

How useful are specific musical cues at expressing Anger?

Riccardo Boiteux
Durham University

ABSTRACT

Music is made up of numerous parameters, such as tempo and timbre, which are each manipulated in various ways to create unique pieces of music. This study uses 6 short instrumental extracts as it seeks to investigate the degree of usefulness of a chosen set of musical cues in pieces which aim to convey the discrete emotion 'Anger'. 30 participants took part in an online survey in which they were asked to rate the usefulness of 5 separate cues (*Timbre, Mode, Tempo, Sound Level* and *Articulation*) using adjective scales, as they listened to 6 short stimuli. The second part to each question asked the participants to briefly describe any of the musical elements which particularly impacted them in an effort to see if the qualitative data from the description yielded results which supported or contradicted the previously asked rating scale question. The hypothesis of this project was that tempo and mode would be the 2 types of cues most likely to be considered most useful during the expression of 'Anger'. However, this study's data demonstrate that there are no significant differences between the relative importance of the 5 cue types in question. Lastly, this study discusses some of its limitations as well as proposing certain manners in which its structure could be improved to ask further relevant questions regarding this topic.

1. INTRODUCTION

Music can be composed to convey a multitude of discrete emotions. Although pieces usually sound completely individualistic, they all share the fact that they are simply created through their unique manipulation of a finite set of musical parameters. Oftentimes, these manipulations are implemented to express specific emotions, or to make listeners feel a certain way. In the first case, music acts "as an object for perception and reflection," in the other, music is said to induce emotion if "it has caused an emotional reaction" (Gabrielsson, 2002: 124). This study focuses on emotion which is perceived through music and looks at how useful some of the aforementioned musical parameters are at supporting this expression.

Previous Research. Primarily, Tuomas Eerola and Jonna K. Vuoskoski's paper was closely considered. Their primary aim hoped to "systematically compare perceived emotions in music using two different theoretical frameworks: the discrete emotion model, and the dimensional model of affect" (Eerola and Vuoskoski, 2011: 18). However, their second objective, which was to establish "a new, improved set of stimuli for the study of music-mediated emotions" (Ibid., 18) is most applicable for this study. Their list of 110 short film music excerpts which provide "representative examples of five discrete emotions (anger, fear, sadness, happiness and tenderness)" (Ibid., 18), has in fact been crucial to the study at hand. That is because, as well as being of sensible duration, these excerpts are also clearly live recordings.

A second study's captivating intention was to "decode the emotional expression of musical performances" (Juslin, 1997: 229). In his paper, Patrick N. Juslin used "synthesized performances of a short melody" to carry out 2 experiments (Ibid., 225). Before detailing these experiments though, Juslin takes a moment to explain his modified version of Egon Brunswik's (1956) 'Lens Model' (Figure 1). He describes the 'functional achievement' (*ra*) as being how successfully a "performer is able to communicate a certain emotion to a listener", and the 'functional validity' (*rf*) as "the extent to which each cue is actually *used* by the listener" (Juslin, 1997: 227-228). To find the functional validities of each cue, Juslin made sure to use "performances that are systematically varied with regard to the relevant cues" (Ibid., 228). In his first experiment, Juslin tried "to recreate representative cue profiles of five emotional expressions: *happiness, sadness, anger, articulation*, and *tenderness*" (Ibid., 225). He then required the participants to decide on what "the intended emotional expression" of the melody was by manipulating the following cues: timing, vibrato, attack, articulation, spectrum, sound level and tempo (Ibid., 225). During the second experiment, Juslin focused on attack, articulation, spectrum, tempo, and sound level, where he "systematically varied [them] in a factorial design" for the listeners to "describe the emotional expression of each cue combination by ratings on adjective scales" (Ibid., 225). An interesting result from this second experiment was that "the predictive strength of each cue varied a lot depending on the emotion judged" (Ibid., 225). These differing results influenced by the motion in question are highly intriguing and should be considered for further study, however in the present paper, only discrete 'Anger' is being examined.

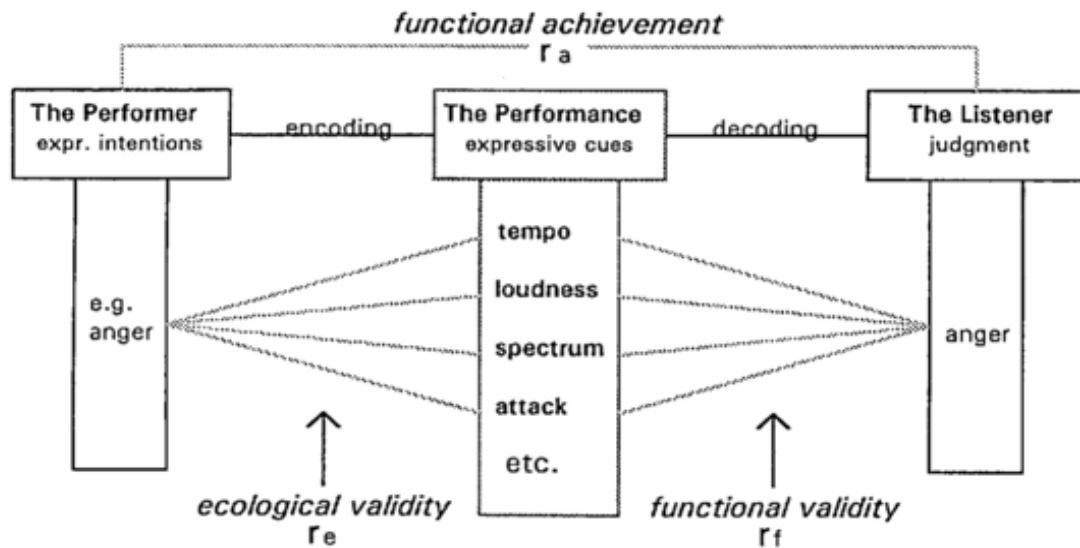


Figure 1: A modified lens model for emotional communication in music performance.

Figure 1. (Juslin 1997)

Furthermore, Juslin's study has a clear limitation; a lack of 'prosodic contours' in the stimuli used. These such contours are delicate fluctuations in the dynamics and speed of the music which overlay a live performance, giving it a human touch (Ibid., 231). These are evidently missing due to the fact that the stimuli used during the experiments are synthesised. Furthermore, even though in his conclusion he explains that he doubts that "more "authentic-sounding" performances would have produced drastically different results" (Ibid., 247), the present study makes it a point to use recordings of real musicians to make sure that this doubt is nowhere in sight.

A final observation to make is that Figure 2 displays a newer 'Lens Model' which was published in 2010. Interestingly, the 'Extended Lens Model' (ELM) has the added complexity of more cues as well as the fact that these cues are split between a composer and performer (Juslin and Lindström, 2010: 338). This ELM figure was very helpful when choosing which cues to focus on in today's paper.

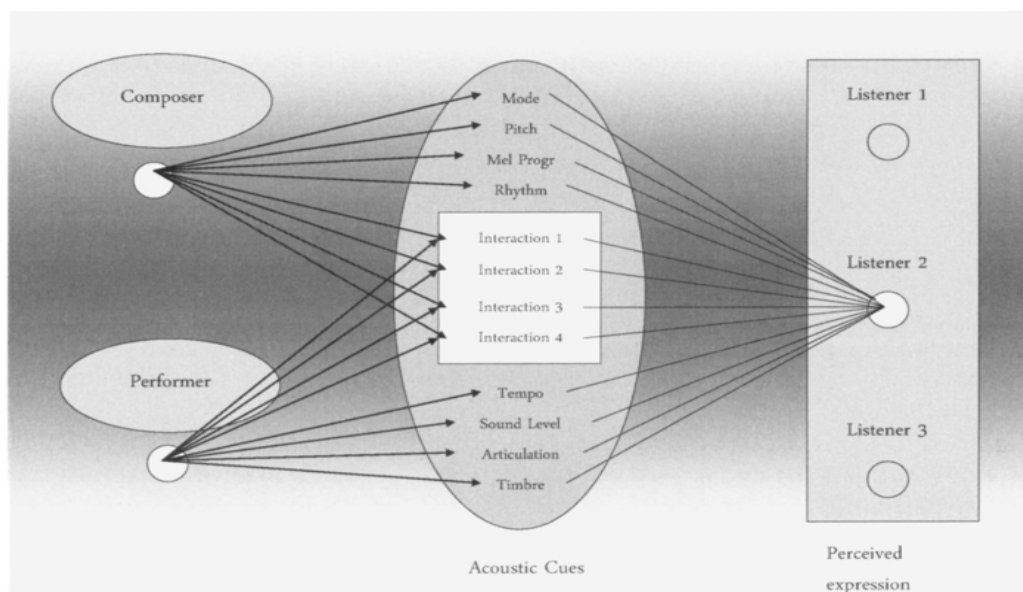


Figure 2. (Juslin and Lindström 2010)

Hypothesis and Research Question. Whilst defining what the features that make a study work as a ‘multiple regression model’ in his article ‘*The psychology of linear judgement models*’, Berndt Brehmer lists a few questions that the study should be asking about judgement, two of which are pertaining to this study. The first is whether one wants to find out which of the “cues are used by the judge”, and the second being if the study is looking to answer what “the relative importance of the different cues” are (Brehmer, 1994: 142). Taking this into consideration, the aim of this research project is to put into practice a ‘multiple regression model’ similar to the one present in Juslin’s 1997 study by using relevant film music stimuli introduced by Eerola and Vuoskoski in 2011.

This study will focus on 6 music excerpts, from the list made by Eerola and Vuoskoski, which they have said express ‘Anger’. The participants will be asked to rate how useful the 5 musical parameters in question (*Timbre, Mode, Tempo, Sound Level* and *Articulation*) are at conveying this emotion. The primary aim of this study is to scrutinise the data regarding the relative importance of the cues across the whole data set. The secondary aim is to see if there is anything noteworthy to report regarding the difference in answers between ‘Musicians’ and ‘Non-Musicians’.

Simone Dalla Bella and colleagues sought to investigate whether both adults and children utilise the same musical properties when attempting to recognise if music sounds sad or happy (Dalla Bella et al., 2011, B1). Their results showed that “5-year-old children exclusively used information about tempo,” but that 6-8-year-old children were able to use both mode and tempo (Ibid., B8). Moreover, Gregory, Worrall, and Sarge sought to see if they could “support the idea of a learned association between mode and emotional response” (Gregory, Worrall and Sarge, 1996: 341). This study’s results demonstrated that children who were “7 to 8 years old showed a significant major-happy and minor-sad connotation”, but that “3 to 4 year-olds did not show any such significant association” (Ibid., 341).

Taking these findings into account, where very young children tend to be able to use both mode and tempo to perceive emotion in music, the present study’s hypothesis is that mode and tempo will be the most useful cue in conveying ‘Anger’. This is because there could be a link between how long participants have been able to understand these musical parameters for, with how comfortable they are at using them as a tool to discern emotion in music. Hence, they may perhaps find them more useful to aid in their perception of the relevant emotion.

2. METHOD

Design. This experiment took the shape of a survey created using *Qualtrics* and was shared with friends and family. The survey’s information sheet gave a brief explanation regarding what the survey would entail as well as letting potential participants know that it would take around 20 minutes to finish. It also shared contact details that participants might need if they wanted clarifications concerning any part of the survey. Finally, before beginning the survey, informed consent was acquired.

Participants. Results from 30 participants, aged 18 to 73 were collected. ($M = 37.1$, $SD = 15.6$). Of these individuals, 12 (40%) were male and 18 (60%) were female. The participants were further divided between ‘Musicians’ and ‘Non-Musicians’, where 18 (60%) were ‘Musicians’ and 12 (40%) were not. No further demographic background such as location and cultural background was collected. In this experiment, ‘Musicians’ were defined as those that knew what each of the cues (*Timbre, Mode, Tempo, Sound Level* and *Articulation*) meant. On the other hand, if the meaning of any or all of the cues were previously not understood, participants were asked to describe themselves as ‘Non-musicians’. Finally, participants were not remunerated for completing the survey.

Materials and Stimuli. As previously mentioned, this experiment uses 6 of the 10 film music excerpts which represent ‘Anger’ as established by Eerola and Vuoskoski in their 2011 paper. The 4 remaining excerpts were purely left out because they could not be found. The excerpts used lasted between 15 to 20 seconds and the fact that they were so short meant that participants did not have to struggle to find which section of the music conveyed ‘Anger’. Furthermore, participants were told that the music was supposed to represent ‘Anger’ because the participants were not being asked about *which* emotion was being expressed. Rather, for the purpose of this study, it was important to let them know they were listening to ‘Anger’ conveying extracts to make sure their answers were completely revolving around the expression of this one emotion. Aside from being of appropriate length, these extracts were also excellent because they indeed had ‘prosodic contours’ (Juslin, 1997: 231). Lastly, a small number of participants reported that they recognised some of the scenes that the film music excerpts accompanied, but this was not an issue for the study at hand. Seen as the participants were told that the music represented ‘Anger’, recognising the excerpts was not considered to be a problem.

Procedure. First, participants were asked if they were ‘Non-musicians’ or ‘Musicians’. Selecting ‘Non-musician’ would take the participant through a fast paced 5-part explanation series explaining the definition of the cues. Contrastingly, those that selected ‘Musician’ were taken straight to the first question of the experiment.

The explanation series was divided so that each section focused on one cue, and ‘Non-musicians’ would watch a short video which explained the definition of that specific cue. Upon having watched the video, they were asked to confirm that they understood what it meant and could only move onto the next part of the explanation series if they had clicked ‘Yes’. After having worked through all 5 parts of the series, participants were asked one last time to confirm that they had understood the signification of every single cue. Again, it was made very clear that it was compulsory for all individuals to understand each cue to advance in the survey. It is important to mention however, that this last page let the participants know that if they needed help clarifying anything at all, they could easily reach out for help using the contact information provided.

After listening to one stimulus, participants were first asked to rate how useful each cue was in allowing the music to convey ‘Anger’. Although the scales were only divided into 5 segments (*Not at all useful*, *Slightly useful*, *Moderately useful*, *Very useful* and *Extremely useful*), the survey design made it so that the scales were out of 100 points. This was implemented to avoid having issues if the results were densely packed. The continuous scale therefore had 5 quartile ranges in which ratings would classify, and these were evenly distributed along the scale with each option occupying 1/5 of the scale. It is also worth noting that the minimum value was 1, ensuring that each of the cues were always taken into consideration. If participants found a cue not to be necessary or useful in expressing ‘Anger’, moving the scale to 1 would clearly convey this opinion.

On the other hand, question 2 asked participants to describe, in a few words, any musical elements which they found to be distinctly salient in their perception of ‘Anger’. It was phrased to imply that participants could mention cues outside of the 5 originally presented on the scales (i.e. *pitch*, *rhythm* etc.) if they deemed it necessary to do so.

The objective for the inclusion of these 2 questions was to see if, even with the liberty to choose what cue to talk about, the data from both sets of results would somewhat complement each other. If not, it would simply be interesting to see how they contrasted to one another.

The 2 questions outlined above were repeated 6 times. Once for each short excerpt.

3. RESULTS

Due to the nature of the first question, a large number of values was collated into a smaller set of grand means and standard deviations. *T-tests* were then run using *RStudio* to check for any significant differences between the answers of ‘Musicians’ and ‘Non-musicians’. Similarly, the results from the second question were put through thematic analysis on *NVivo* to create a word cloud.

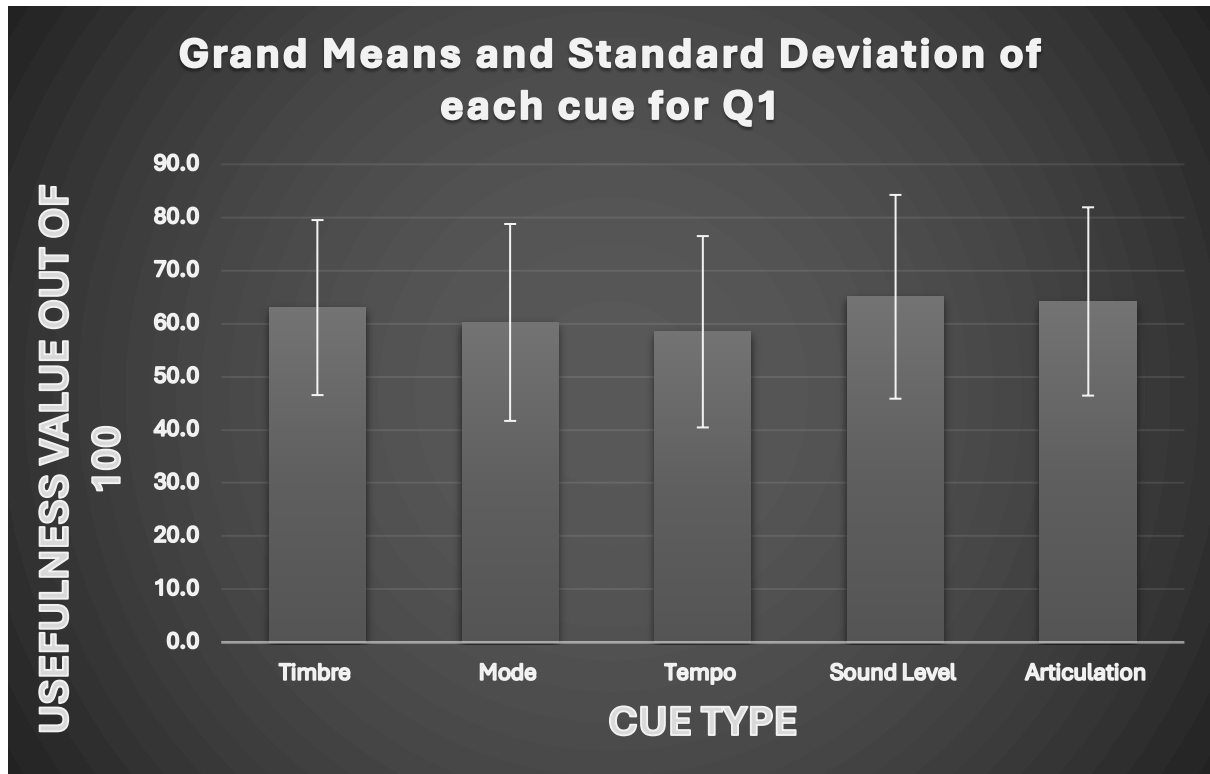


Figure 3. Bar chart showing results of Question 1 for all excerpts

Figure 3 displays the grand means as well as the standard deviations of the usefulness of each cue. Naturally, usefulness is a perceptive ability, however for the purpose of this study, if a cue was useful, it is because it helped the listener perceive the relevant emotion in the excerpt. As is strikingly obvious, the hypothesis was completely disproved. Tempo and mode have the lowest usefulness ratings out of the 5 cues. Tempo has the lowest usefulness rating ($M = 58.5$, $SD = 18.0$), and mode is in penultimate place ($M = 60.2$, $SD = 18.5$). Moving upwards, timbre is shown to be third ($M = 63.0$, $SD = 16.5$), whilst articulation is the second most useful ($M = 64.2$, $SD = 17.7$). Finally, the cue with the highest usefulness rating for the perception of 'Anger' is sound level ($M = 65.1$, $SD = 19.2$). Moreover, as can be seen by the standard deviations listed, these values are relatively quite low suggesting that the participants do not disagree with one another by very much. A final comment on this bar chart is that setting the number of points on the scales to 100 instead of 5 (one for each usefulness level), seems to have helped the clarity of the results.

t-test Musicianship	Timbre	Mode	Tempo	SoundLvl	Articulation
t	1,0817	0,7255	-1,1756	0,18897	-0,94399
p-value	0,2928	0,4749	0,2533	0,852	0,3563
t-test Gender	Timbre	Mode	Tempo	SoundLvl	Articulation
t	-1,1811	-0,50504	0,47662	-1,4983	-0,52004
p-value	0,2475	0,6181	0,6395	0,1457	0,6082

Figure 4. RStudio t-tests

Furthermore, a number of *t*-tests were completed using *RStudio*. The primary aim for these tests was to check if there was a significant difference between the results of the independent variables 'Musician' and 'Non-musician'. It goes without saying that the dependent variables in question were the 5 types of musical parameters. However, as shown in Figure 4, no $p < .05$, and thus there was no difference deemed significant. Disappointed by this outcome, this study attempted a second round of tests, this time making the independent variable the 'Gender' of the participants. Again, as shown in Figure 4, no p were found to be $< .05$ and were thus not significant.



Figure 5. NVivo word cloud showing the results of the thematic analysis of Question 2

The thematic analysis applied to the open-ended question has produced a somewhat contrasting ranking of usefulness compared to the ratings from the bar chart described above. Figure 5 displays a word cloud which has been generated by *NVivo*. Regarding results which support the bar chart, mode and tempo are again in last place with regards to how many times they were referenced. Although, they have swapped place with mode being in last place (32 references) and tempo in 4th (43 references). Hence, the thematic analysis supports the quantitative results which demonstrated that mode and tempo were the least useful at helping listeners perceive ‘Anger’. Moreover, the rankings of the top 3 cues also changed around. Although in the bar chart sound level and articulation boasted the highest 2 ratings, according to the open-ended question they are tied for second with regards to how many times their respective theme nodes were referenced (61 references). The most noteworthy detail of this second set of data, however, is the fact that timbre was mentioned the most times by participants (86 references) by a significant margin.

Additionally, as expected, participants mentioned the usefulness of 3 cues which were not part of the original 5 under investigation. These parameters were pitch and rhythm (9 references each) and melody (10 references). Although these 3 cue types, which are also present on the ‘Extended Lens Model’ (Juslin and Lindström, 2010: 338), were referenced remarkably fewer times than the original 5, these results are still of value. This is because it demonstrates that despite the survey never mentioning them, participants still chose to do so due to how valuable they were at helping the music convey ‘Anger’.

4. CONCLUSION AND DISCUSSION

Conclusion. The results of this study have rather clearly disproved the hypothesis which suggested that mode and tempo would be rated as the most useful cue for the perception of ‘Anger’. On the other hand, sound level, timbre and articulation were seen as more useful in both questions of the experiment, although there was no significant disparity between all of the ratings in the usefulness scale question. Contrastingly, the thematic analysis of the open-ended question showed a much sparser set of references for the cues, with timbre being referenced the most across all the participants. This overall sparsity was certainly due to the fact that participants were not forced to include all of the cues in their answer, and one could suggest that this actually made this second question the more valid of the 2. Moreover, one could propose that, since all of the stimuli employ the instrumentation of symphonic or philharmonic orchestras, it simply is easier for music to convey ‘Anger’ using expansive timbral palettes.

Regarding how the results from this study contribute to the current discourse, first we should look back at Juslin’s 1997 study.

In his introduction, Juslin displayed a table showing what some of the “expressive principles for different emotional expressions” are, including ‘Anger’ (Juslin, 1997: 226). In it, the table states that anger can be conveyed through the use of, for example, “fast tempo, high sound level” and a few other parameters (Ibid., 226). The hope for the current study is that it might add an extra layer of information to this table. Furthermore, if this experiment is expanded to search for the relative usefulness of the cue types in the other discrete emotions

mentioned in the table such as ‘Fear’ and ‘Sadness’, the added layer of information can be introduced to them to.

Additionally, in Eerola and Vuoskoski’s 2011 paper, it is suggested that “a thorough dissection of the acoustical and musical features of the stimuli should be carried out” (Eerola and Vuoskoski, 2011: 42). The hope of the present study is to have somewhat done this, by asking participants to concentrate on the usefulness of the different cues.

Limitations. The first limitation to mention is the sample size of this study. Due to the fact that this study was carried out in an academic context, only 30 participants took part. In a professional context, a much larger and therefore more reliable sample size would be employed for a more serviceable set of results. Again, regarding the context of this study, the participants completed the survey without receiving any payment. This can only have contributed negatively to the overall results due to the fact that it meant that for ‘Non-musicians’, there was not any real pressure to fully learn what the meanings of the musical cues were before beginning the experiment. Another impactful limitation specifically regards the open-ended question. Even though the question was designed so that both ‘Musicians’ and ‘Non-musicians’ were able to answer it, naturally ‘Musicians’ were able to provide a lot more detail. In a future iteration of this study, this limitation could perhaps be solved by creating two separate thematic analysis documents, one for either type of participant, and compare these. The final limitation regards Kate Hevner’s (1936) theory of ‘flexibility in expressiveness’ that suggests the manner in which meaning is conveyed through music depends on a large list of factors. Part of this list includes for example “the attitude of the listeners, their previous experience, ... and their momentary mood and physiological condition” (Hevner, 1936: 247). As clearly shown, the expression of meaning in music is an incredibly delicate and subjective topic which is very difficult to monitor. Unfortunately, these uncontrolled variables negatively impact the reliability of this study’s results.

Future Directions. Primarily, it would be helpful for this study to be revised and then repeated by tackling some of the avoidable limitations. The first limitation which should be fixed is the number of participants, the more there are the more reliable this study would be. The second fix would be to introduce a small payment for every participant, which would hopefully mean that all ‘Non-musician’ participants are enticed to learn the definitions of each cue. Furthermore, if the participants are being rewarded, the study could include the extra complication of the other 3 cue types present in the ‘Extended Lens Model’ (Juslin and Lindström, 2010: 338). That way, all cue types are clearly shown to be of equal relevance to participants. Finally, it would be quite interesting to see similarly structured studies which focus on the usefulness of the cues for other discrete emotions.

REFERENCES

- Brehmer, Berndt. 1994. ‘The psychology of linear judgement models’, *Acta Psychologica*, 87(2-3): 137-154.
- Dalla Bella, S., Peretz, I., Rousseau, L. & Gosselin, N. 2001. ‘A developmental study of the affective value of tempo and mode in music’, *Cognition*, 80(3): B1-B10.
- Eerola, T., & Vuoskoski, J.K. 2011. ‘A comparison of the discrete and dimensional models of emotion in music’, *Psychology of Music*, 39(1): 18-49.
- Gabrielsson, Alf. 2001. ‘Emotion perceived and emotion felt: same or different?’, *Musicae Scientiae*, (Special Issue): 123-147.
- Gregory, A. H., Worrall, L. & Sarge, A. 1996. ‘The Development of Emotional Responses to Music in Young Children’, *Motivation and Emotion*, 20(4): 341-348.
- Hevner, Kate. 1936 ‘Experimental Studies of the Elements of Expression in Music’, *The American Journal of Psychology*, 48(2): 246-268.
- Juslin, Patrik N. 1997. ‘Perceived Emotional Expression in Synthesized Performances of a Short Melody: Capturing the Listener’s Judgement Policy’, *Musicae Scientiae*, 1(2): 225-256.
- Juslin, P. N. & Lindström, E. 2010. ‘Musical Expression of Emotions: Modelling Listeners’ Judgements of Composed and Performed Features’, *Music Analysis*, 29(1-3): 334-364.