

# THE SOUND OF SADNESS: THE EFFECT OF PERFORMERS' EMOTIONS ON AUDIENCE RATINGS

Anemone G. W. Van Zijl, Geoff Luck

Department of Music, University of Jyväskylä, Finland  
Anemone.vanzijl@jyu.fi

## Abstract

Very few studies have investigated the effect of performers' felt emotions on the audience perception of their performances. Does it matter what a performer feels or thinks about when performing? To investigate this, we asked four violinists to play the same musical phrase in response to three different instructions. The first instruction was to focus on the technical aspects of their playing. The second instruction was to give an expressive performance. Following a sadness-inducing mood induction task, the third instruction was to play while focusing on their felt emotions. High quality audio and motion-capture recordings were made of all performances. Subsequently, motion-capture animations, audio recordings, and motion-capture animations combined with audio recordings of the performances were presented to an audience. Thirty audience members rated how much they liked each performance, how skilled they thought each performer was, and to what extent each performance was expressive of sadness. Statistical analysis revealed that, overall, audience members preferred the Expressive performances to the Technical and Emotional ones. In addition, the Expressive performances were rated as played by the most skilled performers. The Emotional performances, however, were rated as being most expressive of sadness. Our results suggest that what performers feel or think about when performing does affect the perception of their performances by an audience.

**Keywords:** performing musicians, felt emotions, audience perception

## 1. Introduction

Many studies have examined listeners' ability to recognize emotions expressed in music (e.g., Gabrielsson & Juslin, 2003; Juslin & Laukka, 2003; Eerola & Vuoskoski, 2013). In addition, the characteristics of the music – in terms of composed features (e.g., mode, harmonic, and rhythmic structure), performance features (e.g., tempo fluctuation, articulation, vibrato), and performer features (e.g., body movement, facial expression) – leading to the identification of certain emotions have been investigated (e.g., Clarke, 1988; Gabrielsson & Juslin, 1996; Gabrielsson & Lindström, 2010; Dahl & Friberg, 2007; Livingstone, Thompson & Russo, 2009). However, little is known about how per-

forming musicians actually try to achieve a performance expressive of emotions, and whether performers' felt emotions play a role in this process (Gabrielsson, 2001-2002). Should musicians feel the musical emotions when expressing them? Or should they rather focus on technique or expressivity when trying to bring a musical message across?

To investigate this, we asked performers to play the same musical phrase in response to three different instructions. This resulted in high quality audio and motion-capture recordings of so-called Technical, Expressive, and Emotional performances. Computational analysis of the audio recordings revealed differences in playing tempo, dynamics, articulatory

features, timbral features, and the extent and rate of vibrato between the three performance conditions. The Expressive performances, for instance, were characterized by the fastest playing tempo, the loudest sound, the brightest and roughest timbre, direct note attacks, and a wide and fast vibrato, as compared to the Technical and Emotional performances (Van Zijl, Toiviainen, & Luck, 2012). Computational analysis of the motion-capture recordings revealed differences in body posture, amount, speed, acceleration, and smoothness of movement of the performers in the three performance conditions. In the Expressive performances, for instance, performers were standing most upright, and moved most, fastest, with the highest acceleration, and lowest smoothness, as compared to the Technical and Emotional performances (Van Zijl & Luck, 2013). Although computational analyses of the recordings revealed differences between performances played with a different focus of attention, the question remains whether these differences would influence audience perception of the performances.

To investigate the effect of performers' thoughts and feelings on audience perception, we asked audience members to rate each performance with regard to three statements. The first statement – 'I like this performance' – was related to preference. Do audience members have a preference for Technical, Expressive, or Emotional performances? The second statement – 'This performer is skilled' – was related to expertise. Do audience members perceive performers who focus on Technique, Expressivity, or felt Emotions as more skilled? The third statement – 'This performance is expressive of sadness' – was related to emotions. Do audience members perceive Technical, Expressive, or Emotional performances as most expressive of, in this case, sadness?

## **2. Method**

### **2.1. Participants**

Participants were thirty Master's Degree students (mean age = 28.07 years, SD = 5.64, females = 18) from a University in Finland. All participants had played a musical instrument

(including voice) for at least one year, while the majority (63.3%) had played a musical instrument (including voice) for more than ten years.

### **2.2. Stimuli**

The stimuli were performances of four violinists (two amateurs and two professionals, all females) who were asked to play the same musical phrase in response to three different performance instructions. The first instruction was to focus on the technical aspects of their playing (i.e., the Technical performances). The second instruction was to give an expressive performance (i.e., the Expressive performances). Following a sadness-inducing mood induction task, the third instruction was to play while focusing on their felt emotions (i.e., the Emotional performances). High quality audio and motion-capture recordings were made of all performances.

Subsequently, motion-capture animations were created using the MATLAB Motion Capture Toolbox (Toiviainen & Burger, 2010). Using the QuickTime 7 software, the motion-capture animations were paired with the audio recordings. A presentation film depicting the motion-capture animations, audio recordings, and motion-capture animations with audio recordings, was created using the iMovie software. The order of the performances was randomised within each presentation mode (see below).

### **2.3. Procedure**

The performances were presented on a big screen in an auditorium. Participants were comfortably seated in the auditorium and the lights were dimmed, so as to resemble a real concert setting. Participants were asked to rate their agreement with the statements 1) I like this performance, 2) The performer is skilled, and 3) This performance is expressive of sadness, on a seven-point bipolar scale (completely disagree – completely agree). Participants were told that they would see or hear 36 performances played by different performers and with different performance intentions. They were neither told how many performers

had provided the performances, nor what the performance instructions had been.

The performances were presented in three blocks, each block containing the same performances but presented in different orders. In the first block, the motion-capture animations were shown without sound (i.e., Vision-only). In the second block, only the audio recordings were played (i.e., Audio-only). In the third block, the motion-capture animations were shown with sound (i.e., Vision & Audio). After each performance, participants had 20 seconds to rate the performance, until a sound signal indicated the start of the next performance.

To make sure all participants understood the rating procedure, data collection was preceded by an example of the same musical phrase performed by a bassoon player. After rating all performances, participants were asked to write down any comments they had about the study and their experiences. Data collection lasted about 45 minutes.

## 2.4. Analysis

Participants' ratings were entered into SPSS and analysed by means of three (one for each statement) three-way repeated-measures ANOVAs with presentation mode (Vision-only, Audio-only, Vision & Audio), expertise of the performer (Amateur, Professional), and performance condition (Technical, Expressive, Emotional) as independent variables. Correlations between ratings of the three statements were analysed by means of Pearson's Correlation Coefficient.

## 3. Results

We present the findings in accordance with the three statements investigating preference, perceived expertise, and perceived emotional expression. Figure 1 depicts the main effects of presentation mode (1A), expertise (1B), and performance condition (1C), as well as the two-way interactions between presentation mode and expertise (1D, 1G, 1J), presentation mode and performance condition (1E, 1H, 1K), and expertise and performance condition (1F, 1I, 1L).

### 3.1. Preference

The Vision-only performances received the highest preference ratings overall (Figure 1A), although the main effect of presentation mode on preference ratings was non-significant,  $F(2, 58) = 1.92, p > .05$ . There was a significant main effect of expertise on preference ratings,  $F(1, 29) = 8.81, p < .01$ , with the performances of the Professionals receiving the highest ratings overall (Figure 1B). The Expressive performances received the highest preference ratings overall (Figure 1C), and the main effect of performance condition on preference ratings was significant,  $F(2, 58) = 13.43, p < .001$ , with Bonferroni-corrected posthoc pairwise comparisons revealing a significant difference between the Technical and Expressive performances ( $p < .001$ ), and between the Expressive and Emotional ones ( $p < .001$ ), only.

As illustrated in Figures 1D and 1E, significant interactions were found between mode and expertise,  $F(2, 58) = 19.51, p < .001$ , and between mode and performance condition,  $F(4, 116) = 2.84, p < .05$ . The interaction between expertise and performance condition (Figure 1F) was non-significant,  $F(2, 58) = 0.40, p > .05$ . In addition, a significant three-way interaction was found between mode, expertise, and performance condition,  $F(4, 116) = 4.06, p < .01$  (not shown).

### 3.2. Perceived expertise

The Vision & Audio performances received slightly lower ratings of perceived expertise (Figure 1A), although the main effect of presentation mode on expertise rating was non-significant,  $F(1.36, 39.39) = 0.42, p > .05$ . There was a significant main effect of the performers' expertise on perceived expertise ratings,  $F(1, 29) = 39.13, p < .001$ , with the performances of the Professionals receiving higher ratings than the Amateur performances (Figure 1B). The Expressive performances received the highest ratings of perceived expertise (Figure 1C). The main effect of performance condition on expertise ratings was significant,  $F(2, 58) = 25.75, p < .001$ , with Bonferroni-corrected posthoc pairwise comparisons revealing a

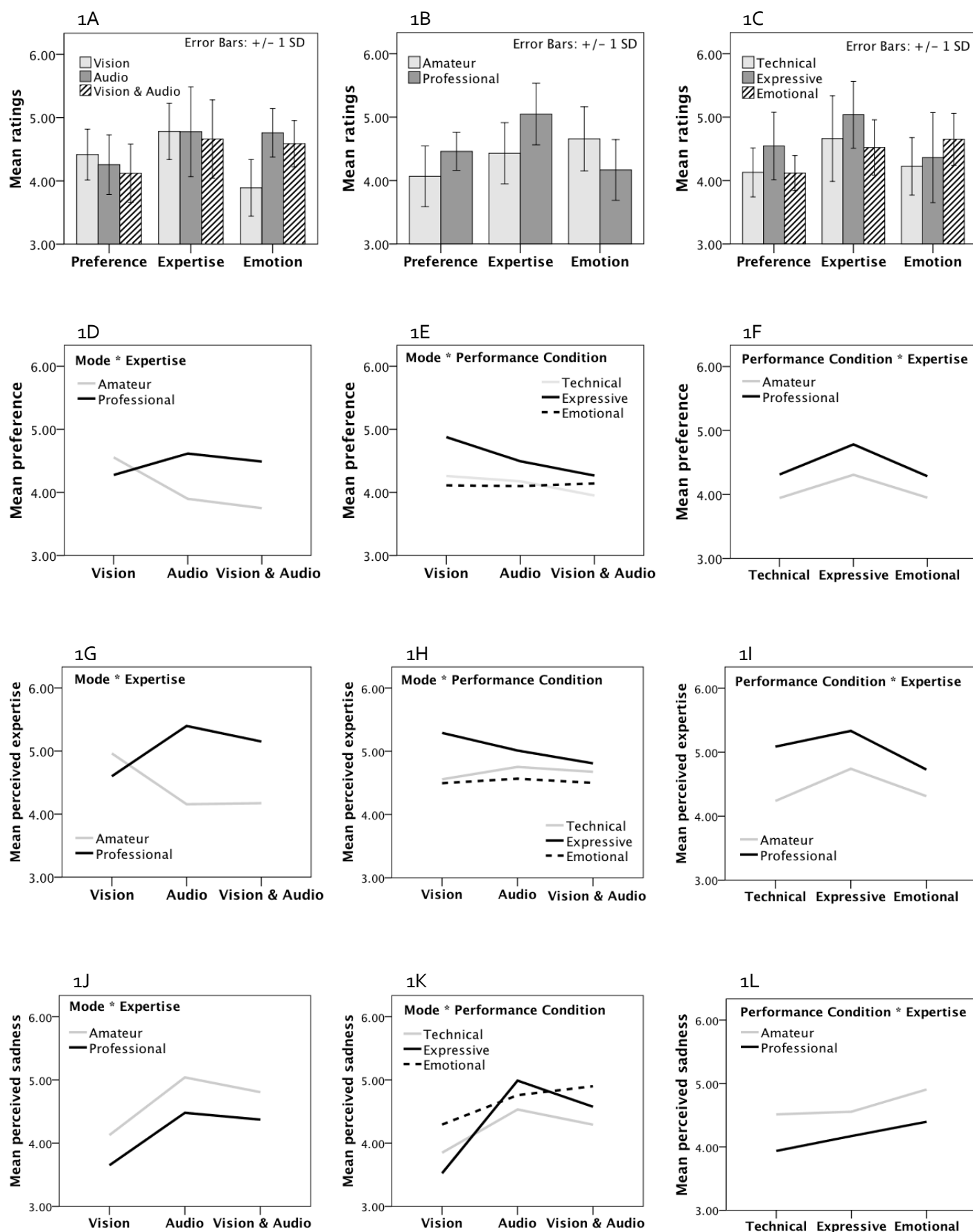


Figure 1. Main effects and two-way interactions of the repeated-measures ANOVAs.

significant difference between the Technical and Expressive performances ( $p < .001$ ), and between the Expressive and Emotional ones ( $p < .001$ ), only.

As depicted in Figures 1G, 1H and 1I, significant interactions were found between mode and expertise,  $F(2, 58) = 60.54$ ,  $p < .001$ , between mode and performance condition,  $F(4, 116) = 3.81$ ,  $p < .01$ , and between expertise and performance condition,  $F(2, 58) = 3.27$ ,  $p < .05$ . In addition, a significant three-way interaction was found between mode, expertise, and performance condition,  $F(3.17, 92.03) = 3.71$ ,  $p < .05$  (not shown).

### 3.3. Perceived emotional expression

The performances in the Vision-only mode received lower ratings of perceived expression of sadness than the performances in the Audio-only and Vision & Audio modes (Figure 1A). The main effect of presentation mode on perceived emotional expression ratings was significant,  $F(2, 58) = 15.38$ ,  $p < .001$ , with Bonferroni-corrected posthoc pairwise comparisons showing significant differences between the Vision-only and Audio-only presentation modes ( $p < .001$ ), and between the Vision-only and Vision & Audio ones ( $p < .01$ ), only. The performances of the Amateurs received higher ratings of perceived expression of sadness than the performances of the Professionals (Figure 1B). The main effect of expertise on perceived emotional expression ratings was significant,  $F(1, 29) = 25.00$ ,  $p < .001$ . The Emotional performances received the highest ratings of perceived expression of sadness (Figure 1C). The main effect was significant,  $F(2, 58) = 10.09$ ,  $p < .001$ , with Bonferroni-corrected posthoc pairwise comparisons showing a significant difference between the Technical and Emotional performances ( $p < .001$ ), and between the Expressive and Emotional ones ( $p < .05$ ), only.

As illustrated in Figure 1K, a significant interaction was found between mode and performance condition,  $F(4, 116) = 7.41$ ,  $p < .001$ . As depicted in Figures 1J and 1L, the interactions between mode and expertise,  $F(1.45, 41.93) = .24$ ,  $p > .05$ , between expertise and performance condition,  $F(2, 58) = .67$ ,  $p > .05$ ,

and between mode, expertise, and performance condition,  $F(4, 116) = 1.21$ ,  $p > .05$  (not shown), were non-significant.

### 3.4. Correlations

A significant correlation was found between ratings of preference and perceived expertise,  $r = .90$ ,  $p < .001$ . No correlation was found between ratings of preference and perceived emotional expression,  $r = -.01$ ,  $p > .05$ , or between ratings of perceived expertise and perceived emotional expression,  $r = -.14$ ,  $p > .05$ .

## 4. Discussion

Does it matter what a performer feels or thinks about when performing? The results of the present study suggest that a performer's focus of attention influences audience perception of a performance.

As illustrated in Figure 1C, statistical analysis of audience ratings revealed that, overall, audience members preferred the Expressive performances to the Technical and Emotional ones. In addition, the Expressive performances were rated as played by the most skilled performers. The Emotional performances, however, were rated as being most expressive of sadness.

When looking at differences between the Amateur and Professional performers, overall, the performances of the Professional violinists were rated higher in terms of preference and perceived skill. The Amateur performances, however, were perceived as being more expressive of sadness, as can be seen in Figure 1B.

The presentation mode, overall, did not really influence the ratings for preference and perceived expertise, as shown in Figure 1A. The presentation mode, however, did affect the ratings for perceived emotional expression. The ratings for perceived expression of sadness were much lower in the Vision-only condition.

The interactions between variables provided a more detailed view of the audience ratings. As depicted in Figures 1D and 1G, the presentation mode affected how the performances of the Amateurs and Professionals were perceived. In the Vision-only condition,

the performances of the Amateurs received higher ratings in terms of preference and perceived expertise. In the Audio-only and Vision & Audio conditions, the performances of the Professionals were rated higher. Analyses of the performers' movements revealed that the Amateurs moved more, more slowly, more smoothly, and with less acceleration than the professionals (Van Zijl & Luck, 2013). It seems that more extensive and more fluent movements were perceived as more pleasing, and were associated with a higher level of musical expertise – in the absence of sound. If we compare the ratings in the Vision & Audio condition to the Vision-only and Audio-only conditions, it becomes clear that the audience members were guided more by sound than by vision in their ratings.

As illustrated in Figure 1J, the performances in the Vision-only condition received the lowest ratings for perceived expression of sadness. This might be explained by the presentation order of the stimuli: When rating the performances in the Vision-only condition, the audience members did not know the piece that was played – although they heard it in the example performance. In addition, it might be difficult to infer the emotional expression of a performance by looking at motion-capture animations without the accompanying sound.

In all presentation modes shown in Figure 1J, the performances of the Amateurs were rated higher than the performances of the Professionals in terms of perceived expression of sadness. In addition to the differences in performers' movements, the differences in audio features of the Amateur and Professional performances might have been of influence here. Analysis of the audio features revealed that the Amateurs played slower, softer, with less direct note attacks, a different timbre, and a wider and slower vibrato, as compared to the Professionals (Van Zijl, Toiviainen, Luck, 2012). It seems that both the movement and auditory characteristics of the Amateur performances were more in line with the characteristics generally associated with the expression of sadness (e.g., Crane & Gross, 2007; Juslin & Laukka, 2003).

As illustrated in Figures 1E and 1H, the Expressive performances received higher ratings

than the Technical and Emotional performances in terms of preference and perceived expertise of the performer in all modes of presentation. The presentation mode interacted in different ways with the Technical, Expressive, and Emotional performances. Again, this might be explained by the audio and movement characteristics of the performances. The Expressive performances were characterised by the fastest tempo, the loudest sound, the most bright and rough timbre, direct note attacks, and a wide and fast vibrato, as compared to the Technical and Emotional performances (Van Zijl, Luck, Toiviainen, 2012). In addition, in the Expressive performances, the performers moved most, fastest, with most acceleration, and lowest levels of smoothness, as compared to the Technical and Emotional ones (Van Zijl & Luck, 2013). The Expressive performances seemed to be of a more extraverted character, which was appreciated by the audience.

As depicted in Figure 1K, the ratings of perceived sadness were different for each presentation mode. In the Vision-only condition the Technical performances were perceived as being most expressive of sadness, followed by the Emotional and Expressive ones. In the Audio-only condition, the Expressive performances were rated highest, followed by the Technical and Emotional ones. And in the Vision & Audio condition, the Technical performances scored highest, followed by the Expressive and Emotional ones. Whereas the pattern in the Vision-only condition might be related to the movement characteristics of the performers (e.g., performers moved least in the Technical condition and most in the Expressive condition), the patterns in the other modes are difficult to explain.

As can be seen in Figure 1F and 1I, the performances of the Professionals received higher ratings in terms of preference and perceived expertise than the performances of the Amateurs. The Expressive performances of both the Amateur and Professional performers received higher ratings than the respective Technical and Emotional ones. As depicted in Figure 1L, in case of perceived expression of sadness the performances of the Amateurs received higher ratings than the Professional

performances. In case of perceived expression of sadness, the Emotional performances of both the Amateur and Professional performers received higher ratings than the respective Technical and Expressive ones. The finding that audience members preferred the Expressive performances and believed they were played by the most skilled performers but perceived the Emotional performances as being most expressive of sadness might suggest that a more external focus of the performer (i.e., 'give an expressive performance') results in a 'better' performance, whereas a more internal focus (i.e., 'focus on felt emotions') results in a performance more expressive of emotion.

Should musicians feel the musical emotions when expressing them? Or should they rather focus on technique or expressivity when trying to bring a musical message across? The findings of the present study suggest that a performer's focus of attention affects the perception of the performance by an audience. It was found that audience members perceived the Emotional performances as more expressive of sadness than the Technical and Expressive ones. It seems that sad feelings of the performer can make a sad piece of music sound sadder. Although we cannot simply equate the lab setting of the present research with a real concert situation, the findings are valuable for music research, education and performance: It does matter what a performer feels or thinks about while performing.

## Acknowledgements

We would like to thank the musicians who provided the stimuli and the students who provided the ratings. The research reported here was supported by the Academy of Finland (project number 7118616).

## References

Clarke, E. F. (1988). Generative principles in music performance. In Sloboda, J. A. (Ed.) *Generative processes in music*. Oxford: Oxford University Press.

Crane, E., & Gross, M. (2007). Motion capture and emotion: Affect detection in whole body

movement. In A. Paiva, R. Prada, & R. W. Picard (Eds.), *Affective computing and intelligent interaction: Lecture notes in computer science* (pp. 95–101). Berlin: Springer-Verlag.

Dahl, S. & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24(5), 433–454.

Eerola, T. & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models and stimuli. *Music Perception*, 30(3), 307–340.

Gabrielsson, A. (2001–02). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae* [Special issue], 123–147.

Gabrielsson, A. & Juslin, P.N. (1996). Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience. *Psychology of Music*, 24, 68–91.

Gabrielsson, A. & Juslin, P.N. (2003). Emotional expression in music. In Davidson, R. J., Scherer, K. R. and Goldsmith, H. H. *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Gabrielsson, A. & Lindström, E. (2010). The role of structure in the musical expression of emotions. In Juslin, P. N. & Sloboda, J. A. (Eds.) *Handbook of Music and Emotion. Theory, Research, Applications*. (pp. 367–400). Oxford: Oxford University Press.

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.

Livingstone, S. R., Thompson, W. F., & Russo, F. A. (2009). Facial Expressions and Emotional Singing: A Study of Perception and Production with Motion Capture and Electromyography. *Music Perception*, 26(5), 475–488.

Toiviainen, P., & Burger, B. (2010). MoCap Toolbox manual. Retrieved from: <http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual>

Van Zijl, A. G. W. & Luck, G. (2012). Moved through music: The effect of experienced emotions on performers' movement characteristics. *Psychology of Music*, 41(2), 175–197.

Van Zijl, A. G. W., Toiviainen, P., & Luck, G. (2012). The sound of emotion: The effect of performers' emotions on auditory performance characteristics. In Cambouropoulos, E., Tsougras, C., Mavromatis, P., Pasiades, K. (Eds.) *Proceedings of the 11th ICMP and 8th ESCOM Conference*. (pp. 1064–1068). Greece: University of Thessaloniki.