

# Emotions perceived and emotions experienced in response to computer-generated music

Maciej Komosinski      Agnieszka Mensfelt

Institute of Computing Science  
Poznan University of Technology  
Piotrowo 2, 60-965 Poznan, Poland

e-mail: [maciej.komosinski@cs.put.poznan.pl](mailto:maciej.komosinski@cs.put.poznan.pl)  
fax: +48 61 8771 525

## Abstract

This paper explores perceived and experienced emotions elicited by computer-generated music. During the experiments, 30 participants listened to 20 excerpts. Each of the excerpts lasted for about 16 seconds and was generated in real-time by specifically designed software. Measurements were performed using both categorical (a free verbal description) and dimensional approaches. The relationship between structural factors of music (mode, tempo, pitch height, rhythm, articulation and presence of the dissonance) and emotions was examined. Personal characteristics of the listener: gender and musical training were also taken into account. The relationship between structural factors and the perceived emotions was mostly congruent with predictions derived from the literature, and the relationship between those factors and experienced emotions was very similar. Tempo and pitch height – the cues common to music and speech – turned out to have a strong influence on the evaluation of emotion. Personal factors had a marginal effect. In the case of verbal categories comparable with the dimensional model, a strong correspondence was found.

**Key words:** computer-generated music, emotions, perceptions, feelings, Russell’s model

## 1 Introduction

The affective algorithmic composition is a relatively young, yet rapidly growing field. It comes as no surprise that the emotional content of artificially generated music has become a matter of interest: most people indicate emotions as their main motivation for listening to music (Juslin & Laukka, 2004). The discipline has already achieved some successes – systems that influence perceived emotions in an intended way are being developed, and various strategies are employed to accomplish this goal: modification of the score (Oliveira & Cardoso, 2008), generation of scores (Wallis, Ingalls, Campana, &

---

The final version of this paper appeared in *Music Perception* 33(4):432–445, 2016.  
<http://dx.doi.org/10.1525/MP.2016.33.4.432>

Goodman, 2011), modification of the performance features (Friberg, Bresin, & Sundberg, 2006) or various combinations of these approaches (Livingstone, Mühlberger, Brown, & Thompson, 2010). Recently, a framework for categorization and evaluation of affective algorithmic composition systems has been proposed (Williams, Kirke, Miranda, Roesch, & Nasuto, 2013). However, the relationship between music and emotions is far from being fully explored. In particular, the potential difference between emotions that are perceived in a musical piece and emotions that are truly experienced by listeners is highly intriguing, and this issue is investigated here in the context of computer-generated music.

## 1.1 Emotions perceived and experienced

There is disagreement in the field of music psychology concerning the quality of emotions induced by music. Some researchers argue that music can express emotions like fear, joy or anger, but it can not induce them in a listener, however music can “move”, arouse the state of being exited about the beauty of the piece, the mastery of the composer etc. (Kivy, 1990). Others argue that music can only arouse low-grade affective states, and only through mediators like memories and associations (Konečni, 2008). However, the claim that music has an ability to arouse ‘real’ emotion in a listener has support from neuroscientific studies (Koelsch, 2010; Kreutz & Lotze, 2007; Peretz, 2010; Panksepp & Bernatzky, 2002) and from the influence of music on subjective feelings, physiology, expressive behavior and action tendency; see (Juslin, 2011) for a review.

The fact that music arouses emotions in humans – sometimes even peak experiences (Gabrielsson, 2010) – raises the question of why it possesses such an ability. Juslin proposed a theoretical framework (Juslin & Västfjäll, 2008; Juslin, Liljeström, Västfjäll, & Lundqvist, 2010; Juslin, 2013) which now covers eight mechanisms – besides a cognitive appraisal (Lazarus, 1991) – of emotion induction by music: brain stem reflex, rhythmic entrainment, evaluative conditioning, emotional contagion, visual imagery, episodic memory, musical expectancy, and aesthetic judgment.

The relationship between emotion perceived and emotion felt is complex; the emotion felt by the listeners can differ from the emotion they perceive. It was found that music perceived as sad can evoke positive emotions (Vuoskoski, Thompson, McIlwain, & Eerola, 2012; Kawakami, Furukawa, Katahira, & Okanoya, 2013). Gabrielson (Gabrielsson, 2002) proposed four types of relationship between emotion perceived and induced: positive, negative, no systematic relation, and no relation at all. Evans and Schubert noted that there could be another relationship that occurs when emotion perceived and felt are different, but not directly opposite (Evans & Schubert, 2008). They also investigated the frequency of each type of relationship proposed by Gabrielson and confirmed his claim that a positive relationship is far from general – however, it is the most frequent one. In their research, a positive relationship was found in 61% of cases, a negative relationship was second in frequency (22%), the third one was no systematic relationship (12%), and finally, no relationship (5%). A positive relationship turned out to be preferred by the listeners.

Recently, 16 studies concerning emotions experienced and perceived in response to music have been reviewed by Schubert (Schubert, 2013). The key finding was that the emotions experienced are generally the same or lower in magnitude than emotions perceived by the listener. Schubert also proposed to reduce the classification of relationship types to two categories: ‘matched’ and ‘unmatched’.

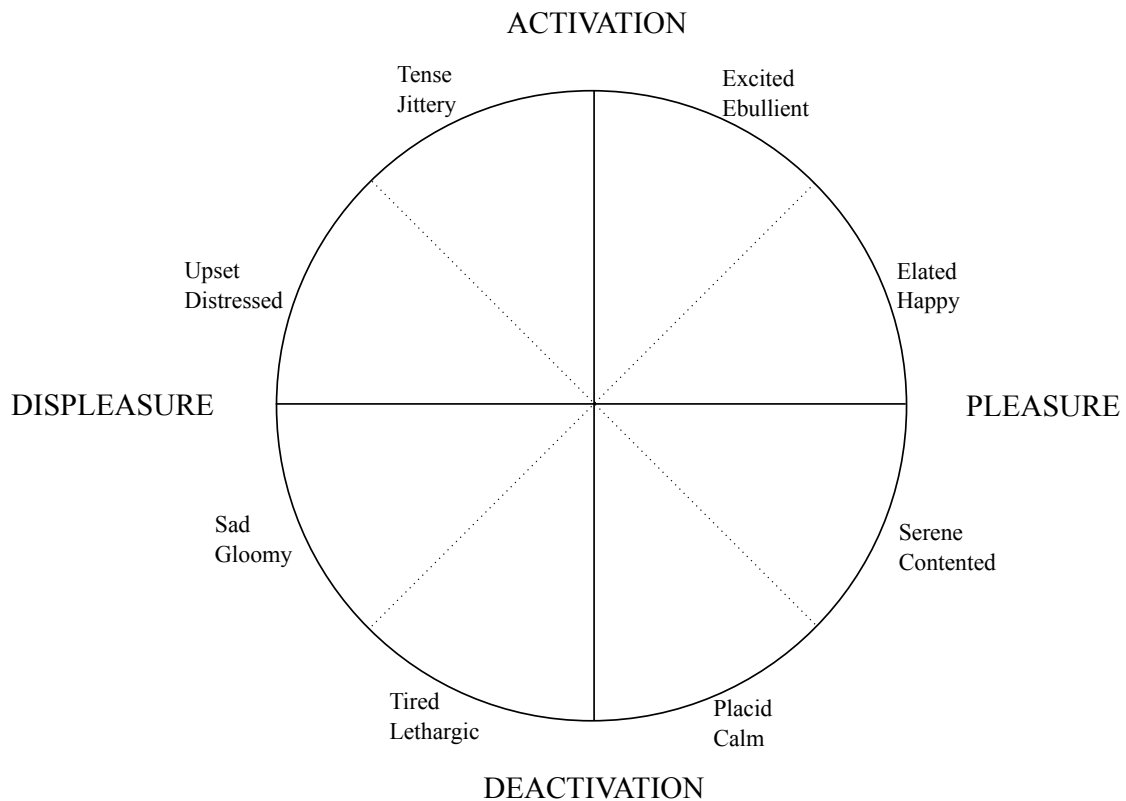


Figure 1: Russell’s model of the core affect (Russell, 2003).

## 1.2 Conceptualization of emotions

There are several conceptualizations of emotions. Often found in music and emotion research is Russell’s circumplex model characterized by two bipolar dimensions: arousal and pleasure-displeasure, and a circle of categories located in this coordinate system (Russell, 1989). This model is supported by research with the use of self-reports, face expression judgments and by research on the similarity of emotional terms (ibidem). It also has some evidence from neuroimaging studies (Kreutz & Lotze, 2007). Still, the model with two dimensions has received some criticism. For instance, it is argued that two is too small a number of dimensions to capture the structure of emotions and the two-dimensional model cannot distinguish between fear and anger (Fontaine, Scherer, Roesch, & Ellsworth, 2007), or that the model does not fit the data (Schimmack & Grob, 2000). In the face of evidence of weaknesses of his model, Russell enhanced it by combining a dimensional approach with categorical and prototype approaches (Russell, 2003). The base of the model is the core affect, described by the pleasure–displeasure and arousal dimensions, and some emotional and non-emotional terms (Fig. 1). The core affect is present all the time – a person is always at some point in this space. If the core affect is significantly changed, an emotional episode takes place, which is interpreted in terms of categories (e.g. fear). A recent finding of good correspondence between the dimensional and categorical models (Eerola & Vuoskoski, 2011) seems to support such an approach.

### 1.3 Factors influencing musical emotions

The multiplicity of mechanisms of musical emotions induction translates to a multiplicity of factors which need to be taken into account while investigating music experience. The effects of interaction between these factors must be considered too. Scherer and Zentner proposed a model where “experienced emotion = Structural features  $\times$  Performance features  $\times$  Listener features  $\times$  Contextual features” (Scherer & Zentner, 2001). Juslin and Sloboda point out the interaction between the music, the listener and the situation, and accentuate personal and contextual factors of this interaction (Juslin & Sloboda, 2011). Personal factors cover, among others, familiarization (Schellenberg, Peretz, & Vieillard, 2008), gender, personality (Liljeström, 2011), and musical training (Hosinho, 2006).

In the case of the influence of structural and performance features, the perception of musical emotions is explored more than the feeling of them. Recently, meta-analysis of over a hundred studies concerning the relationship between structural/performance features and perceived emotions was carried out (Gabrielsson & Lindström, 2011). The influence of structural features on experienced emotions was examined to a lesser extent (Gomez & Danuser, 2007; Coutinho & Cangelosi, 2011).

## 2 The investigation

Motivated by existing doubts and contradicting arguments mentioned in the previous section, a study has been performed aimed at investigating emotions elicited by computer-generated music. The knowledge of the possible relationships between structural factors of music, emotion perception and emotion induction was employed in real-time generation of music. The generated excerpts were short and structurally unsophisticated so that the influence of structural factors could be properly determined. There was no predefined set of stimuli, only the rules for generation of the excerpts were predefined. All of the 600 excerpts presented to the participants came into existence during the investigation. Therefore, even when they were generated according to the same values of parameters, they differed slightly. The influence of slightly differing excerpts generated according to the same parameter values was investigated by evaluating the consistency of responses not only between participants, but also within each participant. To this end, each set of structural factors was used twice for each participant.

The personal factors of listeners were collected in a questionnaire. The subjects had an option to split their answers when their experienced and perceived emotions differed (see Sect. 1.1). Emotions were quantified using both dimensional and categorical approaches. The measurement of the emotions induced by music was performed using self-reports, i.e., participants were asked to describe their emotions themselves.

### 2.1 Participants

Thirty volunteers: 16 females and 14 males participated in the study. Twenty subjects declared no musical training. Nine subjects declared playing instruments/singing as non-professionals, and one declared being a professional. Age, the average number of hours spent listening to music per day and years of musical training of the listeners are summarized in Fig. 2.

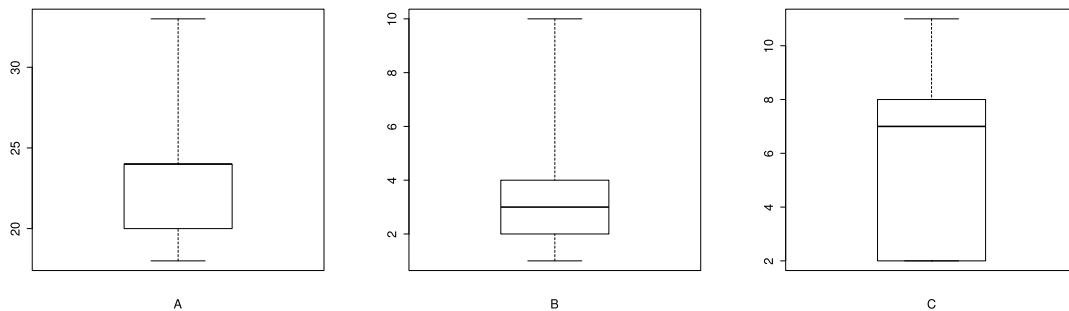


Figure 2: **A**: Age of the participants (in years), **B**: The average number of hours spent listening to music per day, **C**: Years of musical training of those participants who declared playing instruments/singing.

Set	Mode	Tempo*	Pitch**	Rhythm	Articulation	Dissonance
A	major	medium	medium	regular	standard	no
B	minor	medium	medium	regular	standard	no
C	major	fast	medium	regular	standard	no
D	major	slow	high	regular	legato	no
E	major	fast	high	regular	standard	no
F	minor	fast	low	irregular	standard	yes
G	major	medium	medium	regular	legato	no
H	minor	slow	high	regular	legato	no
I	major	slow	medium	regular	standard	no
J	major	medium	medium	regular	staccato	no

Table 1: Sets of parameter values used in the investigation. The A–J letters denote the names (IDs) of the sets. \*Values of the tempo in bpm (beat per minute): slow = 75, medium = 115, fast = 145. \*\*For the medium pitch, the melody track was in the range of C4 to C5 (chords and bass line tracks were one and two octaves lower, respectively). For the low/high pitch, all tracks were an octave lower/higher than for the medium pitch.

## 2.2 Stimuli

The structural factors of music were the parameters of the generation process. Tempo and articulation may be considered performance features rather than structural features, yet in this investigation they are referred to as structural factors in opposition to personal and contextual factors. Each excerpt was generated using values from one of ten predefined sets (Table 1, Fig. 3). Parameter values in each set were combined according to the contemporary knowledge about the relationship between the structural factors of music and the perceived/experienced emotions (Coutinho & Cangelosi, 2011; Gabrielsson & Lindström, 2011; Gomez & Danuser, 2007).

A pilot study revealed that the optimal duration of the investigation was about 20 minutes. Therefore, the number of stimuli and investigated parameters was limited. Values that may be associated with the same emotion, or with the same place in the emotional space in the case of the dimensional approach, were grouped together. The

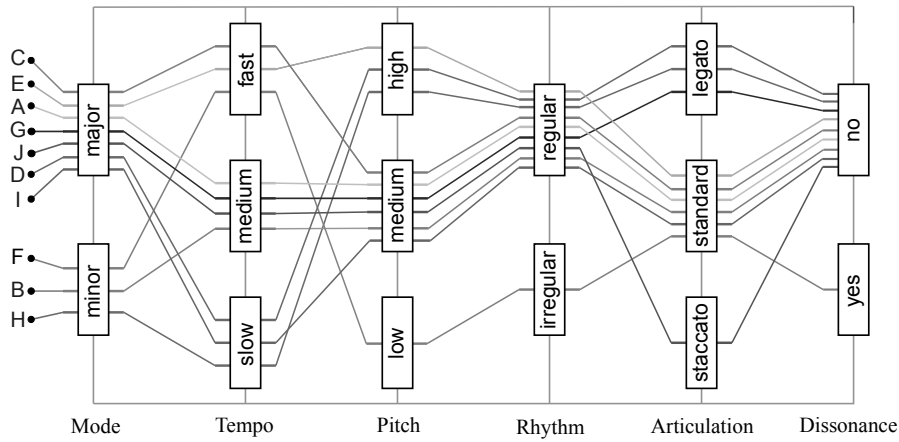


Figure 3: Visual comparison of the sets of parameter values used in the investigation. Rectangles on the vertical lines contain possible values of each parameter.

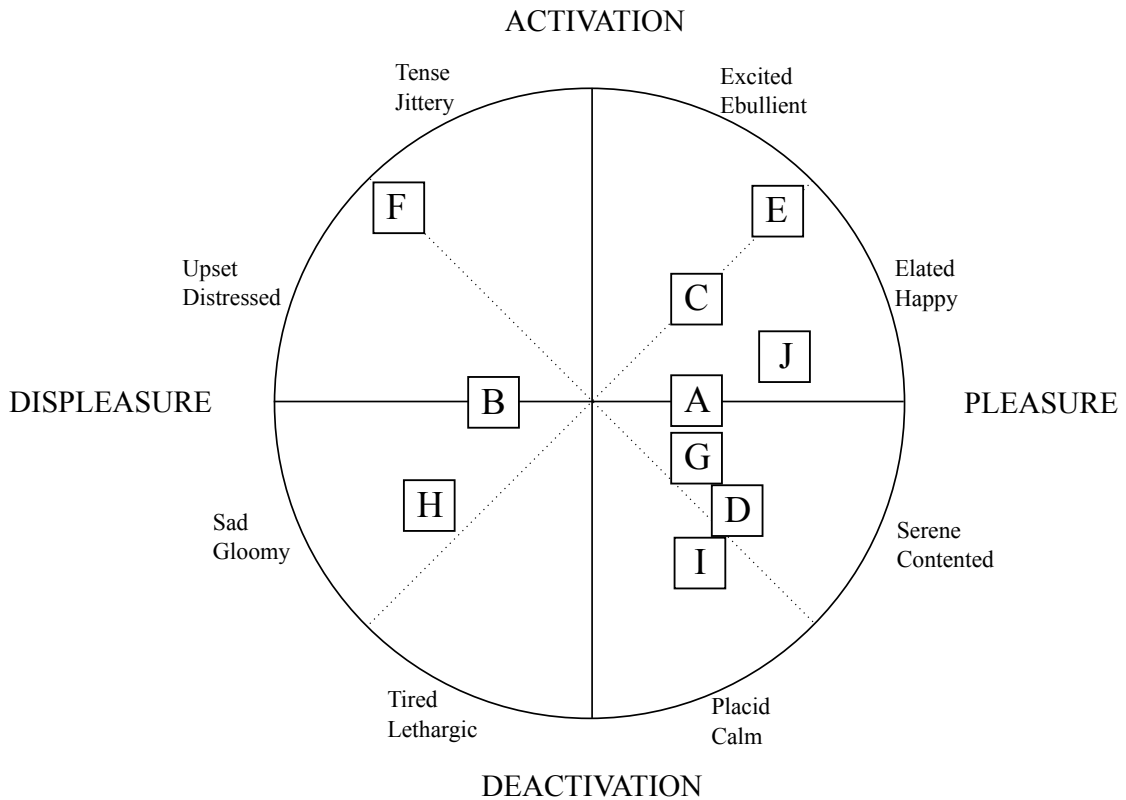


Figure 4: Approximate expected influence of generated music on emotions. Each square demonstrates intended placement of emotions in the affective space, connected with a particular set of parameters (A–J); see also Table 1.

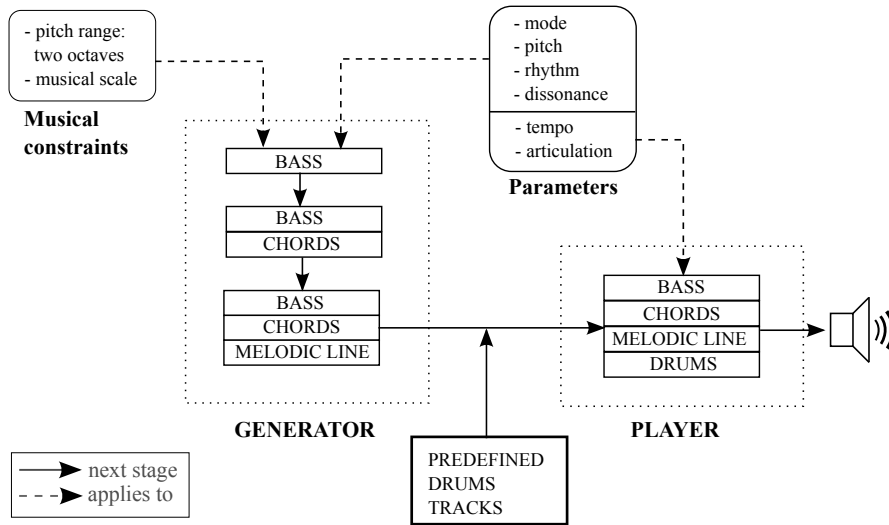


Figure 5: General scheme of the music production process. The program is divided into two main parts: a generator responsible for producing scores and a player responsible for their performance.



Figure 6: A sample score generated and performed by the application during the investigation.

authors decided to choose sets of parameter values that, according to the literature, correspond to every quadrant of the affective space (F, E, H, D). Additionally, to investigate influence of individual parameters, all parameter values were kept at the level which may be considered neutral (like medium pitch height, medium tempo), and the value of only one parameter was varied – e.g., sets A and B differed only in mode. The mode has no neutral value, because both major and minor modes are strongly related to particular emotions (Eerola, Friberg, & Bresin, 2013). In other neutral sets, the mode was arbitrarily set to major. Sets C and I were used to compare the influence of tempo, and sets G and J – to investigate the influence of articulation. An approximate expected influence of music generated with parameters from each set on emotions is presented in Fig. 4.

All excerpts presented to the participants were generated in real-time by the software developed specifically for this purpose. Music was generated in the MIDI format. The production of scores was based on a random choice of notes for a baseline (bass) within the musical constraints. First of all, successive baseline notes had to satisfy the rules of

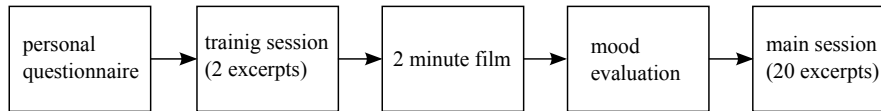


Figure 7: The order of tasks in the investigation. Each task was displayed in a new window. The order was the same for every participant.

chord progressions. Secondly, notes were chosen within one musical scale and within a range of two octaves. The chords and melodic line were then generated according to the baseline. The drums were randomly picked from several predefined tracks. A general scheme of the music generation process is presented in Fig. 5. Each excerpt lasted for about 16 seconds. A sample score generated and performed by the application is presented in Fig. 6.

### 2.3 Procedure

The participants were tested in individual sessions. The sessions took place at listeners' homes which was a trade-off between experimental control and ecological validity – the experimental situation occurred in the natural environment. An average session lasted for about 20 minutes. Participants were first instructed by the researcher. After the instruction had been given, the procedure was continued using a dedicated application run on the laptop (Ubuntu Linux) using Koss Porta Pro headphones (frequency response 15–25,000 Hz).

At the beginning, subjects filled in the questionnaire designed to determine personal factors that could influence musical emotions: age, gender, and musical experience. A training session followed, aimed at familiarizing subjects with the graphical interface of the application. They had to evaluate two excerpts generated using a randomly selected, predefined set of parameter values (A–J, Fig. 3). After the training session and before the main part of the investigation, participants watched two minutes of a nature film clip with a neutral content to induce a neutral mood. The clip did not contain any music, only the nature sounds (mainly singing birds). Then a self-evaluation of the listener's mood was collected in order to determine its influence on musical emotions. The main part of the study followed. Listeners were presented with 20 excerpts (each set of parameter values was used twice) in a random order. For each excerpt, they had to evaluate experienced emotions on negative-positive and low arousal-high arousal dimensions, each on a scale of  $-50$  to  $50$ , using a horizontal slider. They were also asked to describe their emotion verbally, by filling in a text box, although this was not obligatory. They had a possibility to separate experienced emotions from perceived emotions by selecting a check box that activated additional sliders and an additional text box for experienced emotions. The listeners were also asked to rate how much they liked the presented excerpt on a scale from 0 to 10. The procedure is presented in Fig. 7.

### 2.4 Data analysis

The structure of the emotional response data is presented in Fig. 8. The response to an excerpt generated with each set of parameters was collected twice. The response consists of seven components: evaluation of liking (1), and for both perceived and experienced



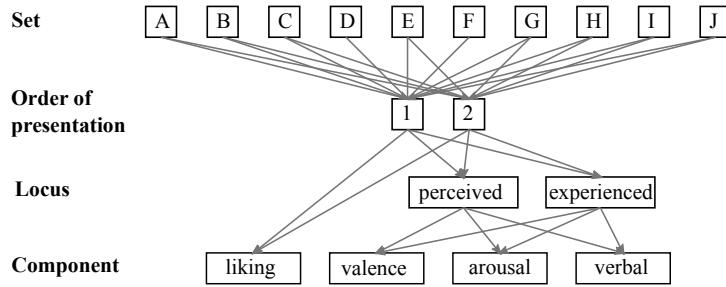


Figure 8: The structure of the emotional response data from a single experiment. Each path in the graph corresponds to one dependent variable (such as, for instance, valence perceived in the second presentation of an excerpt generated with parameter values from the set E), thus there were 140 emotional variables per person.

emotions: valence evaluation (2), arousal evaluation (2) and verbal description (2). For each response, correlations between its components (except for the verbal component) were computed.

Emotional responses were compared with data collected from the questionnaire (gender, declaration of having musical training, hours of listening to music per day) and participant’s mood (evaluated on the valence, arousal and tension dimensions). The following components of the response were compared with personal factors: valence evaluation, arousal evaluation and liking evaluation. Independent samples *t*-tests were performed to investigate possible differences in responses to the music caused by gender and musical training. Correlation coefficients were computed to assess the relationship between the emotional response and the remaining personal factors (hours of listening to music per day and mood evaluation).

Responses to the first and to the second presentation of an excerpt generated with the same set of parameters were compared for each set. Correlation coefficients were computed to assess the relationship between the corresponding components of the responses (the valence evaluation, the arousal evaluation and the liking evaluation). Since not all variables were normally distributed and outliers were present, Spearman’s rank correlation coefficient was used to compute all correlations. The comparison between different sets of parameters was made by plotting the arousal evaluation against the valence evaluation. The valence and arousal ratings were averaged across participants for each set. In cases where a similar location in the affective space was obtained, paired samples *t*-tests were performed to reveal differences along arousal and/or valence dimensions.

In order to compare answers from the dimensional scale with the verbal terms used by participants, the latter were reduced by eliminating the duplication of terms and grouping them in categories from Fig. 1: ‘Tense, Jittery’, ‘Upset, Distressed’, ‘Sad, Gloomy’, ‘Tired, Lethargic’, ‘Placid, Calm’, ‘Serene, Contented’, ‘Elated, Happy’, and ‘Excited, Ebullient’. Additional categories, derived from the literature, were introduced: ‘Mixed feelings’ (Hunter, Schellenberg, & Schimmack, 2010), ‘Aesthetic feelings’ (Konečni, 2008) and also a ‘Neutral’ category. This assignment was done independently by three competent judges; a term was assigned to a category if at least two judgments were congruent. The remaining terms were labeled as “unclassified”.

Set	Arousal		Valence		Liking
	Experienced	Perceived	Experienced	Perceived	
A	0.24	0.24	0.16	0.29	0.16
B	0.38*	0.39*	-0.08	-0.04	-0.14
C	0.69***	0.73***	0.31	0.40*	0.32
D	0.42*	0.51**	0.58***	0.35	0.79***
E	0.62***	0.72***	0.41*	0.45*	0.53**
F	0.60***	0.59***	0.47**	0.49**	0.61***
G	0.66***	0.46*	0.28	0.47**	0.63***
H	0.59***	0.54**	0.43*	0.41*	0.24
I	0.45*	0.47**	0.69***	0.60***	0.65***
J	0.50**	0.43*	0.14	0.38*	0.38*

Table 2: Correlations between responses to the first and the second presentation of an excerpt generated with each set of parameters. Stars denote: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

### 3 Results and discussion

#### 3.1 Correlations between the components of the emotional response

Emotions perceived and felt were strongly positively correlated – an average Spearman’s rho was 0.96 for the arousal evaluation and 0.89 for the valence evaluation. In general, experienced emotion ratings differed from perceived emotion ratings in 9% of cases. It is worth noticing that 68% of differences were generated by 5 out of 30 participants. The valence and arousal dimensions were in most cases independent: a significant positive correlation between valence and arousal evaluation occurred in 32.5% of responses. The results mentioned in this paragraph were significant at the level of  $p = 0.05$  or lower.

#### 3.2 Structural factors

Values of mode (major, minor), tempo (slow, medium, fast), pitch height (low, medium, high), rhythm (regular, irregular), articulation (legato, standard, staccato) and the presence (or the lack) of a dissonance were taken as parameters for the music generation process. Medium tempo, medium pitch height, regular rhythm, standard articulation and the lack of a dissonance may be considered a set of standard parameter values in this study.

Evaluations of the first and the second presentation of an excerpt generated with each set of parameters were compared in order to determine repeatability of the emotional response to music generated in real-time with a given set of parameters. The comparison took into account the following components of the response: experienced and perceived arousal, experienced and perceived valence, and liking; the results are presented in Table 2. In the case of one set of parameters (set A), no correlation occurred. The arousal component – both perceived and experienced – was positively correlated in all other sets of parameters. The experienced valence component was positively correlated in 50% of sets (no correlation occurred in sets A, B, C, G, J). The perceived valence component was positively correlated in 70% of sets (no correlation occurred in sets A, B, D). These findings are consistent with the observation that evaluation of the arousal dimension seems to be easier than evaluation of the valence dimension (Gabrielsson &

Lindström, 2011). The liking component was positively correlated in 60% of sets (there was no correlation in sets A, B, C, H).

It is worth noticing that the least stable sets were A and B with no correlation at all for set A, and no correlation in perceived and experienced valence for set B. Both sets had standard parameter values, and they differed only in mode. The mode alone turned out to be too weak an emotional cue to provide high repeatability of answers. However, evaluations of valence of sets A and B significantly differed, as discussed in the following paragraphs.

The valence and arousal components of the response to excerpts generated with each set of parameters were averaged and plotted; Figs. 9(a) and 9(b) show perceived emotions and experienced emotions, respectively, in a two-dimensional affective space. Since the perceived and experienced emotions were strongly overlapping, corresponding points in both plots have, in most cases, similar location. Paired samples *t*-tests were performed to confirm visible differences between the responses to individual sets. Table 3 compares these differences with differences between sets of parameter values. A detailed discussion of these results follows, along with references to related results reported in the literature.

A number of studies have found that a **major mode** is related to positive valence (happiness in terms of the categorical approach), and minor mode to negative valence (sadness) for both perceived (Eerola et al., 2013; Fritz et al., 2009; Gabrielsson & Lindström, 2011) and experienced (Gomez & Danuser, 2007) emotions. In these experiments, sets A (major mode) and B (minor mode) had, apart from the mode, standard values of parameters. These sets affected valence ratings (A had a positive valence and B had a negative valence) but not arousal ratings (they did not differ significantly in arousal), confirming the relationship of scale type with the valence dimension.

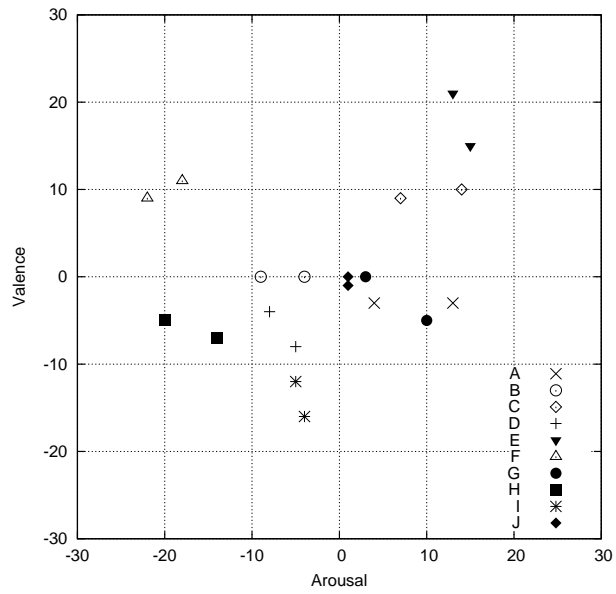
**High tempo** is reported to be associated with happiness (Fritz et al., 2009; Gabrielsson & Lindström, 2011; Juslin & Laukka, 2003) for perceived emotions, and with high arousal and positive valence for experienced emotions (Coutinho & Cangelosi, 2011; Gomez & Danuser, 2007). Our results confirm the connection between a high tempo and high arousal. Set C, which differed from set A in tempo (fast), resulted in similarly positive valence evaluations as set A, but significantly higher arousal evaluations.

Set E – similar to set C except for the **pitch** value (**high**) – raised arousal assessments comparable to set C, confirming the relationship of high pitch and high arousal reported in the literature for perceived emotions (Gabrielsson & Lindström, 2011) and experienced emotions (Coutinho & Cangelosi, 2011).

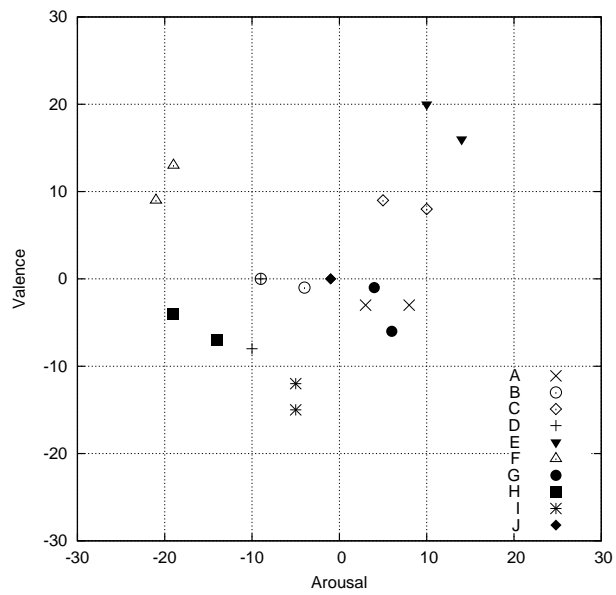
Set F had parameter values that are known (Gomez & Danuser, 2007) to be connected to low valence (**minor mode, irregular rhythm, dissonance**) and high arousal (**high tempo, low pitch, irregular rhythm, dissonance**). Results obtained for this set were consistent with the predictions based on the literature. This finding seems to confirm the relationship of minor mode, high tempo, low pitch, irregular rhythm and dissonance with low valence and high arousal.

Set H, with parameter values that are often connected to sadness: **minor mode, slow tempo, legato articulation** (Gabrielsson & Lindström, 2011), but also **high pitch** related with, among other things, activity (ibidem) – resulted, in these experiments, in low valence ratings and moderately low arousal ratings. Evaluations of arousal were higher than expected. This result suggests that pitch has a strong connection with the arousal dimension.

In sets D and I, **major mode, slow tempo, legato articulation** and **high pitch**



(a) Perceived emotions



(b) Experienced emotions

Figure 9: Averaged valence and arousal ratings of excerpts generated with each set of parameters. Each symbol represents one set of parameters. Parameter values of each set are presented in Table 1. The averages were calculated separately for the first and for the second presentation. The axes of the plot correspond to the axes of Russell's model of core affect (Fig. 1).

Parameter(s)	Differences in values		Differences in the responses
pitch	E high	C medium	E higher in arousal than C
tempo	C fast	A medium	C higher in arousal than A
mode	A major	B minor	A higher in valence than B
articulation	A standard	G J legato staccato	No significant differences
mode tempo pitch articulation	B minor medium medium standard	D major slow high legato	No significant differences
mode tempo	B minor medium	I major slow	B greater in arousal than I
tempo pitch articulation	B medium medium standard	H slow high legato	B greater in valence than H
mode pitch articulation	H minor high legato	I major medium standard	H greater than I in arousal I greater than H in valence
tempo pitch articulation rhythm dissonance	F fast low standard irregular present	H slow high legato regular absent	F greater in arousal than H
tempo pitch rhythm dissonance	F fast low irregular present	B medium medium regular absent	F greater than B in arousal B greater than F in valence

Table 3: Comparison of differences in parameter values and revealed differences in the responses for sets occupying similar location in the affective space. Results were obtained using paired samples *t*-test with a significance level of 0.05.

were the parameters that differed from the standard set. According to the literature, for perceived emotions, the first three structural features may be related to tenderness (Gabrielsson & Lindström, 2011) which, in terms of the dimensional model, is connected with positive valence and low arousal. High pitch may be related to positive, low arousal emotion as well, but, as mentioned earlier, also to activity (ibidem). Set I had standard parameter values except for tempo, which was slow and it had the major mode. The results obtained for this set are moderately low in valence and low in arousal. Set D had similar parameter values, but also had high pitch and a legato articulation. It is similar to set I in valence, but significantly higher in arousal. The latter finding together with the results from set H indicate a relationship of high pitch with high arousal. It also suggests that pitch height has a strong relationship with an emotional evaluation of an excerpt. This is consistent with the latest research on musical cues (Eerola et al., 2013). The register was found to be the third most important cue after mode and tempo.

The results obtained for sets D and I do not confirm the conclusions reported in (Husain, Thompson, & Schellenberg, 2002) that valence may be related only to mode and not to tempo. Valence was evaluated as rather negative for those sets, despite the major mode. This finding is consistent with the study of Gagnon and Peretz which showed the supremacy of tempo over mode in the happy–sad distinction (Gagnon & Peretz, 2003). It was suggested that the possibility of using tempo as a cue for distinguishing between happy and sad music excerpts is acquired earlier in development than the possibility of using mode (Dalla Bella, Peretz, Rousseau, & Gosselin, 2001). The importance of tempo and pitch may be considered as evidence of a close relationship between emotion perception (and possibly emotion induction) in music and in speech, as those two cues are common to music and speech – in contrast to mode, which is specific to music.

Sets A, G, and J had standard parameter values, and they differed only in the type of articulation – standard in the case of set A, legato in the case of set G and staccato in the case of set J. No patterns reported in the literature were found either for legato articulation – sadness, tenderness, solemnity and softness for perceived emotions (Gabrielsson & Lindström, 2011), or for staccato articulation – gaiety, activity, energy, anger and fear for perceived emotions (ibidem), high arousal and positive valence for experienced emotions (Gomez & Danuser, 2007). Results for all those sets were located in a similar position in the affective space, which may suggest that the implementation of articulation in the software that generated the music was not sufficient to express legato and staccato strongly enough.

### 3.3 Personal factors

Independent samples *t*-tests with Welch correction (variances in groups were non-homogeneous) revealed no significant differences (significance level: 0.05) between males and females, with one exception: evaluation of the liking of set H, second presentation<sup>1</sup> (males  $M = 5.21$ ,  $SD = 1.89$ , females  $M = 3.44$ ,  $SD = 1.97$ ;  $t = 2.52$ ,  $p = 0.02$ ) and no differences between musically trained and untrained with one exception: negative–positive evaluation of set D, perceived emotions, second presentation (musically trained  $M = 33$ ,  $SD = 18.89$ , musically untrained  $M = 50.95$ ,  $SD = 21.92$ ;  $t = -2.32$ ,  $p = 0.03$ ). The arousal dimension ( $M = 4.23$ ,  $SD = 2.64$ ) was distributed normally

---

<sup>1</sup>*M* denotes the mean, *SD* denotes the standard deviation.

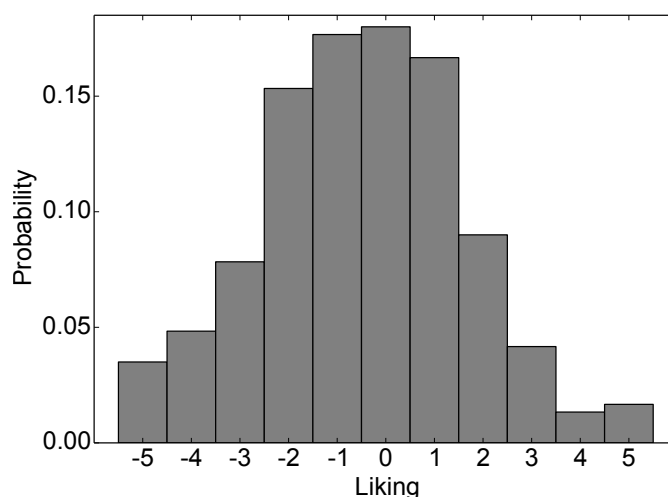


Figure 10: The histogram of the liking evaluations. The level of liking was evaluated using a 11-point scale.

while the tension ( $M = 2.1$ ,  $SD = 1.99$ ) and the valence ( $M = 7.8$ ,  $SD = 1.79$ ) dimensions were not. There was no correlation between the mood dimensions. In most cases, there was no significant correlation between mood and emotional responses (2.5% of cases), and between hours of listening to music and emotional responses (2.5% of cases).

### 3.4 Verbal component

Filling a text box designed for verbal description of emotion was not obligatory; the verbal component is missing in 29.2% of responses. The verbal description of experienced emotions differed from the description of perceived emotions where dimensional description differed as well; such differences occurred in 9% of responses. Due to a disagreement between the competent judges, 12.3% of terms used by participants were not classified in a single category, including terms like tension, boredom, despair, longing, anxiety, or astonishment. 42.4% were classified in one of the categories derived from Russell's model, 1.8% to 'Mixed feelings', 9.3% to 'Aesthetic feelings' and 4.6% to the 'Neutral' category. The infrequency of 'Mixed feelings' could be connected to the fact that parameter values in the sets provided rather congruent cues. The large number of unclassified terms, including terms that are usually connected to musical emotions ('tension' for instance) may suggest that the categories employed in the classification did not necessarily meet their purpose.

### 3.5 Liking component

The level of liking was evaluated using a 11-point scale where 0 corresponded to the default, neutral attitude. Fig. 10 presents the histogram of the evaluation of liking. 68% of the responses are nearly equally distributed in the interval  $[3, 6]$ . It indicates that participants' attitudes were close to neutral with a minor prevalence of slightly negative evaluations. The evaluation of liking was partly related to evaluations of arousal and valence, with the prevalence of the relationship between liking and valence. A significant

positive correlation between liking and valence occurred in 67.5% of responses, and a significant positive correlation between liking and arousal – in 52.5% of responses.

Valence and arousal evaluations corresponding to each category derived from Russell’s model were averaged and plotted: Figs. 11(a) and 11(b) present the results for perceived and experienced emotions, respectively. Again, plots for perceived and experienced emotions are very similar. Note that there is a considerable disparity between some categories, for instance six observations were labeled as ‘Excited, Ebullient’ and seventy nine as ‘Sad, Gloomy’.

## 4 Conclusions

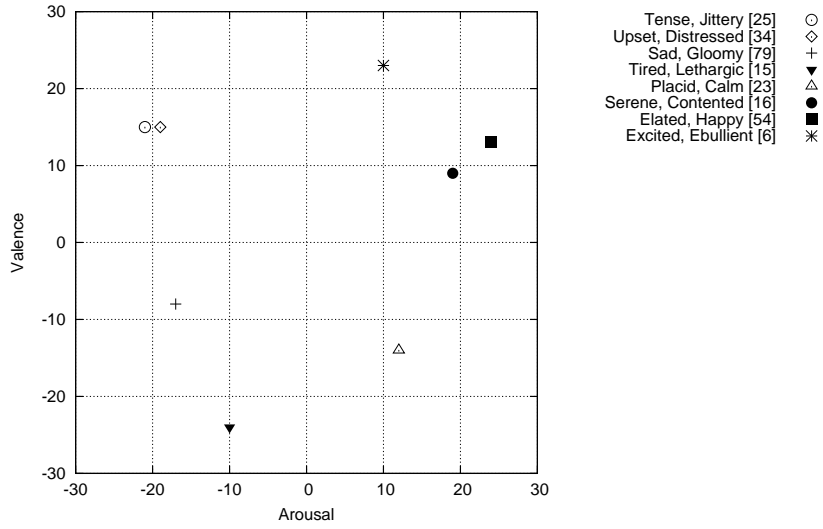
This paper explored issues related to musical emotions in the context of real-time computer-generated music:

- (1) Relationship between perceived and experienced emotions and the following factors:
  - (a) Structural factors of music: mode, tempo, pitch height, rhythm, articulation and the presence of dissonance,
  - (b) Characteristics of the listener: gender and musical experience.
- (2) Correspondence between the categorical and dimensional models of emotion.

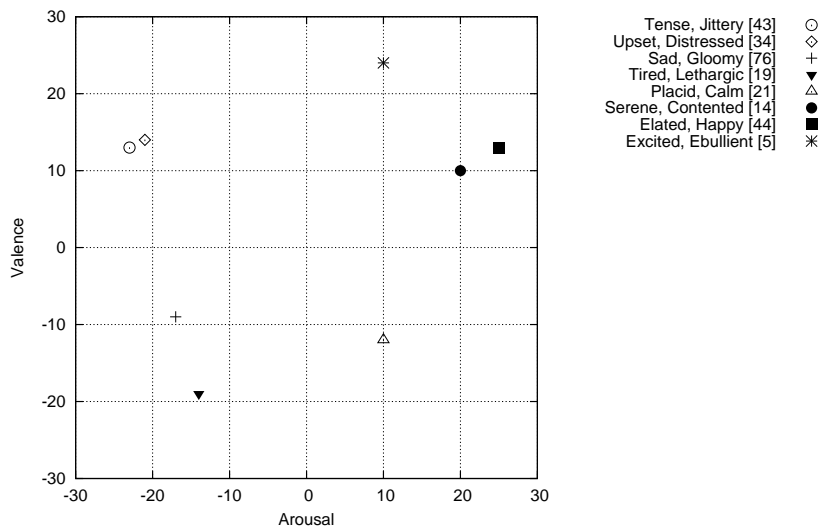
The relationship between the structural features of music and perceived emotions (1.a) was mostly congruent with the current state of knowledge regarding mapping between musical factors and emotions (Fritz et al., 2009; Gabrielsson & Lindström, 2011; Juslin & Laukka, 2004). The results suggest that in the context of simple computer-generated music, the relationship between the above-mentioned factors and experienced emotions is almost the same as in the case of perceived emotions. For both perceived emotions and experienced emotions, the listener characteristics: gender and musical training (1.b) turned out to have a marginal effect. A good correspondence between the two-dimensional model and the categorical model was confirmed (2), although only in the case of verbal categories comparable with the dimensional model. A part of the collected verbal material belonged to the ‘Aesthetic feelings’ and ‘Mixed feelings’ categories which are hard to cover in terms of valence and arousal.

It was reported in the literature that a positive relationship between perceived and experienced emotions is not the only possible relationship (Gabrielsson, 2002). Nevertheless, the positive relationship is prevalent; in previous research it was found in 61% of cases (Evans & Schubert, 2008). In this study, a positive relationship was found in 91% of cases; such a high level of coherence may be caused by the fact that artificially generated stimuli were employed, while in the aforementioned study, real music was used including pieces selected by participants. Self-chosen music was found to elicit more intense and more positive emotions (Liljeström, 2011). As mentioned earlier, Juslin proposed eight mechanisms, apart from the cognitive appraisal, which may be responsible for induction of musical emotions (Juslin & Västfjäll, 2008; Juslin et al., 2010; Juslin, 2013). The use of novel, unfamiliar stimuli may eliminate two of those mechanisms: evaluative conditioning and episodic memory. It is possible that the lack of previous experience with a piece of music makes experienced emotions more similar to





(a) Perceived emotions



(b) Experienced emotions

Figure 11: Averaged valence and arousal ratings for each category derived from Russell's model. Numbers in brackets denote the number of observations in each category.

perceived emotions by increasing the role of the emotional contagion – the mechanism by which perceived emotions are recreated inside the listener.

On the other hand, five from thirty participants were responsible for 68% of all found differences between emotions experienced and felt. It may suggest an influence of factors on experienced emotions that were not covered in this investigation, such as personality-related differences (Vuoskoski & Eerola, 2011). Finally, there is a possibility that not all participants fully understood the concept of perceived and experienced emotions, or they were unwilling or unable to differentiate between these two types of emotion.

In some of the earlier research, gender (Liljeström, 2011) and musical experience (Hosinho, 2006) were reported as factors influencing musical emotions. In this study, these factors had very little influence on emotional ratings. This result may also be related to the type of stimuli – unfamiliar, artificially generated excerpts of music. Additionally, the group that participated in the investigation was quite homogeneous, as it consisted of students aged between 18 and 33.

Comparison of the results obtained on the dimensional scale and the verbal categories derived from Russell’s model: ‘Tense, Jittery’, ‘Upset, Distressed’, ‘Sad, Gloomy’, ‘Tired, Lethargic’, ‘Placid, Calm’, ‘Serene, Contented’, ‘Elated, Happy’ and ‘Excited, Ebullient’ demonstrated congruence. This finding is consistent with a good correspondence between the dimensional and the categorical models reported in the literature (Eerola & Vuoskoski, 2011). In this study, two pairs of verbal categories were very close to each other in terms of dimensional ratings: ‘Tense, Jittery’ with ‘Upset, Distressed’ and ‘Serene, Contented’ with ‘Elated, Happy’. The proximity of the first pair may reflect the often-reported inability of the dimensional model to distinguish between fear and anger (Fontaine et al., 2007).

#### 4.1 Limitations and implications

The group investigated in this study was relatively small and homogeneous; participants were similar in age, which did not allow for cross-age comparisons. The investigation was based on self-report data, so it inherited limitations from this measurement method. Overcoming the problems of self-report is especially important in the context of experienced emotions, therefore, the study would have benefited from employing other measures of emotions: physiological or behavioral. The latter can be applied, for instance, in the context of interactive environments such as computer games.

Another issue is related to the format of the verbal response. Although the free description format enabled communication of the richness of emotional states pertinent to music, the need for categorization of such responses lowered the level of objectivity. A full, systematic exploration of verbal responses was impossible due to a large number of missing answers. An improvement to the study might be to force participants to choose any number of labels from a set of pre-defined labels, and give them a possibility to make an additional free-form comment. This study provided a list of labels used spontaneously by humans to describe their emotions when listening to music, and these labels are good candidates for such a pre-defined set.

This work is the first to focus on the relationship between computer-generated music and both perceived and experienced emotions. Results obtained for perceived emotions provide more evidence for the ability of affective algorithmic composition to express certain emotions. Results obtained for experienced emotions are promising, but they require further validation. Measures of emotion other than self-report could be useful in

reaching this goal. It would be interesting to make a comparison between differences in perceived and experienced emotions for computer-generated music and music composed and performed by humans. The results of such an investigation would shed more light on the involvement of different mechanisms in the induction of musical emotion.

## Acknowledgement

This work has been supported by the Polish National Science Centre, grant no. N N519 441939.

## References

- Coutinho, E., & Cangelosi, A. (2011). Musical emotions: predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion, 11*(4), 921–937.
- Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001, July). A developmental study of the affective value of tempo and mode in music. *Cognition, 80*(3).
- Eerola, T., Friberg, A., & Bresin, R. (2013). Emotional expression in music: contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology, 4*.
- Eerola, T., & Vuoskoski, J. K. (2011, January). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music, 39*(1), 18–49.
- Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae, 12*(1), 75–99.
- Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. (2007). The world of emotion is not two-dimensional. *Psychological Science, 18*, 1050–1057.
- Friberg, A., Bresin, R., & Sundberg, J. (2006, January). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology, 2*(2), 145–161.
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., . . . Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology, 19*(7), 573–6.
- Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Spec Issue*, 123–147.
- Gabrielsson, A. (2010). Strong experiences with music. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: theory, research, applications* (pp. 547–574). Oxford University Press.
- Gabrielsson, A., & Lindström, E. (2011). The role of structure in the musical expression. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: theory, research, applications* (pp. 367–400). Oxford University Press.
- Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to “happy-sad” judgements in equitone melodies. *Cognition & Emotion, 17*(1), 25–40.
- Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion, 7*, 377–387.
- Hosinho, E. (2006). Affective characters of music and listeners’ emotional responses to music: Comparison between musically trained and untrained listeners. In M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th in-*

- ternational conference on music perception and cognition. Alma Mater Studiorum University of Bologna.
- Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychology of Aesthetics, Creativity, and the Arts*, 4(1), 47–56.
- Husain, G., Thompson, W. F., & Schellenberg, E. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception*, 20(2), 151–171.
- Juslin, P. N. (2011). Music and emotion: Seven questions, seven answers. In I. Deliège & J. Davidson (Eds.), *Music and the mind: Essays in honour of John Sloboda* (pp. 113–135). Oxford University Press.
- Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10(3), 235–266.
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological Bulletin*, 129(5), 770–814.
- 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(3).
- Juslin, P. N., Liljeström, S., Västfjäll, D., & Lundqvist, L. (2010). How does music evoke emotions? exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: theory, research, applications* (pp. 605–642). Oxford University Press.
- Juslin, P. N., & Sloboda, J. A. (2011). At the interface between inner and outer world: psychological perspectives. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: theory, research, applications* (pp. 72–97). Oxford University Press.
- Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(5), 559–621.
- Kawakami, A., Furukawa, K., Katahira, K., & Okanoya, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology*, 4(311), 1–15.
- Kivy, P. (1990). How music moves. In *Music alone: Philosophical reflections on the purely musical experience* (pp. 146–172). Cornell University Press.
- Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences*, 14(3), 131–137.
- Konečni, V. J. (2008). Does music induce emotion? a theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 2(2), 115–129.
- Kreutz, G., & Lotze, M. (2007). Neuroscience of music and emotion. In F. Rauscher & W. Gruhn (Eds.), *Neurosciences in music pedagogy* (pp. 143–167). Nova Science Publishers.
- Lazarus, R. S. (1991). Appraisal. In *Emotion and adaptation* (pp. 133–151). Oxford University Press.
- Liljeström, S. (2011). *Emotional reactions to music: Prevalence and contributing factors* (Unpublished doctoral dissertation). Uppsala University, Department of Psychology.
- Livingstone, S. R., Mühlberger, R., Brown, A. R., & Thompson, W. F. (2010). Changing musical emotion: A computational rule system for modifying score and performance. *Computer Music Journal*, 34(1), 41–64.

- Oliveira, A. P., & Cardoso, A. (2008). Modeling affective content of music: A knowledge base approach. In *Proceedings of the 5th sound and music computing conference*.
- Panksepp, J., & Bernatzky, G. (2002). Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioral Processes*, *60*, 133–155.
- Peretz, I. (2010). Towards a neurobiology of musical emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: theory, research, applications* (pp. 99–126). Oxford University Press.
- Russell, J. A. (1989). Measures of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience, vol. 4* (pp. 83–111). Academic Press.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145–172.
- Schellenberg, E. G., Peretz, I., & Vieillard, S. (2008). Liking for happy- and sad-sounding music: Effects of exposure. *Cognition & Emotion*, *22*(2), 218–237.
- Scherer, K. R., & Zentner, K. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (p. 361-392). Oxford University Press.
- Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: A quantitative comparison by means of structural equation modeling. *European Journal of Personality*, *14*, 325–345.
- Schubert, E. (2013). Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives. *Frontiers in psychology*, *4*, 1–18.
- Vuoskoski, J. K., & Eerola, T. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, *15*(2), 159–173.
- Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception*, *29*(3), 311–317.
- Wallis, I., Ingalls, T., Campana, E., & Goodman, J. (2011). A rule-based generative music system controlled by desired valence and arousal. In *Proceedings of the sound and music computing conference*.
- Williams, D., Kirke, A., Miranda, E. R., Roesch, E. B., & Nasuto, S. J. (2013). Towards affective algorithmic composition. In G. Luck & O. Brabant (Eds.), *Proceedings of the 3rd international conference on music & emotion (ICME3)*. University of Jyväskylä, Department of Music.