

# The Sound of Emotion: The Effect of Performers' Emotions on Auditory Performance Characteristics

Anemone G. W. van Zijl<sup>\*1</sup>, Petri Toiviainen<sup>\*2</sup>, Geoff Luck<sup>\*3</sup>

<sup>\*</sup>*Department of Music, University of Jyväskylä, Finland*

<sup>1</sup>anemone.vanzijl@jyu.fi, <sup>2</sup>petri.toiviainen@jyu.fi, <sup>3</sup>geoff.luck@jyu.fi

## ABSTRACT

Do performers who feel sad sound different compared to those who express sadness? Despite an extensive literature on the perception of musical emotions, little is known about the role of performers' experienced emotions in the construction of an emotionally expressive performance. Here, we investigate the effect of performers' experienced emotions on the auditory characteristics of their performances. Seventy-two audio recordings were made of four amateur and four professional violinists playing the same melodic phrase in response to three different instructions. Participants were first asked to focus on the technical aspects of their playing. Second, to give an expressive performance. Third, to focus on their experienced emotions, prior to which they were subjected to a sadness-inducing mood induction task. Performers were interviewed about their thoughts and feelings after each playing condition. Statistical and computational analyses of audio features revealed differences between the performance conditions. The Expressive performances revealed the highest values for playing tempo, dynamics, and articulatory features such as the attack slope. The Emotional performances, in contrast, revealed the lowest values for all of these features. In addition, clear differences were found between the performances of the amateur and professional performers. The present study provides concrete evidence that performers who feel sad do sound different compared to those who express sadness.

## I. INTRODUCTION

Can we hear a difference between a performer who thinks about expressivity, and a performer who mourns the loss of a beloved one? Intuitively, the answer seems to be yes. However, performers usually meticulously practice and plan their performances beforehand (e.g., Chaffin, Imreh & Crawford, 2002). They have to play the same piece over and over again, and move their audience night after night. It is likely that their expressive conveyance becomes highly automatic, and has nothing to do with their actual emotions while performing (Juslin, 2009). Also, many professional performers state that the music occupies centre stage, not the performer or his or her emotions (e.g., Sloboda & Lehmann, 2001). On the other hand, several studies suggest that performers need to feel the emotions in the music to be able to convincingly express them to their audiences (e.g., Persson, Pratt, & Robson, 1992; Persson, 2001; Lindström et al., 2003). Does it make a difference whether a performer feels or expresses the emotions present in the music?

In several studies (e.g. Woody, 2000; Karlsson & Juslin, 2008), emotional and expressive playing are considered as being one and the same. Van Zijl and Sloboda (2011) however, found that music students described 'emotional

playing' as 'just feeling and enjoying the music' and 'expressive performance' as 'the conveyance of a previously constructed musical interpretation to an audience'. With the present study, we aim to explore this proposed difference, and investigate the effect of performers' experienced emotions on the auditory characteristics of their performances.

As far as we know, only two studies addressing the effect of performers' experienced emotions on the produced performances have been published previously. Higuchi, Fornari, and Leite (2010) subjected nine pianists to a severe mood induction procedure before they performed a piece for four hands by R. Schumann – together with the main researcher – once while focusing on each note, and once while thinking of the emotional stimuli. Higuchi et al. found that participants played more *legato* and with less metric precision in the emotional playing condition. Glowinsky et al. (2008) subjected two violinists to a mood induction procedure before they performed several pieces by J.S. Bach in an angry, sad, joyful and peaceful manner. Analyses of the audio, video and physiological recordings revealed differences in duration time of the performances, differences in heart rate, galvanic skin response, and tension in the right arm. The authors did not specify the differences in relation to the performance conditions.

In the present study, we asked performers to play the same musical phrase in response to three different instructions, the last instruction following a mood induction task. Computational analyses enabled us to quantify subtle differences in auditory characteristics between performance conditions. The results suggest a difference in performance characteristics when a performer thinks about expressivity, or mourns the loss of a beloved one.

## II. METHOD

### A. Participants

Eight violinists (4 professionals and 4 accomplished amateurs, all female, mean age 24.3 years,  $SD = 1.8$ ) participated in the study. Participants in both groups were similar in terms of gender, age, and years of playing. They differed in terms of years of lessons, practice hours per week, and number of performances per year.

### B. Musical stimuli

The two musical phrases used in the study were taken from *Three miniatures for oboe and piano: Orientale, Chansonette, A la campagne* by the British composer Sir H. Harty (1911). Only the solo parts were used, not the piano accompaniment.

Two violinists tested the phrases beforehand, to make sure they were suitable for being played on a violin. The composer and title of the piece were removed from the score. The phrases were chosen because of their unfamiliarity, contrasting characters, and limited difficulty despite their beauty. In our analyses, we focus on the first 14-bar phrase in g-minor taken from the movement *Oriente*. The phrase could be characterized as sad and atmospheric. The tempo indication is *Lento ma non troppo*. Figure 1 shows the score of this phrase. The second 18-bar phrase in A-major (not shown) taken from the movement *Chansonette* was included to help ensure that participants were in a positive mood at the end of the data collection (see below).



**Figure 1. Musical score of the phrase composed by H. Harty, as used in the study. (Copyright 1911 Stainer & Bell Ltd, 23 Grunisen Road, London N3 1DZ England www.stainer.co.uk)**

### C. Apparatus

Audio recordings were made using ProTools8 software and a Microtech Gefell M300 directional microphone surrounded by an SE Electronics Reflection Filter. For reference purposes, video recordings were made with four Sony video cameras.

### D. Procedure

To assess their current mood, participants first completed the PANAS state questionnaire (Watson, Clark & Tellegen, 1988). Then, three recording sessions followed in which participants played the same musical phrase in response to three different instructions. Per instruction, participants were encouraged to play the phrase at least three times.

The first instruction was: “While playing, try to focus on the technical aspects of your playing: so, think for instance about the rhythm, think about the dynamics, and think about the articulation.” After the first recording session, participants were interviewed about their performances. They were asked which performance was most representative for the instruction given; what they thought about while playing; whether it felt natural to them to play while thinking about the instruction; and how they felt while playing.

The instruction for the second recording session was: “Play as you would normally play while giving an expressive performance.” Subsequently, the same interview questions as listed above were asked.

Before the third recording session, participants were subjected to a two-part mood induction task. They were told that the phrase was taken from a piece of music written by a British composer, shortly after he had lost his four-year-old son. Participants were asked to imagine how this happened while listening to a story describing the evening the composer

lost his son. In addition, participants were instructed to think of an intense sad emotional experience. They either wrote a few lines about their experience, or told the experimenter about it. The mood induction tasks were inspired by findings in the literature (Juslin, 2003; Van Zijl, 2008; Persson, Pratt, & Robson, 1992; Persson, 2001). Subsequently, participants were asked to play the phrase in response to the third instruction: “Think about the emotional experience you’ve just written/told about. Try to feel the emotions of grief, of great sadness expressed by the music. Don’t think about your playing, just focus on the emotions.”

Following the third recording session, participants completed the PANAS state questionnaire again, to assess whether their mood had changed after the mood induction procedure. Then, the same interview questions as listed above were addressed.

After the participants had completed a background questionnaire, they were debriefed: the composer of the music did not have a son, he was just happily married. To put the participants into a happy mood again, they completed a final recording session in which they played the second ‘happy’ phrase, and received two free movie tickets.

Finally, participants were asked what they thought about the experiment; what they thought were the differences between their performances; what their best recording session was (i.e., in relation to which instruction) and why they thought so; and whether there was anything they wanted to add or ask. In total, data collection lasted about 90 minutes.

### E. Audio feature extraction

Prior to analyses, all performances were manually segmented per bar, using the Sonic Visualiser software (Cannam, Landone & Sandler, 2010). Subsequently, using the MATLAB MIR-Toolbox (Lartillot & Toiviainen, 2007), several audio features were extracted for (each bar of) each performance. Playing tempo was obtained from the segment information. Rhythmical variation was estimated by calculating the average time the lengths of two subsequent bars differed. Dynamics were examined by calculating the root mean square (RMS) value. Articulation was investigated by calculating the attack slope values.

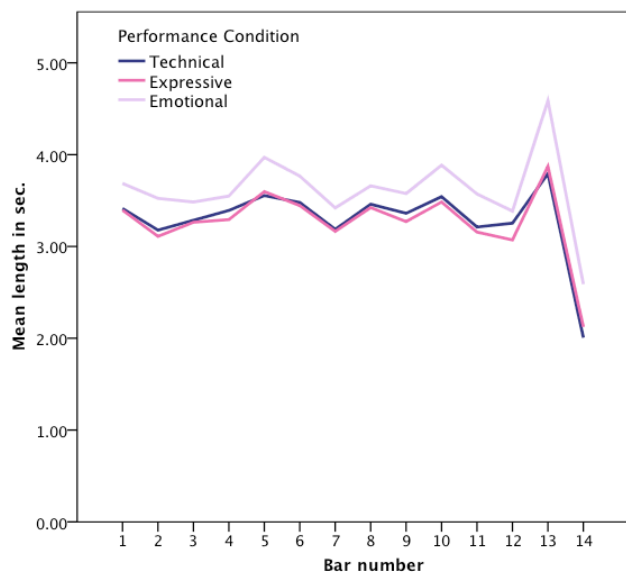
## III. RESULTS

### A. Tempo and Rhythmical variation

The average length of performances was shortest in the Expressive condition ( $M = 45.67$  sec,  $SD = 9.57$ , Tempo = 67.66 BPM), slightly longer in the Technical condition ( $M = 46.12$ ,  $SD = 11.67$ , Tempo = 68.33 BPM), and longest in the Emotional condition ( $M = 50.66$  sec,  $SD = 9.81$ , Tempo = 75.05 BPM). A one way repeated-measures ANOVA revealed a significant main effect of performance condition on performance length,  $F(2, 44) = 15.56$ ,  $p < .001$ . Post hoc pair wise comparisons with Bonferroni correction revealed significant differences in performance length between the Technical and Emotional condition (Mean Difference =  $-4.54$ ,  $p < .01$ ) and between the Expressive and Emotional condition (Mean Difference =  $-4.99$ ,  $p < .001$ ).

An independent t-test revealed that the performances of the professionals ( $M = 41.53$ ,  $SD = 7.92$ , Tempo = 61.53 BPM)

were much shorter than the performances of the amateurs ( $M = 52.94$ ,  $SD = 9.61$ , Tempo = 78.43 BPM), Mean Difference = 11.41,  $t(66.29) = 5.40$ ,  $p < .001$ .



**Figure 2. Average bar length per performance condition, across performers.**

Figure 2 shows the average length of each musical bar per performance condition. As can be seen, the Technical and Expressive performances were very similar in tempo, whereas the Emotional performances were played slower. The pattern of rhythmical variation (i.e., taking more or less time for certain bars) was similar across performance conditions. Performers particularly slowed down in bars 5, 10, and 13. This could be explained by the large amount of short notes in these bars, and by the musical function of these bars: these bars form the end of a musical line within the phrase.

By calculating the mean of absolute differences between subsequent bar lengths, we derived a measure of variation of playing tempo. We found that the variation was highest in the Emotional condition ( $M = 0.53$  sec,  $SD = 0.24$ ), lower in the Expressive condition ( $M = 0.45$  sec,  $SD = 0.15$ ), and lowest in the Technical condition ( $M = 0.44$ ,  $SD = 0.20$ ). A one way repeated-measures ANOVA revealed a significant main effect of performance condition on variation of playing tempo,  $F(1.36, 29.98) = 6.78$ ,  $p < .05$ . Post hoc pair wise comparisons with Bonferroni correction revealed a significant difference in variation of playing tempo between the Technical and Emotional condition (Mean Difference =  $-0.097$ ,  $p < .05$ ).

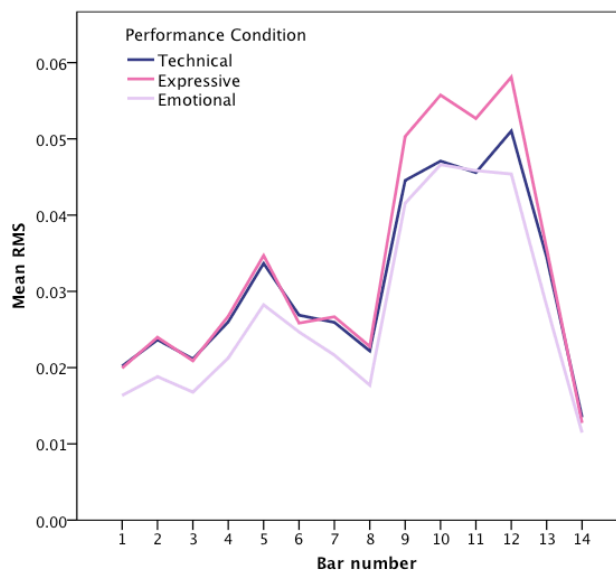
An independent t-test revealed that the playing tempo of the professionals varied significantly less ( $M = 0.38$ ,  $SD = 0.16$ ) than the playing tempo of the amateurs ( $M = 0.56$ ,  $SD = 0.20$ ), Mean Difference = 0.17,  $t(67) = 3.99$ ,  $p < .001$ .

## B. Dynamics

The average level of dynamics, estimated by calculating the root mean square energy (RMS) of each performance, was highest in the Expressive condition ( $M = 0.0336$ ,  $SD = 0.00669$ ), lower in the Technical condition ( $M = 0.0315$ ,  $SD = 0.00788$ ), and lowest in the Emotional condition ( $M = 0.0276$ ,  $SD = 0.00523$ ). A one way repeated-measures ANOVA revealed a significant main effect of performance condition on

RMS value,  $F(1.37, 30.03) = 12.23$ ,  $p < .01$ . Post hoc pair wise comparisons with Bonferroni correction revealed significant differences in RMS value between the Technical and Expressive condition (Mean Difference =  $-0.002$ ,  $p < .05$ ), between the Technical and Emotional condition (Mean Difference = 0.004,  $p < .05$ ), and between the Expressive and Emotional condition (Mean Difference =  $-0.006$ ,  $p < .01$ ).

An independent t-test revealed that the performances of the professionals ( $M = 0.035$ ,  $SD = 0.006$ ) were associated with a higher RMS value than the performances of the amateurs ( $M = 0.027$ ,  $SD = 0.006$ ), Mean Difference = 0.008,  $t(67) = -5.47$ ,  $p < .001$ .



**Figure 3. Average RMS per bar per performance condition, across performers.**

Figure 3 shows the average RMS value per musical bar. As can be seen, the dynamics indicated in the score are followed: *piano* in the beginning, *crescendo* from bar eight leading to a *forte* part, and a *diminuendo* towards the end. The Technical and Expressive performances are very similar in dynamics over the first eight bars, whereas the Emotional performances are played softer. In the last six bars, the Technical and Emotional performances are relatively similar, whereas the Expressive performances are played louder. In bar 11 we notice a dynamical drop in the Expressive and Technical performances. Examination of the videos revealed that this drop might be explained by the bowing used: in case the b-flat is played with an up-bow stroke the sound tends to be softer.

## C. Articulation

We investigated the articulation, or phrasing, of the notes by estimating the attack slope of each note being played. A high attack slope value means that the note is clearly phrased or shaped – clearly pronounced, as one would say about speech.

Overall, the average value of attack slope was highest in the Expressive condition ( $M = 206432.71$ ,  $SD = 125403.00$ ), lower in the Technical condition ( $M = 151828.43$ ,  $SD = 105191.00$ ), and lowest in the Emotional condition ( $M = 134141.14$ ,  $SD = 71147.00$ ). A one way repeated-measures ANOVA revealed a significant main effect of performance condition on attack slope,  $F(2, 44) = 10.10$ ,  $p < .001$ . Post hoc

pair wise comparisons with Bonferroni correction revealed significant differences in attack slope between the Technical and Expressive condition (Mean Difference =  $-54600.00$ ,  $p = .001$ ) and between the Expressive and Emotional condition (Mean Difference =  $-72291.56$ ,  $p < .01$ ).

An independent t-test revealed that the performances of the professionals ( $M = 215900$ ,  $SD = 121966$ ) were associated with a higher attack slope value than the performances of the amateurs ( $M = 116680$ ,  $SD = 58696$ ), Mean Difference =  $-992153$ ,  $t(45.17) = -4.24$ ,  $p < .001$ .

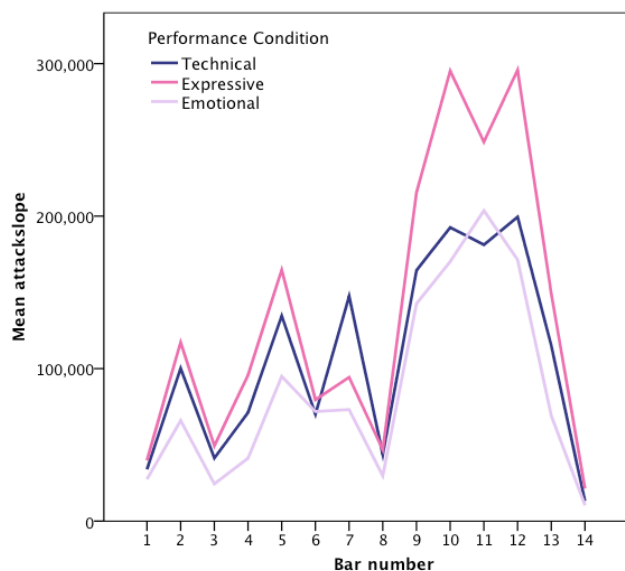


Figure 4. Average attack slope per bar per performance condition, across performers.

Figure 4 shows the average attack slope per bar. If we compare this graph with Figure 3 showing the RMS values, it can be seen that the attack slope patterns are similar but enlarged. Pearson's correlation coefficient indicated a statistically significant relationship between RMS and mean attack slope values,  $r = .82$ ,  $p < .001$ . In musical terms this means that aspects such as bow-changes and variation in bow pressure and speed affect both articulation and dynamics.

#### IV. DISCUSSION

What is the effect of performers' experienced emotions on their auditory performance characteristics? We aimed to investigate this question by asking performers to play a musical phrase in response to three different instructions – focus on technique; focus on expressivity; focus on experienced emotions. We examined the playing tempo, rhythmical variation, dynamics, and articulation of the performances obtained. We found differences between performance conditions.

The Expressive performances revealed the highest values for playing tempo, dynamics, and articulatory features such as the attack slope. In line with the findings of Higuchi, Fornari, and Leite (2010) and Glowinski et al. (2008), the Emotional performances revealed the lowest values for all of these features. The Technical performances revealed intermediate values for the auditory characteristics investigated. These findings seem to suggest an external focus in the Expressive

performances, resulting in more extraverted playing, versus an internal focus in the Emotional performances, resulting in more introverted playing. This interpretation finds support in the interview data obtained. In the words of one of the participants: "when I was thinking that there was an audience [in the Expressive condition]. I played bigger and.. I played to someone, not to myself like I played the last and the first ones." In the words of another participant: "[in the Emotional condition] I was thinking about crying, about blackness, about sadness. (...) And then I got a feeling here [points towards her breast]. I was just feeling that, just physically listening to my body."

In addition, we found differences between the performances of the amateur and professional performers. The playing tempo of the amateurs was slower, the dynamical range of their performances was smaller, and their notes were less distinctively phrased. We might explain these findings by the different level of expertise. The amateurs might have chosen to play the phrase in a slower tempo to make sure they could cope with the technical demands of the phrase. A smaller dynamic range might be related to the bowing technique used: a higher bow speed results in a more intense sound (Gelre, 1991), but requires a higher level of expertise. The phrasing of the individual notes, likewise, is related to bow control, something that increases when expertise develops.

Can we hear a difference between a performer who thinks about expressivity, and a performer who mourns the loss of a beloved one? The findings of the present study indicate that a different performance focus results in differences in auditory features. This finding is interesting in relation to the debate whether felt and posed/portrayed emotions result in similar vocal expressions (e.g., Scherer & Banse, 1996; Bachorowski & Owren, 2003). It is also interesting in relation to the question whether a performer should feel the emotions intended in the music in order to convincingly convey them to the audience, or not. However, ultimately it is the listener's task to decide whether the presence of performers' emotions leads to the desired sound of emotion – which future research should address.

#### ACKNOWLEDGEMENTS

We would like to thank Marc R. Thompson and Olivier Lartillot for their help with the audio feature extraction. The research reported here was supported by the Academy of Finland (project numbers 7118616 and 125710).

#### REFERENCES

- Cannam, C., Landone, C. and Sandler, M. (2010). Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files, in *Proceedings of the ACM Multimedia 2010 International Conference*.
- Chaffin, R., Imreh, G. & Crawford, M. (2002). *Practicing Perfection: Memory and Piano Performance*. Mahweh/ New Jersey/ London: Lawrence Erlbaum Associates.
- Gerle, R. (1991). *The art of bowing practice. The expressive bow technique*. London: Stainer & Bell.
- Glowinsky, D., Camurri, A., Volpe, G., Chiarra, N., Cowie, R., McMahan, E., Jaimovich, J., & Knapp, R. B. (2008). Using induction and multimodal assessment to understand the role of

- emotion in musical performance. In Peter et al., (Eds.) *Proceedings of the workshop Emotion in HCI – Designing for People*. (pp. 8-12). Fraunhof: Fraunhofer Verlag.
- Harty, H. (1911). *Three miniatures for oboe and piano: Orientale, Chansonette, A la campagne*. [Music score] London: Stainer & Bell.
- Higuchi, M. K. K., Fornari, J. E., & Leite, J. P. (2010). The influence of cognition and emotion in pianistic performance. In S. M. Demorest, S. J. Morrison & P. S. Campbell (Eds.), *Proceedings of the 11th International Conference on Music Perception and Cognition (ICMPC11)*. (pp. 186-191). Seattle. Washington. USA.
- Juslin, P. N. (2003). Five Facets of Musical Expression: A Psychologist's Perspective on Music Performance. *Psychology of Music*, 31 (3), 273-302.
- Juslin, P. (2009). Emotion in music performance. In: S. Hallam, I. Cross & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology*. Oxford: Oxford University Press.
- Karlsson, J. & Juslin, P. N. (2008). Musical Expression: an observational study of instrumental teaching. *Psychology of Music*, 36 (3), 309-334.
- Lartillot O. & Toivainen, P. (2007). A MATLAB toolbox for musical feature extraction from audio. In *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*, Bordeaux, France, September 10-15.
- Lindström, E., Juslin, P. N., Bresin, R. & Williamon, A. (2003). 'Expressivity comes from within your soul': A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, 20, 23-47.
- Persson, R. S., Pratt, G. & Robson, C. (1992). Motivational and Influential Components of Musical Performance: A Qualitative Analysis. *European Journal for High Ability*, 3, 206-217.
- Persson, R. S. (2001). The Subjective World of the Performer. In Juslin, P. N. and Sloboda, J. A. (Eds.), *Music and Emotion. Theory and Research* (pp. 275-289). Oxford: Oxford University Press.
- Sloboda, J. A. & Lehmann, A. C. (2001). Tracking Performance Correlates of Changes in Perceived Intensity of Emotion During Different Interpretations of a Chopin Piano Prelude. *Music Perception*, 19 (1), 87-120.
- Van Zijl, A. G. W. (2008). *The role of the performer's experienced emotion in the construction of a musically expressive performance*. Unpublished master's dissertation, Keele University, UK.
- Van Zijl, A. G. W. and Sloboda, J. (2011). Performers' experienced emotions in the construction of expressive musical performance: an exploratory investigation. *Psychology of Music*, 39 (2), 196-219.
- Watson, D., Clark, L. A., Tellegen, A. (1988). Development And Validation Of Brief Measures Of Positive And Negative Affect - The Panas Scales. *Journal of Personality and Social Psychology*, 54 (6), 1063-1070.
- Woody, R. H. (2000). Learning Expressivity in Music Performance: An Exploratory Study. *Research Studies in Music Education*, 14, 14-23.