Moody Music Generator: Characterising Control Parameters Using Crowdsourcing

Marco Scirea[1], Mark J. Nelson[2], and Julian Togelius[3]

[1] Center for Computer Games Research, IT University of Copenhagen, Denmark
`msci@itu.dk`
[2] Anadrome Research, Copenhagen, Denmark, `mjn@anadrome.org`
[3] Department of Computer Science and Engineering, New York University, NY, USA
`julian@togelius.com`

Abstract. We characterise the expressive effects of a music generator capable of varying its moods through two control parameters. The two control parameters were constructed on the basis of existing work on valence and arousal in music, and intended to provide control over those two mood factors. In this paper we conduct a listener study to determine how people actually perceive the various moods the generator can produce. Rather than directly attempting to validate that our two control parameters represent arousal and valence, instead we conduct an open-ended study to crowd-source labels characterising different parts of this two-dimensional control space. Our aim is to characterise perception of the generator's expressive space, without constraining listeners' responses to labels specifically aimed at validating the original arousal/valence motivation. Subjects were asked to listen to clips of generated music over the Internet, and to describe the moods with free-text labels. We find that the arousal parameter does roughly map to perceived arousal, but that the nominal "valence" parameter has strong interaction with the arousal parameter, and produces different effects in different parts of the control space. We believe that the characterisation methodology described here is general and could be used to map the expressive range of other parameterisable generators.

1 Introduction

Music has the power to evoke moods and emotions—even music generated by an algorithm. In fact, in many cases the whole purpose of a music generation algorithm is to evoke a particular mood. This is particularly true of music generators that form part of highly interactive systems such as games, where a common goal of dynamic music systems is to elicit a particular mood from the user on demand, as suits the current state of the system. To take the example of a video game, the music generation could be seen as content within the experience-driven procedural content generation framework [1], where the game adaptation mechanism generates music with a particular mood in response to player actions.

To enable such a capability we need a music generator that can take mood-related parameters and output music that elicits moods that as closely as possible correspond to what was specified by the parameters. This, however, is not a trivial task. While there are a number of features of music that are known to elicit particular moods and that can be incorporated into a music generation, the interplay between these features is complex and it is not clear that any particular change of the generator parameter will have a particular effect.

This paper describes a music generator which is parameterisable in mood space, and a validation of this generator through crowdsourcing. The generator is parameterisable along the two axes of valence and arousal, and uses Pure Data and stochastic processes to produce music. The validation involves letting hundreds of subjects listen to music clips produced by the generator, and express the moods they perceive in the music. Subjects were allowed to express their experiences in free text, which was then preprocessed and plotted on chart to see how particular mood expressions are centred on different on parts of the intended valence–arousal axis. In the following, we will first situate this research with respect to other research on music generation and mood expression, and then describe the music generator and the validation methodology.

2 Background

2.1 Procedurally generated music

Procedural generation of music for games is a broad field. While a good number of games use some sort of procedural music structure, there are different approaches (or degrees), as suggested by Wooller *et al.*: *transformational* algorithms and *generative* algorithms [2].

Transformational algorithms act upon an already prepared structure, for example by having the music recorded in layers that can be added or subtracted at a specific time to change the feel of the music (e.g., *The Legend of Zelda: Ocarina of Time* is one of the earliest games that used this approach). Note that this is only an example and there are a great number of transformational approaches, but we won't discuss them in this paper.

Generative algorithms instead create the musical structure themselves, which leads to a higher degree of difficulty in having the music stay consistent with the game events and generally requires more computing power as the musical materials have to be created on the fly. An example of this approach can be found in *Spore*: the music written by Brian Eno was created with Pure Data in the form of many small samples that created the soundtrack in real time.

Adopting the latter approach, we present generative procedural music generation in games for emotional expression. While the topics of affect [3], semiotics [4] and mood-tagging [5] are also interesting and significant, our focus lies in the **real-time generation of background music able to express moods**.

2.2 Emotions and moods

The topic of emotions has been extensively researched in the field of psychology, although their nature (and what constitutes the basic set of emotions) is still controversial. Lazarus argues that *"emotion is often associated and considered reciprocally influential with mood, temperament, personality, disposition, and motivation"* [6].

Affect is generally considered to be the experience of feeling or emotion. It is largely believed that affect is post-cognitive; emotion arises only after an amount of cognitive processing of information has been accomplished. With this assumption every affective reaction (e.g., pleasure, displeasure, liking, disliking) results from *"a prior cognitive process that makes a variety of content discriminations and identifies features, examines them to find value, and weighs them according to their contributions"* [7]. Another view is that affect can be both pre- and post-cognitive (notably Lerner and Keltner [8]); thoughts are created by an initial emotional response which then leads to producing affect.

Mood is an affective state. However, while an emotion generally has a specific object of focus, moods tends to be more unfocused and diffused [9]. Batson, Shaw, and Oleson say that mood *"involves tone and intensity and a structured set of beliefs about general expectations of a future experience of pleasure or pain, or of positive or negative affect in the future"* [10]. Another important difference between emotions and moods is that moods, being diffused and unfocused, can last much longer (as also remarked by Beedie *et al.* [11]).

In this paper, we focus on moods instead of emotions, for we expect that in games—where the player listens to the background music for a longer time duration than the duration that a particular emotion is experienced— moods are more likely to be remembered by the players after their gameplay. In addition, they are easier for game designers to integrate, since they represent longer-duration sentiment suitable for segments of game play.

2.3 Music mood taxonomy

The set of adjectives that describe music mood and emotional response is immense and there is no accepted standard. For example in the work by Katayose *et al.*[12], the emotional adjective set includes *Gloomy, Serious, Pathetic* and *Urbane.*

Russell [13] proposed a model of affect based on two bipolar dimensions: *pleasant-unpleasant* and *arousal-sleepy*, theorising that each affect word can be mapped into this bi-dimensional space by a combination of these two components. Thayer [14] applied Russell's model to music using as the dimensions of *stress* and *valence*; although the names of the dimensions are different from Russell's their meaning is the same. Also, we find different names in different research while the semantic meanings are identical. We will use the terms *valence* and *arousal*, as they are the most commonly used affective computing research.

Then the affect in music can be divided into the four clusters based on the dimensions of valence and arousal: **Anxious/Frantic** (Low Valence, High
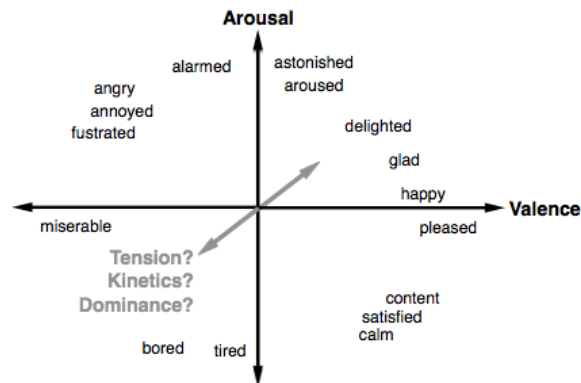
Fig. 1. The Valence-Arousal space, labelled by Russel's direct circular projection of adjectives [13]. Includes semantic of projected third affect dimensions: "tension", "kinetics", "dominance". In our study we haven't considered this third dimension as it's still not very defined.

Arousal), **Depression** (Low Valence, Low Arousal), **Contentment** (High Valence, Low Arousal) and **Exuberance** (High Valence, High Arousal). These four clusters have the advantage of being explicit and discriminable; also they are the basic music-induced moods [15],[16].

3   The generator

To generate music for our study we employed a real-time procedural music generator that we developed using the PD (Pure Data) programming language. Our music generation approach does not take into account chord sequences, leitmotifs, or improvisation. Instead, we aim to create a very minimalistic ambient music created by simple random number generators. This allows us to test our hypothesis of being able to display moods through only the manipulation of the mood defining features we consider.

3.1   Musical mood features

In order to generate mood-based music, we used four musical features intensity, timbre, rhythm, and dissonances, which are mainly inspired by Liu *et al.* [17]. While Liu et al.'s research focused on mood information extraction, we applied their approaches to generate music instead. This section extends our previous approach [18], introducing a new feature called dissonances.

Intensity Intensity is defined by how strong the volume of the music is. It is an arousal-dependent feature: high arousal corresponds to high intensity; low

arousal to low intensity. Intuitively the more stress is present in the music, the more it will have a high volume. Calm pieces of music, in a similar manner, have a lower one.

Timbre Timbre is what we could call the brightness of the music, that is, how much of the audio signal is composed by bass frequencies. It is often associated with *"how pleasing to listeners"* it appears [19]. In previous literature audio features such as MFCC (Mel-Frequency Cepstral Coefficients) and spectral shape features have been used to analyse this timbral feature.

We associated this timbral feature with valence: the more positive the valence, the higher will the timbre be. The brightness of Exuberance music, for example, is generally higher than that of music in Depression, which will result in greater spectral energy in the high sub bands for Exuberance.

Generally, timbre is a factor that is very dependent on the instrumentation choice. In our case we acted on the synthesisers, our instruments, to generate brighter and darker sounds. In our generator we had three different sets of instruments (which are actually the same synthesisers with different settings to make them sound different) for high, low and neutral valence.

Rhythm We included three features related to rhythm: strength, regularity and tempo [17].

- Rhythm strength: how prominent the rhythmic section is (drums and bass). This feature is arousal dependent.
- Regularity: how steady the rhythm is. This feature is valence dependent.
- Tempo: how fast the rhythm is. This feature is arousal dependent. In a high valence/high arousal piece of music, for instance, we can observe that the rhythm is strong and steady. In a low valence/low arousal, on the other hand, the tempo is slow and the rhythm cannot be as easily recognised.

We acted on these features in different ways. To influence rhythm strength, we changed how much the drums and the synthesiser which plays the bass are prominent in the music. Having the instruments play notes on the beat or the upbeat creates different feelings of regularity and irregularity. For example, in Contentment music, we favoured a steady rhythm with notes falling on the beats of the measure. In Depression music, on the other hand, we gave more space to upbeat notes. Finally, to influence the tempo we just acted on the BPMs (Beats Per Minute) of the music.

Dissonances What we mean by dissonance is the juxtaposition of two notes very close to each other: for example C and C♯. The distance between these two is just a semitone, which gives the listener a generally unpleasant sensation.

Dissonance doesn't mean that it always sounds bad. In fact most music pieces contain dissonances, as they can be used as cues expressing something amiss. The listener's ear can also be trained to accept dissonances through repetition.

In general, the bigger the interval between the two dissonant notes, the easier it is on the listener's ear: a C and a C♯ are always dissonant, but the dissonance is more evident if the notes are played from the same octave and not on two different ones.

C.P.E. Bach, in his *Essay on the True Art of Playing Keyboard Instruments* [20], remarks on the affective power of dissonances, although in a more general way: *"... dissonances are played loudly and consonances softly, since the former rouse our emotions and the latter quiet them"*.

Meyer [21] observes that the affect-arousing role of dissonances is evident in the practise of composers as well as in the writings of theorists and critics, remarking how the affective response is not only dependent on the presence of dissonances per se, but also upon conventional association. This means that depending on the conventions of the musical style dissonances might be more or less accepted by the listener and so can arouse different affective reactions. A study on listening preferences on infants conducted by Trainor and Heinmiller [22] shows how even these listeners, which have no knowledge of the musical scale structure, have an affective preference for consonance. Considering that our generator doesn't emulate any musical style, but creates minimalistic, unstructured music, we believe the effect of dissonances would follow this instinctive affective reaction.

Already in our first study we noticed that these features, originally devised to extract mood information, were enough to generate different moods. But we also realised that we could strengthen the impression by introducing dissonances in the music: for Exuberance and Contentment we use a diatonic scale, while for Anxious and Depression an altered one. We believe this is an important feature that cannot be ignored when wanting to show more precise moods in music.

Dissonance feature is valence depending. In our study we just used two scales: a C major scale (C D E F G A B) for positive and a E♭ Harmonic Minor scale minus the third grade (E♭ F [G♭] Ab B♭ B D) for negative valence. Music built on a minor scales is generally considered more sombre than when made in a major key. This is not technically correct in our system because it would require a grade of organisation and harmony that would make plain which is the root note. The notes of the harmonic minor scale are the same as the natural minor except that the seventh degree is raised by one semitone, making an augmented second between the sixth and seventh degrees. For our unstructured music this means that we have a whole-and-a-half interval between B and D and two half intervals (D-E♭ and B♭-B). The removal of the third grade (G♭) makes even more difficult to the listener's ear to identify the key, effectively making the dissonances sound as such.

4   Experiment design

As described in the previous section, we produced a generator intended to be parameterised by two control axes: arousal and valence. Although this construction is based on theoretical motivations and existing work on the relationship

between musical parameters and perceived mood, it does not necessarily follow that these axes *actually* represent arousal and valence. To understand what kind of generative space our music generator actually produces, we designed a study to characterise how the two control axes of our generator influence listeners' perceptions.

Contrary to our previous pilot study [23], we employed a mix of closed-ended questions to validate the axes (e.g. a number of mood expressing words and a Likert scale for valence and arousal), we decided to provide completely open-ended questions to the participants, so as to eliminate as much bias as possible from their answers, and understand the effects of our generator's control parameters in an open-ended way.

We developed the online survey with HTML and PHP, using a MySQL database to hold the data collected; the participants were presented with a page consisting of a music clip and five blank boxes where they were asked to write emotional words that they thought the music expressed.

After each five responses we introduced a special page where the participants could review their answers, listen to the previous five clips again and see some of the most recent answers from other users for the same clips. We created this page to give feedback to the users and to make the survey, hopefully, more interesting for them by giving them the opportunity to confront their answers with the ones other users provided.

The experiment has no pre-defined end: the user is able to continue answering until he/she wants, and can close the online survey at any time.

4.1 Music clip generation

We generated 100 clips of 30 seconds of music using our music generator, each of these expressing a randomly chosen point in the bi-dimensional mood space we described in section 2.3.

The music clips have been generated by linearly connecting the features and the respective axis, even though we are conscious that the relationships are probably more complex; in fact we hope the data collected through this study will help us better define these.

The maximum and minimum bounds we gave to the various musical features were:

– **Tempo**: 100-136
– **Intensity**:

	Synth 1	Synth 2	Synth 3 (Bass)	Drum machine
Minimum values	69	56	35	60
Maximum values	98	119	83	128

– **Rhythm strength**: -20% to +20% intensity to *Synth 3* and *Drum machine*.
– **Timbre**: three different settings for the synthesisers: the lowest is selected when the valence is less or equal of 33, the middle between 33 and 66 and the higher above 66 (valence goes from 0 to 100 in our system).

– **Steadiness**: three settings dependent on the valence axis as the Timbre: *steady*, *medium_steady* and *unsteady*. On the *steady* rhythm all the notes fall on the beat of the measure, on the *medium steady* rhythm all instruments play notes on the beat, while the drum machine plays off-beat. Finally for the *unsteady* rhythm only Synth 1 (which is the higher pitched instrument, and the one more resembling a lead voice) plays on the beat while all the other instruments play offbeat. Note that if all the instruments played on the offbeat the listener would have no way of telling the beat from the offbeat, effectively perceiving a steady rhythm.
– **Dissonances**: as discussed in section 3.1, we use a C major scale (C D E F G A B) for positive and an E♭ harmonic minor scale minus the third grade (E♭ F [G♭] Ab B♭ B D) for negative valence.

5 Results and analysis

We collected a total of 2020 free-text labels from 192 distinct users. We can consider patterns in these labels to constitute an open-ended, nonparametric characterisation of how users perceive the music's mood as we vary the control parameters intended to represent arousal and valence. The obvious question is then: are there any patterns, and do they provide any insight into the effects of these control parameters? With free-text labels, it is not entirely implausible that there could end up being no easily discernible patterns in the data. However there turn out to be some strongly localisable responses, particularly among the labels volunteered relatively frequently. Although listeners could in principle respond with any English word or phrase, some words recur often, e.g. "mysterious" was volunteered 34 times.

In order to characterise the control parameters using these labels from the users themselves, for each label we calculate the average (mean) arousal and valence of the clips for which that label was volunteered. The goal of doing so is to localise the label somewhere in the two-dimensional control space. We would like to say things such as: the label "rushed" appears on average in the high-arousal, high-valence part of the space, while the label "relaxed" is given on average to low-arousal, low-valence clips.

Of course, if a given label was only volunteered a few times, an average is not very reliable. Therefore we choose the 20 labels which are *best localised*, in the sense that we have enough data to more reliably determine their average location in the control space. To determine how well localised a label is, we rank labels by the standard error of their mean location on the arousal/valence axes (standard errors summed over both axes). The standard error of a mean, $se = \frac{stddev}{\sqrt{n}}$, will in general be lower for labels with lower sample standard deviations, and for labels which appear more times in the data set.[4]

---

[4] Since we make no assumption about the distribution of data, we can't use the standard error as a basis for a confidence interval. Nonetheless, it is useful as a proxy for how well we can localise a label in the arousal/valence space, relative to other labels in our data set.
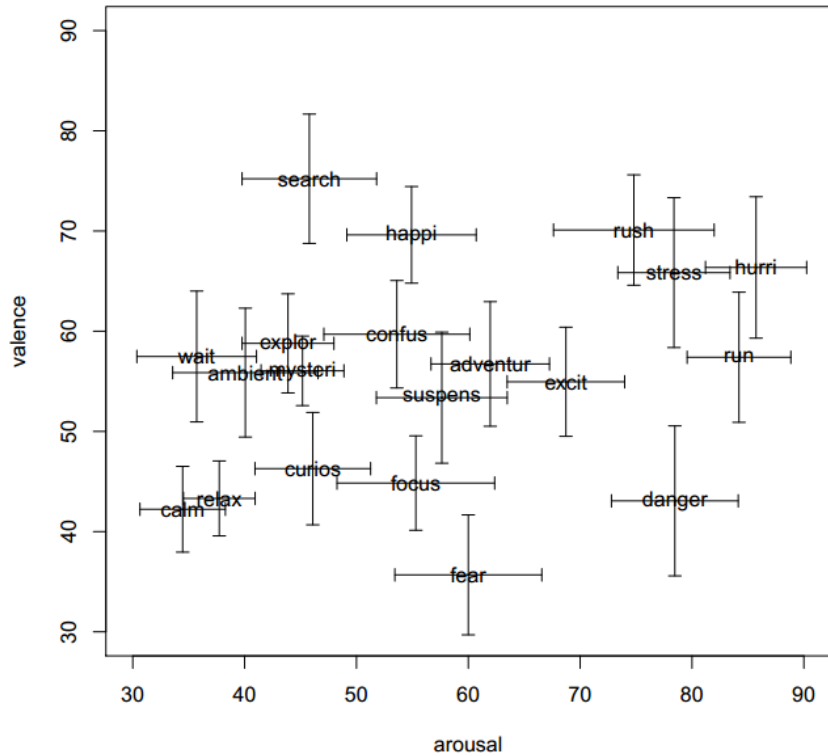
Fig. 2. Free-text, crowdsourced characterization of moods across the generator's two-dimensional control space. Plotted labels are the 20 best-localised labels (post-stemming), plotted at the average location for which they were volunteered as labels. Error bars represent the standard error of the mean.

Before ranking labels by standard error, we perform two preprocessing steps on the data. First, we stem the words using the Snowball stemmer,[5] in order to aggregate minor part-of-speech variations of labels—for example, *relaxed* and *relaxing* are both mapped to the stem *relax*. In addition, we exclude labels that appear fewer than 5 times in the data set even after stemming, because the sample standard deviation is an unreliable measure for extremely small n.

Figure 2 plots the 20 best-localised label stems, at the average location of the (arousal, valence) parameter settings that elicited that label as a response. The standard errors of the mean are plotted as indicative error bars. This plot alone is surprisingly informative as a characterisation of the control parameters' effect on perception of musical moods. Especially considering that users volunteered free-text labels rather than selecting categories, the trends in the axes are rather striking.

---

[5] http://snowball.tartarus.org/

We can make a few qualitative observations on the basis of these 20 well-localised labels. In a general sense, the "arousal" and "valence" theory that drove the development of our control axes does not seem to precisely align, in this setting, with the effect of the axes to which we've nominally given those labels, though arousal is closer than valence.

Arousal maps to something like a calm vs. stress axis (which is, in fact, the definition of arousal). Low-arousal words include (unstemmed): curious, waiting, calm, relaxing, ambient. High-arousal words include: rushed, stressed, hurried, run, dangerous. Valence seems to be largely dominated by arousal, but modifies it in a way that has strong interaction effects.

Looking at high-valence clips, when coupled with a high arousal setting, they elicit labels that accentuate a feeling of being rushed: rushed, stressed, hurried. So, raising valence, rather than being perceived as positive valence, instead emphasises a kind of speed in the context of high arousal, with even a somewhat negative sentiment. Low-valence, high-arousal clips are most often labelled as "danger" instead.

With a mid-level arousal setting, valence does seem to act as a relatively straightforward valence setting: high-valence clips are characterised by "happy", and low-valence clips by "fear". As arousal drops, however, the effect of the "valence" setting becomes particularly inconsistent with the intent that it be a knob used to vary perceived valence. The nominally low-valence clips, when coupled with low arousal, elicit quite *positive* labels: calm, relax. As "valence" increases while arousal stays low, the main effect is to get somewhat more active: from "calm" to "wait" and "explore" at mid-valence, to "search" at high valence settings.

6 Conclusion and future work

The goal of this study was an open-ended characterisation of the perceived affects produced by a music generator parameterised by a two-dimensional valence/arousal control axis. We crowdsourced labels for clips in randomly selected parts of the control space, and looked for labels that are *well-localized*, i.e. where they are volunteered by users most often for specific parts of the control space. Although the generator was designed with a valence/arousal control axis, the purpose of this study was simply to characterise what perceived effect in listeners our two-dimensional control axis actually has, without regard (in the study design or in communication to study participants) for where these two axes came from or what they were supposed to represent.

We designed this experiment as to avoid as much bias from our part as possible by having the participants answer to completely open-ended questions. Overall, we collected a total of 2020 free-text labels from 192 distinct users. We found some interesting results from this study that seem to show how we can express moods through the manipulation of our musical features which can be recognised from the listener relatively correctly in our mood space.

The results are very promising regarding the recognition of the arousal axis, which seems to map very well to the answers provided by the participants. The valence axis, on the other hand, behaves as intended only on medium arousal values, becoming more confused—and exhibiting interesting interplay with the arousal axis—when approaching the extremes of the arousal axis (both positive and negative). This result is in line with previous studies that show valence is harder to characterize [24], although we notice that in our final data sample the emotion words that express positive/negative affect are fewer in number than the ones expressing arousal, suggesting more study of the interaction between the two axes may be useful. We also notice a slight bias towards more positive valence and towards higher arousal. This is unexpected, but might be symptom of an intrinsic bias caused by the music produced by the generator (the sounds and unstructured nature of the music).

An interesting avenue for future work is to investigate the degree to which cultural background of respondents may influence the labels volunteered. Although we describe the study as open-ended in an attempt to be unbiased, the presence of certain frequent words such as *explore* is intriguing; they can be seen as somewhat game-related, whereas this study did not involve games at all (though they are a future application of ours). This may be simply due to the shared background and therefore shared vocabulary and shared attention to salient features of respondents in this study. We plan to conduct a study with respondents from different cultural backgrounds to investigate this issue.

On the application side, we plan to connect this generator with the experience-driven Super Mario level generator described by Shaker *et al.* [25], in order to explore whether we can observe a significant difference in the players' emotional responses through the use of music that reinforces or diverges from the intended emotions elicited by the level generator. Finally we will continue our work in applying this research on expressing narrative clues through music [26].

References

1. Yannakakis, G.N., Togelius, J.: Experience-driven procedural content generation. IEEE Transactions on Affective Computing **2**(3) (2011) 147–161
2. Wooller, R., Brown, A.R., Miranda, E., Diederich, J., Berry, R.: A framework for comparison of process in algorithmic music systems. In: Generative Arts Practice 2005 — A Creativity & Cognition Symposium. (2005)
3. Birchfield, D.: Generative model for the creation of musical emotion, meaning, and form. In: Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence. (2003) 99–104
4. Eladhari, M., Nieuwdorp, R., Fridenfalk, M.: The soundtrack of your mind: mind music-adaptive audio for game characters. In: Proceedings of Advances in Computer Entertainment Technology. (2006)
5. Livingstone, S.R., Brown, A.R.: Dynamic response: Real-time adaptation for music emotion. In: Proceedings of the 2nd Australasian Conference on Interactive Entertainment. (2005) 105–111
6. Lazarus, R.S.: Emotion and Adaptation. Oxford University Press (1991)

7. Brewin, C.R.: Cognitive change processes in psychotherapy. Psychological Review **96**(3) (1989) 379–394
8. Lerner, J.S., Keltner, D.: Beyond valence: Toward a model of emotion-specific influences on judgement and choice. Cognition & Emotion **14**(4) (2000) 473–493
9. Martin, B.A.: The influence of gender on mood effects in advertising. Psychology & Marketing **20**(3) (2003) 249–273
10. Batson, C.D., Shaw, L.L., Oleson, K.C.: Differentiating affect, mood, and emotion: Toward functionally based conceptual distinctions. (1992)
11. Beedie, C., Terry, P., Lane, A.: Distinctions between emotion and mood. Cognition & Emotion **19**(6) (2005) 847–878
12. Katayose, H., Imai, M., Inokuchi, S.: Sentiment extraction in music. In: Proceedings of the 9th International Conference on Pattern Recognition. (1988) 1083–1087
13. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology **39**(6) (1980) 1161–1178
14. Thayer, R.E.: The Biopsychology of Mood and Arousal. Oxford University Press (1989)
15. Kreutz, G., Ott, U., Teichmann, D., Osawa, P., Vaitl, D.: Using music to induce emotions: Influences of musical preference and absorption. Psychology of Music **36**(1) (2008) 101–126
16. Lindström, E., Juslin, P.N., Bresin, R., Williamon, A.: "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. Research Studies in Music Education **20**(1) (2003) 23–47
17. Liu, D., Lu, L., Zhang, H.J.: Automatic mood detection from acoustic music data. In: Proceedings of the International Symposium on Music Information Retrieval. (2003) 81–87
18. Scirea, M.: Mood dependent music generator. In: Proceedings of Advances in Computer Entertainment. (2013) 626–629
19. Aucouturier, J.J., Pachet, F., Sandler, M.: "The way it sounds": Timbre models for analysis and retrieval of music signals. IEEE Transactions on Multimedia **7**(6) (2005) 1028–1035
20. Bach, C.P.E., Mitchell, W.J., John, W.: Essay on the true art of playing keyboard instruments. WW Norton (1949)
21. Meyer, L.B.: Emotion and meaning in music. University of Chicago Press (2008)
22. Trainor, L.J., Heinmiller, B.M.: The development of evaluative responses to music:: Infants prefer to listen to consonance over dissonance. Infant Behavior and Development **21**(1) (1998) 77–88
23. Scirea, M., Cheong, Y.G., Bae, B.C.: Mood expression in real-time computer generated music using pure data. In: Proceedings of the International Conference on Music Perception and Cognition. (2014)
24. Livingstone, S.R., Brown, A.R., Muhlberger, R.: Influencing the perceived emotions of music with intent. In: Proceedings of the Third International Conference on Generative Systems in the Electronic Arts. (2005)
25. Shaker, N., Yannakakis, G.N., Togelius, J.: Towards automatic personalized content generation for platform games. In: Proceedings of the 2010 Conference on Artificial Intelligence and Interactive Digital Entertainment. (2010)
26. Scirea, M., Cheong, Y.G., Bae, B.C., Nelson, M.J.: Evaluating musical foreshadowing of videogame narrative experiences. In: Proceedings of Audio Mostly 2014. (2014)