

Abstract

Previous research has shown that – compared to audio-only presentations – the audiovisual presentation of a musical performance consistently enhances observers’ appreciation. Furthermore, it has been hypothesized that observation of a musical performance through multiple channels – auditory and visual – may result in a heightened emotional response due to the broader range of actions and cues available (Livingstone & Thompson, 2009). In this study, we set out to investigate how the mode of presentation – audio-only (AO), video-only (VO), and audiovisual (AV) – affects participants’ emotional responses to a musical performance. Nineteen adults took part in an experiment in which they were presented with AO, VO, and AV versions of a recorded piano performance. Their emotional responses to the stimuli were measured using self-report combined with psychophysiological indices of experienced emotions (skin conductance and heart rate). In contrast to the predictions arising from previous work, skin conductance responses indicated that emotional arousal was highest in the audio-only presentation mode, compared to both audiovisual and video-only presentation modes. Self-reports of felt emotions did not reveal any significant differences between the AV and AO presentation modes, although both were rated as eliciting more intense emotional responses than the VO mode, and the AO presentation mode elicited more pleasant emotional responses than the VO mode. These findings do not support the view that audiovisual presentations would increase the appreciation of a musical performance as compared to audio-only presentations. Potential explanations for these findings are discussed.

Keywords: Music-induced emotion, musical performance, multisensory perception, psychophysiology, skin conductance

Introduction

Music is capable of conveying emotional meaning and eliciting emotional responses. In musical performance, emotional expression is communicated through auditory cues such as articulation, timbre, vibrato, dynamics, and timing (e.g., Gabrielsson, 1999; Juslin & Timmers, 2010; Palmer, 1997), as well as by means of visual cues such as body posture, gestures, and facial expressions (e.g., Dahl & Friberg, 2007; Davidson, 1993, 1994; Thompson, Russo, & Quinto, 2008). Research conducted over the past two decades has unequivocally established that visual cues contribute significantly to observers' perceptions of the expressivity (Davidson, 1993, 1994; Vuoskoski, Thompson, Clarke, & Spence, 2014) and emotional expression (Dahl & Friberg, 2007; Krahé, Hahn, & Whitney, 2013; Thompson et al., 2008) of a musical performance, but less is known about the role of visual cues in the emotional responses that are elicited. Although some previous studies have attempted to investigate the effect of visual performance cues on observers' experienced emotions (e.g., Krahé et al., 2013; Vines et al., 2011), the challenges involved in distinguishing experienced emotion from perceived emotion using self-report (cf. Konecni, 2008) renders the implications of the findings unclear.

The psychological mechanisms of music-induced emotion have chiefly been considered from a unisensory, auditory perspective (e.g., Juslin & Västfjäll, 2008; Scherer & Zentner, 2001). Despite the fact that Juslin and Västfjäll actually proposed visual imagery as one of the potential mechanisms by which music can induce emotional responses in listeners, they discuss this mechanism in terms of non-musical visual imagery that is conjured up by the music (such as images of nature, for example) rather than performance-related imagery. However, some theories of music-induced emotion directly predict that the observation of a musical performance through multiple sensory modalities – auditory and visual – ought to lead to an enhanced, or more intense, emotional response (Livingstone &

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Thompson, 2009; Molnar-Szakacs & Overy, 2006). Molnar-Szakacs and Overy proposed that through the coupling of perception and action – a kind of ‘simulation’ of the actions and affective cues present in musical performance – the mirror neuron system might account for some of the emotional responses that are elicited by music. Mirror neurons – first discovered in the macaque monkey – fire both when we perform a goal-directed action, and when we observe (see or hear) another agent performing the same action (Di Pellegrino et al., 1992; Gallese, 2006; Kohler et al., 2002). In the context of musical performance, the mirror neuron system may respond to sound-producing actions as well as to the emotional expression conveyed by the performance (i.e., resonating with those auditory and gestural features in the music that resemble vocal and motor expressions of emotion; Overy & Molnar-Szakacs, 2009). Although we are able to hear and mirror these actions and affective cues by means of auditory information alone (Kohler et al., 2002; Molnar-Szakacs & Overy, 2006), it has been hypothesized that the activation of multiple channels – auditory and visual – may result in a heightened emotional response due to the broader range of actions and cues available for mirroring (Livingstone & Thompson, 2009).

This view, proposed by Livingstone and Thompson (2009), subsequently received empirical support: A recent meta-analysis revealed that – compared to audio-only presentations – audiovisual presentations consistently enhance observers’ appreciation (defined as “liking,” perceived “expressiveness,” “overall quality,” and “overall impression”) of a musical performance (Platz & Kopiez, 2012). **Moreover**, Chapados and Levitin (2008) reported that observers’ emotional reactivity to clarinet performances – measured using psychophysiological indices of emotional arousal – was heightened in the audiovisual presentation condition (as compared to the audio-only and video-only conditions). Physiological measures of emotional arousal circumvent the problem of differentiating perceived emotion from felt emotion in subjective reports, as well as the difficulties (and inter-individual variability) involved in distinguishing and conceptualizing emotional responses (see, e.g.,

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Barrett, 2006; Gohm & Clore, 2000). The limited reliability of introspection – especially when it comes to emotional responses and their relative intensities (cf. Nisbett & Wilson, 1977) – makes self-report measures of music-induced emotions potentially vulnerable to bias and demand characteristics. For example, if asked to rate the intensity of their emotional responses to a musical performance, participants may base their ratings on perceived expressivity (i.e., the expressive cues conveyed by the performer) rather than on their own experienced emotions (cf. Vuoskoski et al., in press).

Conversely, physiological indices of experienced emotion – such as skin conductance – have been shown to reflect emotional responses even when participants are not consciously aware of the stimuli that elicited the emotion (e.g., Öhman & Soares, 1994). Skin conductance has been reliably associated with the arousal dimension of experienced emotions (e.g., Bradley & Lang, 2000), with higher skin conductance responses reflecting more intense emotional reactions, irrespective of their valence (Winton, Putnam, & Krauss, 1984). The relationship between the valence dimension and physiological responses appears to be less straightforward, although facial muscle electromyography, heart rate, and eye-blink startle reflex do to correlate with emotional valence (e.g., Bradley & Lang, 2000). In the context of music-induced emotions, differentiated psychophysiological responses to different kinds of music have been reported in several studies (e.g., Gomez & Danuser, 2007; Krumhansl, 1997; Nyklíček et al., 1997), with the temporal features of music correlating with arousal-related responses (Gomez & Danuser, 2007). Chills or frissons, which are pleasurable piloerection responses sometimes evoked by music (see, e.g., Huron & Margulis, 2010), have also been associated with a specific pattern of psychophysiological responses. Guhn, Hamm, and Zentner (2007; see also Salimpoor et al., 2009) found that those musical passages that evoked frissons also elicited the greatest increases in heart rate and skin conductance, thus further affirming that certain

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

psychophysiological measures are associated with the intensity of emotional responses in musical contexts.

The aim of this study was to investigate the effect of mode of presentation – audio-only, video-only, and audiovisual – on participants’ emotional responses to a musical performance. Emotional responses were measured using self-report combined with psychophysiological indices of experienced emotions, namely skin conductance and heart rate. In line with Livingstone and Thompson’s (2009) theoretical predictions, we hypothesized that the audiovisual presentation would induce more intense emotional responses than either the audio-only or video-only presentations. Although Chapados and Levitin (2008) have already provided preliminary support for this hypothesis, certain aspects of their study design and method are unsatisfactory. First, only 12 participants were tested in a between-participants design with three conditions, leaving only four participants per condition. As emotional responses to music (and the psychophysiological responses associated with them) are highly subjective (e.g., Hodges, 2010; Juslin, 2009), it is difficult to determine how much of the variance between conditions was due to inter-individual variability rather than the presentation mode. Second, because they used performances of an atonal piece for solo clarinet without any notated meter as their musical material (the second of Stravinsky’s *Three Pieces for Clarinet Solo*), it is unclear whether their findings would also apply to other kinds of music and other instruments. In order to further explore the effect of visual cues on music-induced emotions – and to extend Chapados and Levitin’s findings – we utilized a within-participants rather than a between-participants design, and used a recorded performance of Romantic piano music as our stimulus material.

Method

Participants

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Nineteen healthy participants (5 male), aged between 19 and 58 years ($M = 28.31$, $SD = 11.47$), took part in the experiment. Participants were asked to indicate their level of musical expertise by choosing a classification that best described them (non-musician, amateur musician, or professional musician). Twelve of the participants identified themselves as amateur musicians, while none considered themselves a professional musician. All of the participants gave their informed consent prior to start of the experiment. The participants were recruited via the participant pool of *REMOVED FOR REVIEW*, and received a monetary incentive (£5) in return for their participation. All of the experimental procedures followed the University of *REMOVED FOR REVIEW* Policy on the Ethical Conduct of Research Involving Human Participants and received the relevant approval from the Research Ethics Committee.

Stimuli

The experimental stimuli consisted of audiovisual (AV), video-only (VO), and audio-only (AO) versions of a recorded piano performance. A female pianist with international concert experience performed the Intermezzo in B Minor (Op. 119) by Johannes Brahms on a Yamaha Disklavier grand piano. The piece was selected from the existing repertoire of the pianist on the basis that the performer herself found it moving and expressive, and that it was considered engaging and accessible for listeners. The performance was videoed from a front-left angle (see Figure 1 for a sample frame) using a digital HD-camcorder mounted on a tripod. This angle (rather than a side angle) was chosen so that participants would be able to see the performer's facial expressions in addition to her head and torso movements. In other research (e.g. Thompson & Luck, 2011) the head and upper torso have been documented to convey most of the expressive information in the context of a piano performance. The audio was recorded in wav-format using a stereo microphone, a MOTU 828mk2 audio interface, and Reaper v4.402 software running on Mac OS X. We also generated neutral filler stimuli, which

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

consisted of nature scenes and soundscapes recorded in a park. The purpose of these filler stimuli was to minimize any potential carry-over effects between successive trials.

Audiovisual and video-only versions of the stimuli were generated using Final Cut Pro (version 10.0.6) running on Mac OS X. The audiovisual and video-only stimuli were presented in mp4-format with a resolution of 1280×720 , while the audio-only stimuli were presented in mp3-format. The audio was presented via headphones (Sony MDR-ZX600). When the audio-only stimuli were presented, the participants saw a black screen. The duration of each piano performance stimulus (AV, VO or AO) was 4 minutes and 3 seconds, while the duration of each filler stimulus was 3 minutes.

The audiovisual version of the piano performance was rated by 7 expert judges in terms of auditory and visual expressivity, and the representativeness/typicality of the visual and auditory expression (in relation to other performances of the same style). Both expressivity and typicality were evaluated using scales ranging from 0 to 100. The scale extremes were “Not at all expressive” and “Very expressive” for visual and auditory expressivity, and “Very unconventional / eccentric” and “Very representative / typical” for the typicality of the visual and auditory expression. The expert judges were doctoral students and postdoctoral researchers from 3 different institutions, working in the field of music psychology, and had a minimum of 10 years of musical instrument training ($M = 13.57$, $SD = 3.41$). The judges were not aware of the aims or hypotheses of the study. The mean expressivity ratings were 69.0 ($SD = 17.3$) for auditory expressivity, and 53.7 ($SD = 11.5$) for visual expressivity. The mean typicality ratings were 79.0 ($SD = 23.1$) for auditory expression, and 70.4 (12.8) for visual expression. In sum, the auditory component of the performance was rated as somewhat more expressive than the visual component, while both visual and auditory components were rated as representative of a typical piano performance in the Romantic tradition.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Physiological sensors

Sensors for skin conductance and blood volume pulse were attached to each participant's non-dominant hand. Galvanic skin responses (GSR, also called skin conductance, SC) were recorded from electrodes (model: SA9309M) placed on the second and fourth fingertips, and blood volume pulse (BVP) from a photoplethysmograph (model: SA9308M) placed on the third finger. Data from the sensors were collected by an 8-channel encoder (Pro-Comp Infinity), at a sampling rate of 100 Hz. Both sensors and the encoder were manufactured by Thought Technology Ltd., Montreal, CA.

Procedure

At the beginning of the experiment, the participants were invited to take a seat in front of a computer screen. They were informed about the experiment and the task that they would be asked to perform, and were asked to read and sign the consent form. Subsequently, the participants were connected to the physiological sensors. The experimenter carefully explained the function of each sensor in order to make the participants feel at ease. The experimenter informed the participants that they could withdraw from the experiment at any moment should they wish to do so. The signal from each sensor was then checked by the experimenter on a second computer screen outside the experiment room. In order to obtain an optimal signal, the physiological sensors were attached to each participant's non-dominant hand, and participants were asked to keep their non-dominant hand relaxed and still for the duration of the experiment. Once the signal was considered stable and clear, the experimenter asked the participant to move the mouse cursor over the "start" button positioned on the bottom left of the computer screen to start the experiment.

A purpose-designed computer program was written in the C# programming language to acquire the physiological data, deliver the stimuli, and to collect the participants' ratings. The stimuli were

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

presented in a random order. After each stimulus, participants rated the intensity of the emotional effect the stimulus had on them on a 7-point scale (1 = did not move me at all, 7 = moved me very strongly), and described their emotional reactions using 6 unipolar scales representing the three bipolar dimensions of core affect (Schimmack & Grob, 2000): The scales measured positive valence (represented by the adjectives pleasant, good, and positive), negative valence (unpleasant, negative, bad), **high** energy arousal (awake, wakeful, alert), **low** energy arousal (sleepy, tired, drowsy), **high** tension arousal (tense, clutched up, jittery), and **low** tension arousal (relaxed, calm, at rest). Each group of adjectives was rated on a 7-point scale (1 = does not describe my emotional reaction at all, 7 = describes my emotional reaction very well). The presentation order of these six scales was randomized. After describing their own emotional reaction, the participants were also asked to describe their perception of the emotional expression conveyed by the stimuli. The same six scales were again used in the ratings, ranging from 1 (does not describe the performance at all) to 7 (describes the performance very well), and presented in a randomized order. Felt emotions were always rated prior to rating perceived emotions, as we were primarily interested in participants' experienced emotions. While the focus of this study is on felt emotions, we anticipated that the rating of both perceived and felt emotions would help participants to better distinguish their felt emotions from perceived emotions.

Results

Self-reported felt emotion

To explore participants' overall emotional responses to the piano performance stimuli, the ratings for the unipolar scale pairs representing each of the three core affect dimensions were compared using paired Wilcoxon signed rank tests (the ratings of each participant were averaged across the three presentation modes). The analyses revealed that participants' emotional responses to the stimuli were characterized by higher positive ($Mdn = 4.67$) than negative ($Mdn = 2$) valence, $Z = -2.48$, $p = .013$;

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

neither high ($Mdn = 3.67$) nor low ($Mdn = 3.67$) energy arousal, $Z = -0.04$, $p = ns$; and low ($Mdn = 5.33$) rather than high ($Mdn = 2$) tension arousal, $Z = -3.57$, $p < .001$. The mean ratings and standard deviations for the different presentation modes are displayed in Table 1. To investigate the effect of presentation mode on participants' ratings of felt emotion, seven repeated-measures analyses of variance (ANOVA) with one within-participants factor, Stimulus Type (AO, VO or AV), were carried out. There was a significant main effect of Stimulus Type in the ratings of the overall intensity of emotional response, $F(2,36) = 14.34$, $p < .001$, η_G^2 (generalized eta squared; Bakeman, 2005) = .36. Multiple comparisons of means (paired t-tests, $p < .05$ significance level adjusted using the Holm-Bonferroni method; Holm, 1979) revealed that the intensity ratings obtained in the AV and AO presentation modes were significantly higher than those obtained in the VO mode. Similarly, there was a significant main effect of Stimulus Type in the ratings of felt positive valence, $F(2,36) = 6.36$, $p = .004$, $\eta_G^2 = .11$. Post-hoc comparisons of means revealed that the AO stimulus was rated as evoking significantly more positive valence than the VO stimulus. The comparisons also indicated that the ratings of positive valence obtained in the AV presentation mode were higher than those obtained in the VO mode, though the result was not significant once the Holm-Bonferroni correction was applied. There were no significant main effects in the ratings of felt negative valence, energy arousal (high or low), or tension arousal (high or low).

Perceived emotion

Next, we investigated the effect of presentation mode on the ratings of perceived emotional expression (the mean ratings and standard deviations are displayed in Table 2). Again, the within-participants variable was Stimulus Type. There were significant main effects of Stimulus Type in the ratings of positive valence, $F(2,36) = 6.89$, $p = .003$, $\eta_G^2 = .15$, and the ratings of high energy arousal, $F(2,36) =$

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

6.06, $p = .005$, $\eta_G^2 = .14$. Post-hoc comparisons of means revealed that in the AO presentation mode, the performance was evaluated as conveying significantly more positive valence than in the VO mode.

Post-hoc comparisons also revealed that in the AV and AO presentation modes, the performance was evaluated as conveying significantly more energy arousal than in the VO mode.

Physiological measures

Electrodermal activity. The extraction of skin conductance features was performed using the MATLAB toolbox Ledalab. Four different skin conductance features were extracted by means of continuous deconvolution analysis (CDA; Benedek & Kaernbach, 2010), which separates the continuous GSR signal into tonic and phasic components. The slowly-varying tonic skin conductance (i.e., skin conductance level) reflects general psychological arousal, whereas phasic activity, manifested by wave-like peaks in the skin conductance signal, is a good indicator of event-related responses (e.g., Lovibond, 1992). The means and standard deviations for each measure are displayed in Table 3. Absolute (rather than baseline-adjusted) measurements of skin conductance were used in all the analyses, as each participant was only compared to himself/herself across the three conditions. The entire length of the stimulus was used as the response window. For four of the participants we were not able to obtain usable data, either because of excessive movement artefacts, or else there was no clear evidence of dynamic electrodermal responses. Thus, $n = 15$ for the following analyses.

The mean tonic skin conductance within the response window (i.e., within each stimulus) was analyzed using a mixed-design ANOVA with one within-participants factor (Stimulus Type; AO, VO, or AV), and one between-participants factor (Presentation Order). Presentation Order was included in the analyses in order to account for the possible habituation of the skin conductance response that might occur during the course of the experiment (see e.g., Roth, Dawson, & Fillion, 2012; Grewe et al., 2007). There was a significant main effect of Stimulus Type, $F(2,20) = 16.6$, $p < .001$, $\eta_G^2 = .06$, but no

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

significant main or interaction effects related to Presentation Order. Post-hoc comparisons of means revealed that participants' tonic skin conductance was significantly higher in the AO and AV presentation modes compared to the VO mode, and higher during the AO mode compared to the AV mode. Next, we looked at the average phasic driver (the average magnitude of event-related skin-conductance responses within the response window; see Benedek & Kaernbach, 2010), which reliably reflects phasic activity. Once again, there was a main effect of Stimulus Type, $F(2,20) = 7.42, p = .004, \eta_G^2 = .14$, but no significant main or interaction effects related to Presentation Order. Post-hoc comparisons revealed that the participants exhibited more phasic activity in response to AO and AV stimuli as compared to the VO stimuli. Finally, two additional features related to phasic activity were investigated: the number of significant skin conductance responses, and the sum of the amplitudes of significant skin conductance responses. In both cases, there was a significant main effect of Stimulus Type, $F(2,20) = 5.13, p = .016, \eta_G^2 = .02$, and $F(2,20) = 9.22, p = .001, \eta_G^2 = .10$, respectively. Once again, post-hoc comparisons revealed that there was more phasic activity present in the AO and AV presentation modes than in the VO mode. There were no significant main or interaction effects related to Presentation Order.

Heart rate. Recordings from the BVP sensor were analyzed in terms of heart rate (HR) and heart rate variability (HRV). The raw data was inspected visually in order to detect and remove any artefacts. For each participant, the BVP signal was transformed to Z-points, after which a Butterworth band pass filter (cut-off frequencies 1 HZ and 8HZ) was applied to the data. A cubic spline was fitted to the data in order to compute the first derivative and to identify the peaks of the signal where the derivative was null (Haag, 2004). The number of samples between two signal peaks was calculated and divided by the recording sampling frequency of 100 Hz so that the peak-to-peak interval could be represented in

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

milliseconds, and the progression of the values defined the inter-beat interval (IBI) sequence. From the IBI sequence we obtained two measures of heart rate variability, as defined by Buccelletti (Buccelletti et al., 2012); the standard deviation of IBI intervals, and the root mean square the differences between adjacent IBI intervals (RMSSD). Repeated-measures ANOVAs with one within-participants factor (Stimulus Type) were used to analyze the variables, but no main effects were found. The means and standard deviations for HR and HRV are displayed in Table 3.

Dynamic electrodermal activity. In order to investigate whether the temporal profile of the skin conductance differed between presentation modes, the skin conductance response curves of the 15 participants obtained in the 3 different presentation modes were analyzed using functional data analysis (FDA). FDA allows each curve to be used as a datum that can be analyzed using techniques comparable to classic inferential statistics (Levitin, Nuzzo, Vines, & Ramsay, 2007). In order to perform the functional analysis, the raw data were scaled using z-score transformations before fitting B-splines on each curve (considered to be the best technique for approximating most non-periodic functional data; Ramsay & Dalzell, 1991). Z-transformation was performed using the mean and standard deviation of each individual response window (i.e., individually for each participant in each presentation mode), so as to focus the analysis on the temporal dynamics of the signal, rather than on the absolute values, which varied considerably among participants. A further reason to compute the mean and the standard deviation for the z-score transformation from the data of each individual response window (rather than from all three response windows of each participant, for example) was the nature of the skin conductance signal. The tonic component of SC (Lim, Rennie, Barry, Bahramali, Lazzaro, Manor, & Gordon, 1997) has a relatively slow temporal dynamic, which makes it possible for the absolute value of SC to be influenced to a significant degree by the order of presentation of the stimuli.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Three two-way functional analyses of variance (fANOVAs) with random projection (Cuesta-Albertos & Febrero-Bande, 2010) were used with modality and participants as the independent-group variable and the random variable, respectively. Analyses were performed using the software R, with the implemented library *fda.usc* (Febrero-Bande & Oviedo de la Fuente, 2012). The analysis comparing the AO and VO modes approached significance ($RP=240$, $df=2,14$, $p = .058$), prompting further investigation. In this further investigation, the values of the electrodermal signals for each pair of functions (obtained by fitting B-splines to the normalized raw signal) were analyzed by means of paired t-tests using a critical level of $p < .05 / 3$. Figure 2 highlights the results for the paired t-tests on each frame (frame rate: 100 Hz). No statistically significant differences were found for the comparisons AO vs. AV, or VO vs. AV. However, the electrodermal responses in the AO presentation mode significantly differed from the VO mode in four instances during the musical performance ($p < .05$). These instances are illustrated in Figure 2.

Correlations between self-reports and physiological measures. Finally, we investigated the associations between self-reported felt emotion and physiological indices of emotional arousal. In order to account for inter-individual variability in physiological responses and scale use, all variables were mean-centered within individuals. All data points were used in the analyses, but the threshold for statistical significance was adjusted to correspond to the actual degrees of freedom (13 for EDA and 15 for HR). Due to the limited degrees of freedom as well as the number of tests, no statistically significant correlations emerged. However, the correlation coefficients, displayed in Table 4, indicate clear positive trends between skin conductance level and felt intensity, positive valence, and high energy arousal, and between phasic skin conductance responses and felt positive valence.

Discussion

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

In contrast to previous findings (e.g. Chapados & Levitin, 2008) as well as our own hypothesis, we found that – relative to the audio-only presentation – the audiovisual presentation of a performance of Romantic piano music did not lead to a more intense emotional response. In fact, skin conductance responses indicated that the emotional arousal was highest in the audio-only presentation mode, compared to both audiovisual and video-only presentation modes. Overall, the piano performance stimuli used in the present study evoked pleasant (rather than unpleasant) and relaxing (rather than tense) emotions with a moderate degree of energy arousal. Self-reports of felt emotions failed to reveal any significant differences between the AV and AO presentation modes, although both were rated as eliciting more intense emotional responses than the VO mode, and the AO presentation mode elicited more pleasant emotional responses than the VO mode. The ratings of perceived emotional expression reflected a similar pattern, with the AO version rated as more pleasant than the VO version, and the AO and AV versions rated as more energetic than the VO version.

The findings for the self-reported intensity of experienced emotion are in line with those reported by Vines et al. (2011), who also found (using a between-participants experimental design) that the overall intensity of self-reported felt emotion did not differ between the AO and AV conditions, but both evoked more intense emotional responses than the VO condition. However, the finding that the AV version did not evoke more pleasant emotions than the AO mode is somewhat discrepant with previous studies, as the meta-analysis by Platz and Kopiez (2012) showed that AV presentations tended to enhance observers' appreciation of a musical performance relative to AO presentations. However, it should be noted that Platz and Kopiez's definition of "appreciation" ("liking," "expressiveness," "overall quality," or "overall impression") is somewhat different from felt positive valence, and thus this discrepancy may also reflect differences in the measured concepts. The skin conductance responses obtained in the present study – although in line with the self-report data – are also discrepant with previous findings. Chapados and Levitin (2008) found that the AV condition elicited the highest levels

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

of electrodermal amplitude (i.e., tonic skin conductance) relative to the AO and VO conditions, while the AO presentation led to the highest tonic and phasic skin conductance activity in the present study (although the difference between the AO and AV modes was significant only in the case of tonic activity). No effects of condition were observed in the heart rate measures, although the trend in heart rate variability was in line with the skin conductance measures. The lack of significant effects may be partly due to the wide age range of the participants, as older people have less variability in HR (see Etzel et al., 2006, for an example).

There are several possible explanations for the difference between our findings and those of Chapados and Levitin (2008). First, we utilized a within-participants design in which all participants were exposed to all three presentation modes, while Chapados and Levitin (2008) conducted a between-participants comparison with a sample of 12 participants. Although repeated exposure to a musical piece can have an effect on liking and thus on the emotional response induced by the piece (e.g., Hunter & Schellenberg, 2011), we accounted for this potential source of error by randomizing the presentation order of the stimuli, and by including Presentation Order as a factor in **some of** our analyses. In contrast, there is no reliable way of accounting for inter-individual variability in music-induced emotional responses – especially in a study conducted on a small sample. Second, the stimulus material used in the present study differed from that used in Chapados and Levitin's study in terms of instrumentation (clarinet vs. piano) and musical style (atonal vs. Romantic tonal). It could be argued that the emotional expression in atonal music is less apparent and less accessible to listeners than in Romantic music (which emphasizes emotional expression, e.g., Kennedy, Bourne Kennedy, & Rutherford-Johnson, 2012; and is more familiar to the majority of listeners), with the consequence that the combined AV mode in Chapados and Levitin's study may have provided participants with more salient additional information (regarding emotional expression) than in the present study. **Indeed, a later study by Vines et al. (2011) supports this interpretation: Using the same stimuli as Chapados and**

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Levitin (plus three additional performances of the same piece), Vines et al. demonstrated that the variations in expressive intention had a significant effect on participants' felt emotions when the performances could be seen, whereas no such effects were observed in the audio-only condition.

Furthermore, compared to the piano, the clarinet may allow more freedom of movement for the performer. The participants in Chapados and Levitin's study saw the entire body of the performer, while participants in the present study only saw the head, face, and upper torso of the participant. Thus, it may be that the clarinetist in Chapados and Levitin's study conveyed more expressive cues via body movements than the pianist in the present study. However, the facial expressions of a performer can also convey important information regarding emotional expression, and these are arguably more restricted in the case of wind instruments where the mouth is engaged in sound production.

We should also consider potential explanations for the failure of our findings to support the hypothesis that the presentation of a musical performance through multiple modalities would lead to a more intense emotional response than a unimodal presentation. Why did the AO presentation mode lead to more arousing emotional responses than the AV mode? Although the expert judges rated the auditory component of the piano performance as somewhat more expressive than the visual component, the visual component was nevertheless considered to convey an average level of expression. Thus, it is unlikely that the participants would have found the visual component of the piano performance so unexpressive that it would have deducted from their emotional experience in the audiovisual presentation mode. It may be that in the absence of visual stimulation, participants engaged in emotionally engaging visual imagery. Visual imagery is one of the musical emotion-induction mechanisms suggested by Juslin and Västfjäll (2008), where a listener conjures up – intentionally or unintentionally – visual images while listening to music. Previous laboratory investigations of music-induced emotions have found that music-induced imagery appears to be a rather common phenomenon (especially when listening to instrumental music; see, e.g., Vuoskoski & Eerola, 2012, 2013). Music-

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

induced imagery has the potential to intensify listeners' emotional responses to music (Vuoskoski & Eerola, 2013), and one could argue that the participants were freer to engage in visual imagery in the AO presentation mode (in which they faced a black computer screen) than in the AV mode. The emotional and expressive cues present in the actual visual component of the AV stimulus may not be as strong or effective in eliciting emotional responses as the potential visual images conjured up by the participants themselves. It may also be that the participants simply found it more pleasant to immerse themselves in imagery or other kinds of thoughts while listening to music, compared to having to attend to the videoed performance stimulus.

A second possible explanation is related to musical expectancy, another mechanism underlying music-induced emotions, where an emotional response is evoked when a specific feature in the music violates, delays, or confirms the listener's expectations about the continuation of the music (Juslin & Västfjäll, 2008; Meyer, 1956). Unexpected musical events such as sudden changes in texture, tempo, rhythm, or harmony have been associated with music-induced chills or frissons, which are arguably some of the most powerful and pleasurable responses evoked by music (Panksepp, 1995). In a musical performance, some of the performer's movements and gestures may serve to articulate the underlying musical structure, emphasize phrase boundaries, and anticipate emotional changes in the music (MacRitchie, Buck, & Bailey, 2013; Vines et al., 2006). Seeing these gestures may make the unfolding musical events more predictable and less surprising to the observer (especially in the case of unfamiliar music), and thus reduce the tension generated by musical expectancies. In other words, it may be that participants found the musical events more 'surprising' in the AO presentation mode, and more predictable in the AV mode, which could help to explain the higher skin conductance levels recorded in the AO mode. Conversely, atonal music (such as the clarinet piece used in the study by Chapados and Levitin) offers fewer cues for the formation of clear musical expectations, so that additional visual cues

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

might be particularly important in the formation of musical expectations (and subsequent surprises) when attending to a performance of atonal music.

As this study is based on one performance by one performer, our findings cannot be generalized across other performers and performances. Although both the visual and auditory expression of the performer were rated by expert judges as representative of a typical piano performance in the Romantic tradition, the auditory component of the performance was rated as somewhat more expressive than the visual component. However, it is important to note that the visual component was not rated as unexpressive, but as conveying an average level of expressivity (53.7 on a scale from 0 to 100). It may also be that in the case of the specific piece used in the present experiment, the Intermezzo in B Minor (Op. 119) by Johannes Brahms, the potential effect of visual expressive cues was limited by the significant role that auditory expressive cues such as loudness and timing variation (cf. Bhatara et al., 2011) play in the performance of Romantic music. The specific musical and expressive features (e.g., slow tempo, minor key, and considerable variation in loudness and timing) of the Intermezzo may also explain – at least in part – the mixed pattern of relaxation and emotional arousal (i.e., intensity of felt emotion) reported by the participants. As different felt emotions have differing implications for physiological responses, it is possible that a piece evoking more tension or less emotional arousal, for example, might lead to a diverging pattern of results also on the level of psychophysiology.

In conclusion, the present study could not reject the null hypothesis that attending to a musical performance through both auditory and visual modalities does not enhance the emotional response compared to the audio-only condition. In fact, skin conductance measures revealed that the participants in this study experienced the highest levels of emotional arousal when only listening to the piano performance. These findings therefore suggest that, contrary to what has been proposed elsewhere (see Platz & Kopiez, 2012), audiovisual presentations do not necessarily increase the appreciation of a

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

musical performance as compared to audio-only presentations, and they suggest the need for further research into the complexities of audiovisual interactions in musical performance. Future studies should endeavour to use more varied stimuli – in terms of musical style, instrumentation, and emotional expression – and to use both self-report and psychophysiological measures to gain a more reliable account of music-induced emotional responses.

References

- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379-384.
- Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, 10(1), 20-46.
- Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190, 80-91.
- Bhatara, A., Tirovolas, A. K., Marie Duan, L., Levy, B., & Levitin, D. J. (2011). Perception of emotional expression in musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 37(3), 921-934.
- Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37, 204-215.
- Buccelletti, F., Bocci, M. G., Gilardi, E., Fiore, V., Calcinaro, S., Fragnoli, C., Maviglia, R., & Franceschi, F. (2012). Linear and nonlinear heart rate variability indexes in clinical practice. *Computational and Mathematical Methods in Medicine*, 2012, article ID 219080, doi:10.1155/2012/219080.
- Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, 108(3), 639-651.
- Cuesta-Albertos, J. A., & Febrero-Bande, M. (2010). A simple multiway ANOVA for functional data. *Test*, 19(3), 537-557.
- Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24(5), 433-454.
- Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, 21(2), 103-113.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

- Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive invention to an audience? *Journal of Human Movement Science*, 6, 279-301.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91(1), 176-180.
- Febrero-Bande, M., & Oviedo de la Fuente, M. (2012). Statistical computing in functional data analysis: The R package fda.usc. *Journal of Statistical Software*, 51(4), 1-28.
- Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd Ed., pp. 501-602). San Diego, CA: Academic Press.
- Gallese, V. (2006). Intentional attunement: A neurophysiological perspective on social cognition and its disruption in autism. *Brain Research*, 1079(1), 15-24.
- Gohm, C. L., & Clore, G. L. (2000). Individual differences in emotional experience: Mapping available scales to processes. *Personality and Social Psychology Bulletin*, 26(6), 679-697.
- Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, 7(2), 377.
- Guhn, M., Hamm, A., & Zentner, M. (2007). Physiological and musico-acoustic correlates of the chill response. *Music Perception*, 24(5), 473-484.
- Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion recognition using bio-sensors: First steps towards an automatic system. *Affective Dialogue Systems: Lecture Notes in Computer Science*, 3068, 36-48.
- Hodges, D. (2010). Psychophysiological measures. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 279-312). Oxford, UK: Oxford University Press.
- Hunter, P. G., & Schellenberg, E. G. (2011). Interactive effects of personality and frequency of exposure on liking for music. *Personality and Individual Differences*, 50(2), 175-179.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

- Huron, D., & Margulis, E. H. (2010). Musical expectancy and thrills. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 575-604). Oxford, UK: Oxford University Press.
- Juslin, P. N. (2009). Emotional responses to music. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 131-140). Oxford, UK: Oxford University Press.
- Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 453-489). Oxford, UK: Oxford University Press.
- Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31, 559-575.
- Kennedy, M., Rutherford-Johnson, T., & Kennedy, J. (2013). *The Oxford Dictionary of Music*. Oxford, UK: Oxford University Press.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297(5582), 846-848.
- Koneni, V. J. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 2(2), 115-129.
- Krahé, C., Hahn, U., & Whitney, K. (2013). Is seeing (musical) believing? The eye versus the ear in emotional responses to music. *Psychology of Music*. DOI: 0305735613498920.
- Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 51(4), 336.
- Lavy, M. M. (2001). *Emotion and the experience of listening to music: A framework for empirical research* (Unpublished doctoral dissertation). University of Cambridge, UK.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

- Levitin, D. J., Nuzzo, R. L., Vines, B. W., & Ramsay, J. O. (2007). Introduction to functional data analysis. *Canadian Psychology/Psychologie Canadienne*, 48(3), 135-155.
- Lim, C. L., Rennie, C., Barry, R. J., Bahramali, H., Lazzaro, I., Manor, B., & Gordon, E. (1997). Decomposing skin conductance into tonic and phasic components. *International Journal of Psychophysiology*, 25(2), 97-109.
- Livingstone, R. S., & Thompson, W. F. (2009). The emergence of music from the Theory of Mind. *Musicae Scientiae*, 13(2 suppl), 83-115.
- Lovibond, P. F. (1992). Tonic and phasic electrodermal measures of human aversive conditioning with long duration stimuli. *Psychophysiology*, 29, 621-632.
- MacRitchie, J., Buck, B., & Bailey, N. J. (2013). Inferring musical structure through bodily gestures. *Musicae Scientiae*, 17(1), 86-108.
- Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago: The University of Chicago Press.
- Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: From motion to e'motion. *Social Cognitive and Affective Neuroscience*, 1, 235-241.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231.
- Nyklíček, I., Thayer, J. F., & Van Doornen, L. J. (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, 11(4), 304-321.
- Öhman, A., & Soares, J. J. (1994). "Unconscious anxiety": Phobic responses to masked stimuli. *Journal of Abnormal Psychology*, 103(2), 231-240.
- Palmer, C. (1997). Music performance. *Annual Review of Psychology*, 48(1), 115-138.
- Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception*, 13(2), 171-207.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

- Platz, F., & Kopiez, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception*, 30(1), 71-83.
- Ramsay, J. O., & Dalzell, C. J. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 539-572.
- Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., & Zatorre, R. J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS ONE*, 4(10), e7487.
- Scherer, K. R. & Zentner, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 361-392). Oxford: Oxford University Press.
- Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: A quantitative comparison by means of structural equation modeling. *European Journal of Personality*, 14(4), 325-345.
- Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion*, 22, 1457-1470.
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, 118, 157-170.
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1), 80-113.
- Vuoskoski, J. K., & Eerola, T. (2012). Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories. *Psychology of Aesthetics, Creativity, and the Arts*, 6(3), 204-213.
- Vuoskoski, J. K. & Eerola, T. (2013). Extramusical information contributes to emotions induced by music. *Psychology of Music*. DOI: 10.1177/0305735613502373.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

- Vuoskoski, J. K., Thompson, M. R., Clarke, E., & Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception, & Psychophysics*, 76(2), 591-604.
- Vuoskoski, J. K., Thompson, M. R., Spence, C., & Clarke, E. (in press). Interaction of sight and sound in the perception and experience of musical performance. *Music Perception*.
- Winton, W. M., Putnam, L. E., & Krauss, R. M. (1984). Facial and autonomic manifestations of the dimensional structure of emotion. *Journal of Experimental Social Psychology*, 20(3), 195-216.

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Tables

Table 1. Mean ratings (and standard deviations) of emotions felt in response to the three different types of stimuli (video-only, audio-only, and audiovisual).

	Presentation mode		
	Video-only	Audio-only	Audiovisual
Overall intensity	2.74 (1.52)	4.74 (1.48)	4.89 (0.94)
Positive valence	3.74 (1.52)	5.05 (1.62)	4.74 (1.85)
Negative valence	2.21 (1.44)	2.37 (1.77)	2.84 (1.95)
High energy arousal	2.95 (1.39)	3.63 (1.46)	3.84 (1.71)
Low energy arousal	3.63 (1.92)	3.79 (1.96)	3.11 (1.79)
High tension arousal	2.11 (1.52)	2.00 (1.45)	2.47 (1.65)
Low tension arousal	4.63 (1.74)	5.32 (1.34)	5.05 (1.39)

Note: The adjectives used to represent each of the scales were *pleasant*, *good*, and *positive* (positive valence); *unpleasant*, *negative*, and *bad* (negative valence); *awake*, *wakeful*, and *alert* (high energy arousal); *sleepy*, *tired*, and *drowsy* (low energy arousal); *tense*, *clutched up*, and *jittery* (high tension arousal); and *relaxed*, *calm*, and *at rest* (low tension arousal).

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Table 2. Mean ratings (and standard deviations) of perceived emotional expression in the three different types of stimuli (video-only, audio-only, and audiovisual).

	Presentation mode		
	Video-only	Audio-only	Audiovisual
Positive valence	3.68 (1.46)	5.11 (1.37)	4.74 (1.52)
Negative valence	1.74 (0.93)	1.84 (1.17)	2.42 (1.30)
High energy arousal	2.37 (0.96)	3.58 (1.26)	3.53 (1.84)
Low energy arousal	4.47 (1.87)	4.32 (1.86)	3.79 (1.99)
High tension arousal	1.53 (0.84)	1.74 (1.24)	2.00 (1.25)
Low tension arousal	4.63 (1.67)	5.26 (1.24)	4.95 (1.51)

Note: The adjectives used to represent each of the scales were *pleasant*, *good*, and *positive* (positive valence); *unpleasant*, *negative*, and *bad* (negative valence); *awake*, *wakeful*, and *alert* (high energy arousal); *sleepy*, *tired*, and *drowsy* (low energy arousal); *tense*, *clutched up*, and *jittery* (high tension arousal); and *relaxed*, *calm*, and *at rest* (low tension arousal).

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Table 3. Means and standard deviations of the electrodermal measures (average tonic skin conductance, average phasic driver, the number of significant skin conductance responses, and the sum of the amplitudes of significant skin conductance responses) and heart rate variables (heart rate and heart rate variability) recorded in the three different presentation modes (video-only, audio-only, and audiovisual).

	Presentation mode		
	Video-only	Audio-only	Audiovisual
Tonic skin conductance (μ S)	1.05 (0.94)	1.66 (1.33)	1.31 (1.09)
Average phasic driver (μ S)	.038 (.035)	.085 (.076)	.067 (.059)
Number of sig. SC responses	55.2 (101.0)	87.4 (125.6)	69.1 (87.4)
Sum of sig. SC response amplitudes (μ S)	1.71 (2.93)	4.77 (6.07)	3.31 (3.97)
Heart rate (BPM)	71.2 (5.02)	72.6 (6.09)	71.8 (5.67)
Heart rate variability (RMSSD; BPM)	5.89 (5.02)	6.39 (3.71)	4.73 (3.84)

EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Table 4. Pearson correlation coefficients between self-reported felt emotions and physiological responses. All variables were mean-centred within individuals. All data points were used in the analyses, but degrees of freedom were 13 for the skin conductance measures, and 15 for the heart rate measures. None of the correlations were statistically significant.

	Tonic skin conductance (μ S)	Average phasic driver (μ S)	Heart rate (BPM)	Heart rate variability (RMSSD; BPM)
Overall intensity	.35	.17	-.05	-.18
Positive valence	.48	.34	-.11	.04
Negative valence	-.13	-.14	-.06	-.33
High energy arousal	.46	.29	.08	-.06
Low energy arousal	-.08	-.17	-.27	.12
High tension arousal	-.13	-.30	.11	-.05
Low tension arousal	.17	.27	-.31	-.03

Figures

Figure 1. A sample frame from the video component of the stimuli.



EMOTIONAL RESPONSES TO MUSICAL PERFORMANCE

Figure 2. The averaged skin conductance response curves obtained in the 3 different presentation modes. The electrodermal responses in the AO presentation mode significantly differed from the VO mode in four instances ($p < .05$; marked with grey).

