706	1.846	1.120	0.885	0.909	Tension
14	1.115	1.033	0.385	0.697	1.885	0.909	0.692	1.050	1.962	0.958	0.885	1.107	Tension
15	1.077	0.935	1.000	1.131	1.423	1.027	1.154	1.156	1.346	1.093	0.462	0.706	Power
Avg	1.072	0.997	0.874	1.003	1.508	1.055	1.118	0.961	1.228	0.985	0.559	0.813	

Table 4. Correlation coefficients between ratings of emotions. Statistically significant values are bolded. Using the Pearson r table for critical values, a correlation coefficient of over absolute value of 0.388 was considered statistically significant (n=26, df=24, and two-sided level of significance = 0.05).

	Transc.	Peaceful	Power	Joyful A.	Tension	Sadness
Transc.		**-0.392**	**0.644**	**-0.459**	**0.452**	0.145
Peaceful.	**-0.392**		**-0.632**	**0.593**	**-0.847**	**-0.587**
Power	**0.644**	**-0.632**		-0.146	**0.543**	-0.021
Joyful A.	**-0.459**	**0.593**	-0.146		**-0.733**	**-0.842**
Tension	**0.452**	**-0.847**	**0.543**	**-0.733**		**0.702**
Sadness	0.145	**-0.587**	-0.021	**-0.842**	**0.702**	

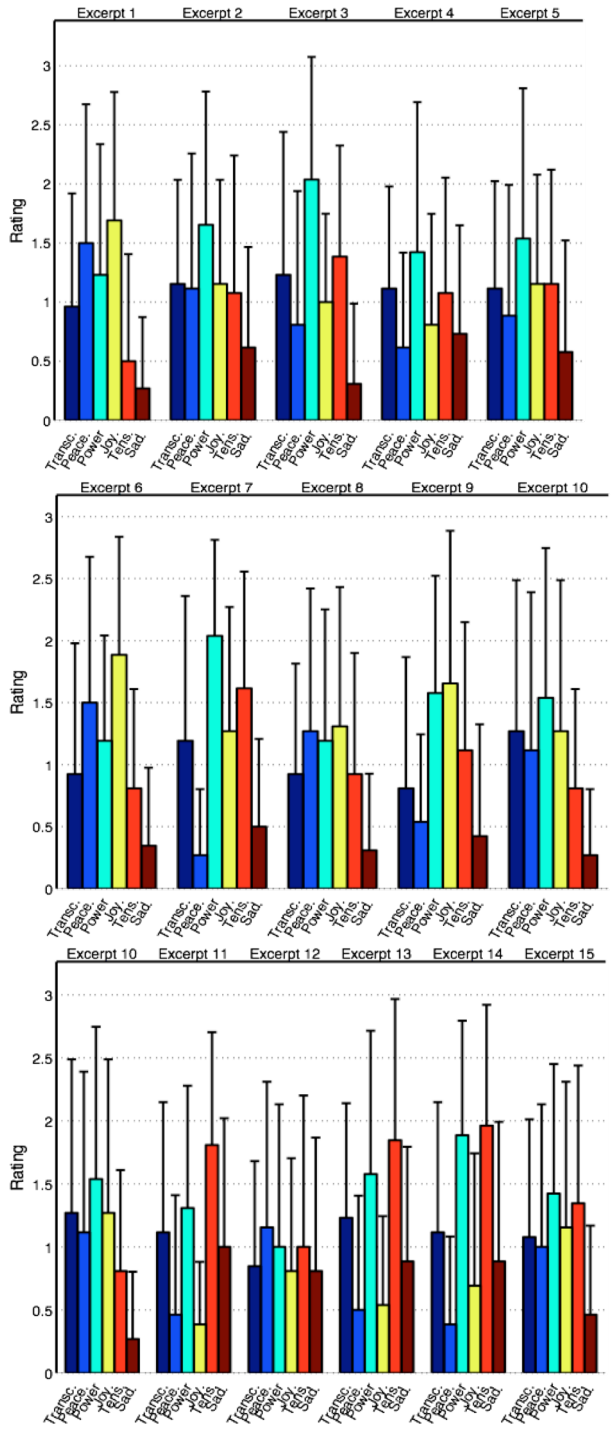

Figure 1. Mean and standard deviations in all subjects' ratings of six emotions, for each of fifteen excerpts.

difference between groups was determined using two samples t-tests (for groups of two) or a one-way ANOVA (for groups of three).

The three criteria in assessing levels of musical experience were years of musical education (Group 1: 6 years or less, Group 2: 7 years or more; the average years of musical education across all subjects was 6.3), level of exposure to classical music (Group 1: "Almost never" and "A few times a year", Group 2: "A few times a month" and "Almost every day"), and familiarity with the *Eroica* (Group 1: "Never heard", Group 2: "Somewhat", Group 3: "Very familiar").

There was no significant difference in standard deviations between subjects who had 6 or fewer years of musical education and subjects who had 7 or more years of musical education (two sample t-test, p=0.1592). Neither was there a significant difference between subjects who were very, somewhat, or not at all familiar with the *Eroica* (one way ANOVA F(2,267)=0.56, p=0.5726).

There was, however, a significant difference between standard deviations between subjects who listened to classical music frequently and subjects who listened to classical music infrequently (two sample t-test, p=0.02). Subjects who listened to classical music "almost never" or only "a few times a year" showed a smaller average standard deviation in their emotional ratings, compared to subjects who listened to classical music "a few times a month" or "almost every day".

5) Variations between subjects of varying demographic backgrounds. The second point of interest was whether listeners had different emotional responses to the same music depending on their demographic backgrounds. We looked at standard deviations between subjects of different countries of origin (Spain, USA, and Japan), gender, and age (Group 1: 36 years or younger, Group 2: 37 years of older; the average age was 36.8). We split standard deviations of the ratings for each emotion, for each excerpt, into groups of two or three, depending on the criteria (Table 7).

No statistically significant difference was found between subjects of different countries (one-way ANOVA F(2,267)=0.62, p=0.5376). Neither was any statistically significant difference was found between male and female subjects (two sample t-test, significance level 0.2922). However, a significant difference was found between subjects of different age groups (two sample t-test, significance level 0.0228). Subjects of age 36 or younger showed smaller

Table 5. Description Correlation coefficients Table 6. Description Correlation coefficients between emotional ratings and musical descriptor values. Statistically significant values are bolded. Using the Pearson r table for critical values, a correlation coefficient of over absolute value of 0.388 was considered statistically significant (n=26, df=24, and two-sided level of significance = 0.05)

	RMS mean	RMS standard deviation	Low energy rate	Bright-ness	Spectral centroid	Zero crossing rate	Tempo mean	Tempo SD	Note onsets /sec	Modality	Key clarity
Transc.	**0.447**	**0.461**	**-0.504**	0.095	0.034	0.146	0.208	0.247	0.218	0.246	-0.223
Peace.	-0.087	-0.090	0.122	0.218	0.303	0.190	-0.018	0.382	-0.343	-0.253	**0.437**
Power	0.177	0.076	**-0.405**	-0.096	-0.148	0.041	0.092	-0.197	0.344	0.115	0.043
Joyful A.	0.024	-0.110	-0.040	0.197	0.334	0.196	-0.207	0.041	-0.081	-0.354	**0.459**
Tension	0.006	-0.073	0.103	-0.054	-0.175	0.004	-0.031	-0.320	0.158	**0.440**	**-0.497**
Sadness	0.048	0.107	0.204	-0.035	-0.173	-0.048	0.165	-0.315	0.097	0.299	**-0.620**

Table 6. Mean standard deviations for each emotion (averaged over fifteen excerpts, and averaged over subjects who qualified for the criteria) for various criteria regarding musical experience and exposure.

Emotion	All subjects	Years of music education		Exposure to classical music		Familiarity with the *Eroica*		
		6 or less	7 or more	Infrequent	Frequent	Never heard	Somewhat familiar	Very familiar
Trans	0.997	0.953	1.065	0.923	1.053	0.968	0.925	1.013
Peace	1.003	1.003	1.010	0.977	1.013	0.923	0.957	1.079
Power	1.055	1.076	1.041	1.062	1.057	1.126	0.948	1.162
Joyful	0.961	0.922	1.004	0.937	0.972	0.911	0.968	0.922
Tension	0.985	1.019	0.955	0.942	1.021	0.991	0.982	0.981
Sadness	0.813	0.694	0.886	0.699	0.880	0.621	0.939	0.697
Av.	0.969	0.944	0.993	0.923	0.999	0.923	0.953	0.976

Table 7. Mean standard deviations for each emotion (averaged over fifteen excerpts, and averaged over subjects who qualified for the criteria) for various criteria regarding demographic and cultural background.

		Country of origin			Gender		Age	
Emotion	All subjects	Spain	USA	Japan	Male	Female	36 or less	37 or more
Trans.	0.997	1.028	0.882	1.084	0.9760	1.0210	1.020	0.965
Peace	1.003	1.008	1.017	0.984	1.0294	0.9888	0.964	0.933
Power	1.055	1.089	1.064	1.018	1.0397	1.0804	1.053	0.986
Joyful	0.961	0.940	1.146	1.024	0.8959	0.9985	0.810	1.016
Tension	0.985	0.998	1.076	0.921	1.0025	0.9715	0.854	1.055
Sadness	0.813	0.642	0.536	0.989	0.7521	0.8439	0.666	0.876
Av.	0.969	0.951	0.953	1.003	0.9493	0.9840	0.895	0.972

standard deviations (i.e. more agreement on what emotions they felt) than subjects of age 37 or older.

C. Participant Comments

In the online survey, a text box was included beneath each excerpt to allow subjects to include comments on the particular excerpt to which they just recorded their emotional responses. The comments section was optional. Table 8 is a summary of all the comments received for each excerpt. The comments are divided into those not regarding emotion (e.g. regarding the length of the excerpt), regarding emotion, and "reasoning" (user-given reasoning or explanations for why they recognized certain emotions, often based upon the musical content or instrumentation of the excerpt).

A common complaint was that several exceprts were too short. To ensure that subjects remained alert throughout all fifteen exceprts, excerpts were kept to under thirty seconds; however, this feedback indicates that longer excerpts may be necessary to sufficiently recognize an emotion. Another common comment was that the listening environment influenced listeners' ability to perceive emotions. For example, listening to music on low volume through headphones provides a weaker emotional experience than listening to a live performance in a concert hall.

Participants seemed to associate diminished chords with transcendence, quiet dynamics with peacefulness, and brass instruments with power, to name a few examples. However, it is evident that participants also frequently disagree with one another. For example, one participant described an excerpt as "suitable for a holiday morning" suggesting peacefulness, whereas another perceived tension. What one participant described as "peaceful," another participant said conveyed "a

Table 7. Comments regarding each excerpt, given by subjects. If more than one person wrote an identical comment, (such as "the excerpt was too short") the number of people who gave that comment is notated in parentheses.

Ex	Not regarding emotion	Regarding emotion	Reasoning
1		"feels like resolving to something, but not sure what," "nicer than previous one (subject had commented about "tension" in previous excerpt)"	
2		"slightly sad but somehow gives power," "feels like something's not quite right," "nice but not strong enough to grab me"	
3	"too short"	"didn't do much emotionally," "imagined being a strained person, depressed, not knowing what to do," "liked this"	
4	"too short"	"powerful but not the single most powerful excerpt"	
5		"calming," "beautiful," "floating along all emotions"	
6	"too long" "too short" (wanted to see where leading)	"peaceful," "liked it," "some parts seem different from others emotionally; mainly joyful, but at the end some tension," "dynamic," "sense of urgency/hurried"	"gradual crescendo, playfulness of half-step and stress-release motif"
7	"too short" (3)	"strong enough to be emotionally influenced", "very strong emotion," "felt tense"	
8		"enjoyable," "nicely energetic," "pretty"	
9		"mixed signals, no one clear emotion"	
10	"too short"	"joyful, happy" (2), "peaceful, pretty," "playful with a hint of an unknown undercurrent"	
11	"low volume" (2)	"perhaps sadness, not sure," "felt sadness and tension, but felt as if had taken courage to confront the problem; I felt more powerful and determined"	
12	"may have felt transcendence if hearing in a concert hall"	"feels like spring," "building towards resolution or happy ending"	"diminished chords led to feeling of grandeur", "associate the dynamic (small energy) with peacefulness"
13		"feel 'conflict' rather than 'tension,'" "nice"	"sounds more solemn than sad, because of soft dynamic level and minor tonality", "strong sounds of violin", "background strings really pushing the tension meter"
14	"lack of musical experience makes me feel like I'm missing something"		"jubilant, running strings with fugal lines; love it!", "tension in strings builds up with punctuation of power in horns"
15		"suitable for a relaxed, holiday morning," "musical tension in some moments, sounds dramatic," "felt a little bit of tension"	

sense of urgency." The variety in comments regarding the same excerpts of music suggest major differences in how listeners perceive the emotional content of music.

VI. DISCUSSION

Limitations of this study, like many other studies in the field of music and emotion, lie in the question of the selection of the approach of emotional categorization, the choice of musical genre, and the selection of excerpts.

Some findings of this study are worth contemplating. One notable result in the correlations of musical descriptors with emotional ratings is the prevalence of "key clarity" as one of the musical descriptors that correlated significantly with emotional ratings (included in four out of six of the emotions). This seems to suggest that the ambiguity of key influences what emotions listeners recognize, more so than the actual modality (i.e. the degree of "minor-ness" or "major-ness") of the music. Another point to notice is that at least four MFCC coefficients correlated significantly with each emotion (except for "power" with which only one MFCC coefficient correlated significantly). This suggests that timbre has an important influence on emotion recognized in listeners, which points to the importance of orchestration in classical pieces such as Beethoven.

To place a heavier focus on individual emotions, the correlation between RMS mean and RMS standard deviation with transcendence ratings suggests that listeners experience feelings of "transcendence" when listening to loud music of greatly altering dynamic. Furthermore, transcendence and power are similar in their positive correlations with low energy rate, indicating that they both share some brightness in sound.

Why did subjects who listen to classical music more infrequently agree more upon their emotional responses, compared to subjects who listen frequently to classical music? Secondly, why did younger subjects agree more on their emotional responses than subjects of the older generation? A possible explanation for the first question is that listeners with frequent exposure to classical music associate their own biases, interpretations or memories to classical music. These

personal "interpretations" tend to diversify more. On the other hand, listeners with less frequent exposure to classical music rely on first impressions, which may be more agreed upon.

A plausible answer to the second question could be that due to the increasing prevalence of the internet, people of younger generations are more exposed to similar music, more so that people of older generations. Another, possibly more likely, explanation is the younger subjects have less experience with music, and thus are naïve listeners. Naïve listeners may rely on surface-level musical cues (such as rhythm or modality) to experience emotion, whereas more experienced listeners pay attention to a wider range of musical cues and thus their responses are more divergent.

The user-generated comments give rise to philosophical considerations surrounding the emotional experience of listening to music. The first point is the universality vs. cultural specificity of emotional responses to music. There is a common belief that music is the "universal 'language of the emotions'; that is… the suggestion that expressiveness can be recognized cross-culturally". However, as Davies (2001, p. 37) puts it, culture plays a surprisingly influential role in how one perceives, and emotionally reacts to, music.

> "Until one appreciates the belief systems that determine the significance of the social settings in which emotions are situated, and then recognizes the connection of music with all this, it will not be a simple matter to read off expressiveness from foreign music."

A point to take into consideration is that differences in how individuals emotionally respond to music is affected not only by their cultural upbringing, but by other demographic factors, such as age. As observable in the comments (Table 8), individuals can have nearly opposite emotional reactions to the exact same excerpt of music. While it is known that all cultures have some form of music (i.e. the appreciation of music is a cultural universal), the emotional appraisal of music is not necessarily the same across cultures—music may not necessarily be a "universal language of emotions."

The last point to mention, which perhaps applies not only to emotions recognized in music, but to emotions in all aspects of life, is that emotions are rarely, if ever, straightforward. Some comments received included, "paradoxical (tension and peacefulness)" or "tinge of sadness (even though "joyful activation"). The often paradoxical combinations of contrasting emotions that music conveys is perhaps a reflection of emotions experienced in everyday life. Music reflects the complexity, many shades of grey, of emotion, that we experience in real life.

VII. CONTRIBUTIONS AND FUTURE WORK

The present study contributes several findings to the field of music psychology: musical descriptors that correlate with real listeners' emotional responses to music (specifically excerpts from Beethoven's Third Symphony), quantifiable measurements and correlations between these descriptors and emotional ratings. The focus was on specific measurable audio descriptors that can be automatically extracted from audio.

Furthermore, we contribute findings regarding what elements of musical experience/exposure and demographic /cultural background influence how listeners respond emotionally to the same clips of music. Subjects with less frequent exposure to classical music agree more than subjects with frequent exposure. Younger subjects agree more than older subjects. Such findings might say something about how people perceive music and respond to it emotionally.

Future work could examine the association between emotional responses and musical descriptors in a wider range of musical genres. The present study focused on a limited set of excerpts from only one symphony by Beethoven. It would be worthwhile to see whether these correlations found in this study also apply to other classical pieces and other genres of music, or world music. An extension of this study would also benefit from collecting responses from a larger group of people with a wider range of cultural backgrounds.

Many music listeners would agree that music and emotion are intertwined and strongly related, and yet to quantitatively prove the relationship between these two rather intangible forces is a challenge—even more so when the relationship differs between listeners, depending on their cultural or musical background, or even just individual preferences. This study has been an attempt to take a step towards a more concrete understanding of how music can influence emotions. There are still many more paths in the field of music and emotion research to be examined, and deepening our understanding of how exactly we are emotionally moved by the sound of music may have many yet unexplored uses.

ACKNOWLEDGMENT

I would like to thank the MIT-Spain program, through which my visit to the MTG was funded, as well as the members of the PHENICX initiative for supporting my visit. This research was partially supported by the European PHENICX STREP project No. 601166. I am grateful for Emilia Gómez for advising me throughout this project, and Agustín Martorell and Perfecto Herrera for their invaluable advice. Finally, I would like to thank the people who participated in the online study for their time and thoughtful feedback.

REFERENCES

Balkwill, L.L. & Thompson, W.F. (1999). A cross- cultural investigation of the perception of emotion in music. *Music Perception, 17(1),* 43-64.

Blood, A.J. & Zatorre, R.J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, 98,* 11818-11823.

Davis, S. (2001). Philosophical perspectives on music's expressiveness. In *Music and Emotion: Theory and Research.* (pp. 23-44). New York: Oxford University Press.

Farbood, M.M. (2006). A quantitative, parametric model of musical tension (Doctoral dissertation). Retrieved from the Massachusetts Institute of Technology Libraries, http://opera.media.mit.edu/projects/tension.html

Fritz, T., Jentschke, S., Gosselin, N. Sammler, D., Peretz, I., Turner, R., Friederici, A.D., Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology, 19(7),* 573-376.

Gabrielsson, A. (2001). Emotions in strong experiences with music. In *Music and Emotion: Theory and Research*. (pp. 431-449). New York: Oxford University Press.

Grewe, O., Kopiez, R., & Altenmuller, E. (2009). The chill parameter: Goose bumps and shivers as promising measures in emotion research. *Music Perception: An interdisciplinary Journal, 27(1),* 61-74.

Hevner, K. (1936). Experimental studies on the elements of expression in music. American Journal of Psychology, 48(2), 246-267.

Hu, X., Chen, Z., & Yang, D. (2009). Lyric-based song emotion detection with affective lexicon and fuzzy clustering method. In *Proc. of the International Society for Music Information Conference,* Kobe, Japan.

Hu, X., Downie, J., Laurier, C., Bay, M., & Ehmann, A. (2008). The 2007 MIREX audio mood classification task: Lessons learned. In *Proc. of the International Conference on Music Information Retrieval,* Philadelphia, PA.

Juslin, P.N. & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research, 33(2),* 217-238.

Juslin, P.N. & Sloboda, J.A. (2001a). Music and emotion: introduction. In *Music and Emotion: Theory and Research*. (pp. 3-20). New York: Oxford University Press.

Juslin, P.N. & Sloboda, J.A. (2001b). Psychological perspectives on music and emotion. In *Music and Emotion: Theory and Research*. (pp. 71-104). New York: Oxford University Press.

Kim, Y.E., Schmidt, E.M., Migneco, R., Morton, B.G., Richardson, P., Scott, J., Speck, J.A., Turnbull, D. (2010). Music emotion recognition: A state of the art review. *11th International Society for Music Information Retrieval Conference*.

Krumhansl, C.L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51(4),* 336-352.

Laurier, C., Sordo, M., Serra, J., & Herrera, P. (2009). Music mood representation from social tags. *Procession of the International Society for Music Information Conference.,* Kobe, Japan

Laurier, C. (2011). *Automatic classification of musical mood by content-based* analysis (Doctoral dissertation). Provided in print by Universitat Pompeu Fabra.

International Music Information Retrieval Systems Evaluation Laboratory (IMIRSEL). (2012). *MIREX 2012: Audio Train/Test: Mood Classification – MIREX08 Dataset* [Online report] Retrieved from http://www.music-ir.org/nema_out/mirex2012/results/act/mood_report/index.html

Nyklicek, I, Thayer, J.F., Van Doornen, L.J.P. (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology, 11(3),* 304-321.

Orio, N. (2006). Music retrieval: A tutorial and review. *Foundations and Trends in Information Retrieval, 1(1)*, 1-90.

Paiva, R.P. (2012). From Music Information Retrieval to Music Emotion Recognition [PowerPoint slides]. Retrieved from http://www.slideshare.net/donruca/from-music-information-retrieval-to-music-emotion-recognition

Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception, 13,* 171-201.

Pike, A. (1972). A phenomenological analysis of emotional experience in music. *Journal of Research in Music Education, 20,* 262-267. [Abstract]

Russel, J.A. (1980). A circumspect model of affect. *Journal of Psychology and Social Psychology, 39(6),* 1161.

Schubert, E. (2003). Update of the Hevner adjective checklist. *Perceptual and Motor Skills, 96*, 1117-1122.

Siegwart, H. & Scherer, K.R. (1995). Acoustic concomitatnts of emotional expression in operatic singing. The case of Lucia in Ardi gli incense. *Journal of Voice, 9,* 249-260.

Sloboda, J.A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19,* 110-120.

Torres-Eliard, K., Labbe, C., Grandjean, D. (2012). Towards a dynamic approach to the study of emotions expressed by music. Proceedings 4th International ICST Conference on Intelligent Technologies for Interactive Entertainment, Genoa, Italy.

Zentner, M., Grandjean, D., Scherer, K.R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*, 494.