

Goodchild, Meghan, Jonathan Wild, and Stephen McAdams. "Exploring Emotional Responses to Orchestral Gestures." *Musicae Scientiae*, (May 2017). doi:[10.1177/1029864917704033](https://doi.org/10.1177/1029864917704033).

Exploring emotional responses to orchestral gestures

Meghan Goodchild, Jonathan Wild & Stephen McAdams

Schulich School of Music, McGill University

Accepted for publication in *Musicae Scientiae*

Abstract

Research on emotional responses to music indicates that prominent changes in instrumentation and timbre elicit strong responses in listeners. However, there are few theories related to orchestration that would assist in interpreting these empirical findings. This paper investigates listeners' emotional responses to four types of orchestral gestures—large-scale timbral and textural changes that occur in a coordinated, goal-directed manner—through an exploratory experiment that collected continuous responses of felt emotional intensity for musician and nonmusician listeners. A time series regression analysis was used to predict changes in emotional responses by modeling changes in several musical features, including instrumental texture, spectral centroid, loudness, and tempo. We demonstrate the application of a new visualization tool that compiles the emotional intensity ratings with score-based and performance-based musical features for qualitative and quantitative analysis. The results suggest that the response profiles differ for the four gestural types. Following the increasing growth of instrumental texture and loudness, the emotional intensity ratings climbed steadily for the gradual addition types. The ratings for the sudden addition gestures sharply increased in response to the rapid textural change, peaking toward the end of the excerpt. There was a slight tendency for musicians, but not nonmusicians, to anticipate the moment of sudden addition with heightened emotional responses. The responses to the reductive excerpts, both gradual and sudden, feature a plateau of lingering high emotional intensity, despite the decrease of other musical features. The visualization provided a method to observe the evolution of listeners' emotional reactions in response to the orchestral gestures and assisted in interpreting the results of the time series regression analysis.

Keywords

music and emotion, emotional intensity, orchestral gestures, musical timbre, instrumental texture

Introduction

Orchestration—the art of structuring music with timbre—became a fundamental aspect of musical composition as a structural and expressive device in the nineteenth century (Todd, 1986). The shaping of instrumentation has the capacity to demarcate formal divisions, to provide broad contrasts of weight and colour, and to supply expression and emotional content. As McAdams (2013) states:

Larger-scale changes in timbre can ... contribute to the expression of higher-level structural functions in music ... Orchestration can play a major role in addition to pitch and rhythmic patterns in the structuring of musical tension and relaxation schemas that are an important component of the aesthetic response to musical form (60).

Despite the evident integral function of timbre on the listening experience, a disparity exists between the importance of orchestration on the one hand and the amount of scholarly attention paid to the subject on the other. Empirical studies have reported that changes in instrumentation and texture induce strong emotional responses in listeners. In their analyses of Romantic musical excerpts where chills (i.e., *frisson*) were experienced in coordination with distinct patterns of heart rate and skin conductance, Guhn, Hamm, and Zentner (2007) found that the musical passages were marked by particular dynamic, harmonic, and structural characteristics. Among these musical features, passages featuring the contrast of a solo instrument and the orchestra induced strong emotional responses. In his study of survey responses based on self-reports, Sloboda (1991) reported ten musical features directly associated with emotional responses, including a sudden dynamic or textural change associated with chills, tears, and an increase in heart rate. Along these lines, Panksepp (1995) found that peaks in the number of experienced chills corresponded to dramatic crescendos and he proposed that the same feeling could be achieved through a striking reduction of instrumental forces (e.g., a solo instrument emerging out of the orchestra). Although these studies have made significant steps

towards investigating the links between music and emotion, the analyses rarely consider the musical features apart from the local, isolated context.

The findings from music and emotion research suggest that texture and instrumentation profoundly affect the listening experience, but a lacuna exists in the field of music theory related to orchestration, resulting in a limited understanding of the function and perception of large-scale orchestral shaping (Boulez, 1987; Sandell, 1995; Slawson, 1985). Goodchild (2016a, 2016b) has proposed a typology of orchestral gestures based on large-scale timbral and textural changes. As shown in a visual representation in Figure 1, the four types are defined by changes in instrumentation, organized into a two-by-two design based on time course (gradual or sudden changes) and direction (additive or reductive changes). The x-axis represents time and the y-axis represents the number of instrumental parts that are engaged (described further below). Each gesture adds or removes instruments, gradually or suddenly, as a coordinated change over time. The four types relate to descriptions from orchestration treatises of an “orchestral crescendo” (gradual addition), the reverse process (gradual reduction), and timbral contrasts including a rapid switch to full orchestral forces (sudden addition) and the drop-off to a contrast choir or soloist (sudden reduction). Goodchild (2016a, 2016b) characterizes large-scale orchestral shaping as gestures, representing their goal-directed sense of motion and their structural and expressive role in the musical discourse. The dramatic shaping of the orchestral gestures coordinates disparate musical elements (e.g., instrumental texture, dynamics, tempo and other musical parameters), creating a sense of agency and emotional force.

[Insert Figure 1]

The literature presents inconclusive predictions for listeners' emotional responses to orchestral gestures. Schubert (2004) suggests that there is a causal relationship between combinations of musical features and emotional response. In a regression study, he reported that between 33% and 73% of the variation in perceived arousal could be explained mainly by loudness and tempo, but there were inconsistent associations for instrumental texture and timbre. Gabrielsson and Lindström (2010) suggest that a predictable relationship exists between parameters of loudness and timbre on the one hand and emotional intensity on the other; listeners reported that higher degrees of loudness and upper harmonics were associated with emotions of higher intensity. Several studies have reported connections between chill responses and sudden or unexpected changes in texture and loudness (Grewe, Nagel, Kopiez, & Altenmüller, 2007; Guhn et al., 2007; Sloboda, 1991). In particular, one of the most common acoustical correlates of chills is a rapid increase in loudness (Huron & Margulis, 2010).

Given these experimental results, the nature and extent of listeners' emotional responses to the various types of orchestral gestures is unclear. How do variations in musical features that shape orchestral gestures contribute to emotional response profiles of listeners? Would the additive gestures result in higher emotional intensity on average compared to reductive gestures, since environmental sounds moving towards the listener (i.e., looming sounds) have been found to be more biologically salient than those moving away (Hall & Moore, 2003)? It is uncertain whether there would be differences in the number of chills experienced for the additive gestures compared to the reductive ones. Additionally, would listeners' experiences of orchestral gestures differ based on the degree of musical training? Previous research has shown that nonmusicians have been found to be equally as receptive to the perceived emotional quality of musical excerpts (Bigand & Poulin-Charronnat, 2006; Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005);

however, musicians were also found to be more sensitive to changes in many musical features in a recent study on perceived musical tension (Farbood, 2012).

In this exploratory behavioural experiment, we investigated listeners' experiences of orchestral gestures through continuous ratings of felt emotional intensity and retrospective ratings. We employed two approaches to investigate how emotional responses to large-scale orchestral gestures unfold in time in relation to various musical features. First, a time series regression study was used to predict changes in emotional response by modeling changes in various musical features (such as instrumental texture, loudness, and tempo). Second, we developed a new type of visualization that illustrates the relative textural density of each instrumental family over time and maps musical-feature overlays of time-varying musical parameters and the emotional intensity ratings as an interpretive analytical tool. Our aim is to investigate one aspect of the temporal dynamics of the listening experience by researching the musical features in orchestral music that elicit emotional responses.

A multitude of personal factors may influence an individual's affective responses to music (Eerola & Vuoskoski, 2013). Therefore, we collected retrospective ratings in order to account for moderating factors of preferences and familiarity with musical stimuli. To complement the continuous ratings, participants also reported the number of experienced chills, defined as goose bumps, shivers, or tingles down the spine. Although rare, chills provide insight into strong emotional experiences, which are powerful, salient, and memorable (Whaley, Sloboda, & Gabrielsson, 2009). To our knowledge, this study is among the first to isolate various textural and timbral changes in relation to musical structure and to explore their potential affective responses.

Method

Participants

We recruited 45 participants (22 musicians and 23 nonmusicians) from the Montreal area via McGill University email lists and classified advertisements. The musician listeners (11 females) had an average age of 24 years ($SD = 5.7$) and were required to have studied a minimum of 2 years in an undergraduate degree in music in order to obtain a high level of instrumental and theoretical training. The nonmusician listeners (12 females) had an average age of 24 years ($SD = 10.3$) and had less than two years of musical training in early childhood. Musician and nonmusician listeners were included in order to assess the role of musical expertise on listening experience. All participants were pre-screened to ensure normal hearing and a history of strong emotional responses to music (Whaley et al., 2009).

Stimuli

Twelve musical excerpts were chosen to fit within the four types of gestures (three excerpts per type), as shown in Table 1. The musical passages are categorized by only one overarching orchestral gesture (as described above) and were drawn from the symphonic repertoire of the late nineteenth and early twentieth centuries, where large-scale gestures were likely to occur. The excerpts were 1 to 3 minutes in duration in order to provide a listening context. Well-known excerpts were avoided to minimize the possibility that participants would respond to personal or external associations (such as episodic memory and evaluative conditioning mechanisms) (Juslin & Västfjäll, 2008). All excerpts were extracted from commercial CD recordings in uncompressed waveform audio file format (WAV) sampled at 44.1 kHz with 16-bit amplitude resolution. The recordings were chosen qualitatively for power of

expression. A two-second fade-in or fade-out was added with Audacity (Audacity Team, audacityteam.org) if the excerpt did not begin or end with silence, respectively.

The stimuli were stored on a Mac Pro computer (Apple Computer, Inc., Cupertino, CA), amplified through a Grace Design m904 monitor (Grace Design, Boulder, CO), and presented over Dynaudio BM6a loudspeakers (Dynaudio International GmbH, Rosengarten, Germany). In separate pilot experiments, the level of each excerpt was set to a comfortable level. Sound levels for each excerpt were measured with a Brüel & Kjær Type 2205 sound-level meter (A-weighting) placed at the level of the listener's ears in the front and rear rows (the experimental context is described in the next section), with maximum values for each excerpt ranging between 81.6 and 97.0 dB ($M = 91.4$, $SD = 3.9$).

[Insert Table 1]

Procedure

The experimental sessions took place at the Music Perception and Cognition Lab at the Schulich School of Music of McGill University. Six group-listening sessions were held to represent a concert experience. In each session, a maximum of eight participants listened in a dimmed room with acoustic absorbing material on the walls at a comfortable temperature (22°C). On arrival participants were randomly assigned a seat within three rows of five chairs facing the two loudspeakers arranged at $\pm 45^\circ$. An empty chair between participants reduced distraction and allowed the placement of the questionnaire and slider (described below) when not in use.

Participants were asked to sign an informed-consent form and read the instructions on a clipboard. The first part of the experiment consisted of two practice trials, which allowed participants to ask questions and to familiarize themselves with the experimental procedure. The

two practice pieces consisted of a gradual addition excerpt from Scriabin's Symphony 1, movement III, and a sudden reduction excerpt from "Great Gate of Kiev" from *Pictures at an Exhibition* by Mussorgsky, orchestrated by Ravel. Participants were told that the practice pieces were representative of the music they would hear in the main experiment. The data from the practice stage were not analyzed. The main experiment consisted of the 12 stimuli (listed in Table 1 above) played in a randomized order for each session. The experimental session lasted approximately one hour, and the listeners were paid for their participation.

While listening to each excerpt, the participants used a handheld slider to continuously rate emotional intensity that they felt in response to the music. They were instructed to rate the strength of their emotional response regardless of whether it was a positive or negative feeling. They were told that the slider position should reflect a personal reaction to the music and not what they imagine the music is trying to express. In the typical dimensional model of emotion, emotions are understood as varying degrees of arousal (low to high) and valence (negative to positive) (Eerola & Vuoskoski, 2013; Russell, 1980; Schubert, 2004). Following previous work in the field, the unidimensional emotional intensity scale would seem to tap into arousal or activation of emotion, rather than the valence component (McAdams, Vines, Viellard, Smith, & Reynolds, 2004; Schubert, 2010). Given the exploratory nature of the experiment, we were interested in the evolution of emotional intensity over time and its connection to the musical features that sculpt the orchestral gestures.

After each excerpt, participants completed retrospective questionnaires indicating their familiarity, preference, and the number of chills experienced. Questions about several other subjective experiences will not be reported here due to the scope of the paper. After the main

experiment concluded, participants completed a final questionnaire outlining their musical training, listening habits and preferences, and socio-demographic characteristics.

Apparatus

The experiment was programmed with the PsiExp computer environment (Smith, 1995). Slider responses were measured simultaneously using a MIDI-based acquisition hardware, which converted the slider position to a 7-bit value and sent MIDI timings to the controlling computer (maximum response latency of 10 ms) (Fléty, 2002). A custom-made handheld mechanical slider box provided force feedback through physical resistance with an elastic band. As the slider was moved in the direction of more emotional intensity, the participant increasingly felt more resistance to the movement. The elastic band would return the slider to the starting position if released. In pilot studies, participants reported that the force feedback on the slider was an intuitive method of rating that obviated the need for visual attention to the device. Participants held the slider box on their lap and moved the slider forward and backward to indicate higher and lower emotional intensity, respectively. They began each trial with the slider at the bottom (low emotional intensity) and were told that the maximum intensity point at the “strong” end of the scale, where the slider can be pushed no further, should correspond to the strongest emotional response to orchestral music they have experienced. This anchoring was chosen to ensure that participants used the majority of the range of the slider while preventing them from overuse of the maximum end that would create a flat-line response.

Data analysis

Emotional intensity slider responses

The emotional-intensity slider data were resampled at 2 Hz using nearest neighbor interpolation. The data were range-normalized between 0 and 1 for each participant across all excerpts. Normalization reduces the inter-participant variability due to different ranges of rating. Further data processing is described below.

Musical features

Musical features were coded as time series to be included in the regression analysis and in the visualizations. Three performance-based features (loudness, spectral centroid, and tempo) and three score-based features (instrumental texture, onset density, and contour) were selected due to their importance in descriptions of orchestral gestures in treatises and in similar research (Adler, 2002; Schubert, 2004). Loudness and spectral centroid were extracted from the commercial recording audio files (16-bit, 44.1 kHz) with PsySound3 (Cabrera, Ferguson, & Schubert, 2007) which analyzes as a default setting in a succession of 93 ms overlapping windows (23 ms hop size). Using the auditory filter model of steady-state sounds by Moore, Glasberg, and Baer (1997), the loudness was calculated in sones, calibrated to the measured sound pressure level in dB of each excerpt (as explained above). The loudness parameter, therefore, provides a measure of the variation in dynamics during the course of the excerpt. Spectral centroid was extracted in Hz (frequency), measuring the center of gravity of the relative weights of the frequencies present in the signal. Spectral centroid has a robust connection to perceived brightness (McAdams, 2013).

Tempo, instrumental texture, onset density, and contour were manually coded. Using Sonic Visualiser (Cannam, Landone, & Sandler, 2010), the inter-onset interval (IOI) between each beat (at event n) from the recording was marked by the first author, independently verified by a musician colleague, and then converted into time series in beats per minute (bpm) using the following equation:

$$\text{bpm} = 60 \cdot \left(\frac{1}{\text{IOI}}\right) \quad (1)$$

where:

$$\text{IOI} = \text{onset time}_{n+1} - \text{onset time}_n \quad (2)$$

If there were no instruments with an attack on a particular beat, the timing was interpolated based on the surrounding context. Automated methods of beat extraction were unsuccessful due to the textural density of the orchestral sound files.

Following Schubert's (2004) approach, instrumental texture was objectively coded as the number of independent voices (or parts) sounding at each beat calculated from the score. An instrument that doubles the same pitch of another player of the same instrument was omitted from the count (e.g., flutes 1 and 2 on the same pitch would count as one voice, but a flute and oboe on the same pitch would count as two voices). If one instrument produces two or more simultaneous notes, each is counted. Notes that are attacked on a previous beat and sustained into the beat under evaluation are included in the count.

For the visualizations, instrumental texture was supplemented to include information about the contributions of each instrumental family (strings, woodwinds, brass, and percussion), as well as the organ or harp. The texture parameter is displayed as a stacked bar graph with the

instrument family indicated by colour. With this approach, one can view the textural thickness of each family and its evolution over time.

The onset density counts the number of onsets or attacks per beat, following a similar procedure as instrumental texture. Onset density distinguishes between sustained pitches or long note values (low onset density) and short note values or rapid figuration (high onset density) in sections that may have the same count for instrumental texture. At every beat, the contour variable tracks the highest and lowest notated pitches. The distance in semitones between the highest and lowest pitches, also known as the *ambitus*, provides a measure of the range of the pitch content over time.

The data processing was conducted in MATLAB (Mathworks, Inc). All graphs were created in Excel (Microsoft, 2013) and Illustrator (Adobe Systems, 2007). All of the musical feature time series were resampled at 2 Hz. For the regression study, all of the features and slider data were filtered to smooth the high-frequency noise and remove extraneous information using a low-pass digital filter (4th-order phaseless Butterworth) with a cutoff frequency at 0.2 Hz. For the visualizations, onset density, and contour were not smoothed due to the discrete nature of these values. For all of the excerpts, the first two seconds of data were cut, except for excerpt 3 in which the first five seconds were cut. The cut was due to fact that the beginnings of the recordings have only a few instruments playing, which created a spurious high spectral centroid from the background noise of the recording. Two to five seconds was deemed to be an acceptable data loss, especially given the initial orientation time, a period of acclimatization during which time the responses are unreliable, for a minimum of 2 seconds (median of 8 seconds) (Schubert, 2012).

Results

Retrospective Ratings

Familiarity

The familiarity scores as recorded on a seven-point scale (1 = not at all, 7 = extremely) were low overall, with a global median of 1 (range = 6.0). A nonparametric Mann-Whitney U test was run to determine if there were differences in median familiarity ratings between the musicians and nonmusicians for each excerpt. Table 2 reports the results along with the median and range. Compared to the nonmusicians, musicians were more familiar with E04, E05, E09, and E12.

[Insert Table 2]

Preference

The preference scores as recorded on a seven-point scale were moderate overall, with a global median of 4.0 (range = 6.0). A Mann-Whitney U test was run to determine if there were differences in median preference ratings between musicians and nonmusicians (Table 3). No statistically significant differences in ratings were found.

[Insert Table 3]

Chills

Participants experienced between 0 and 18 chills across the entire experiment, with a mean across all excerpts of 4.18 ($SD = 5.01$). Twenty-nine participants (64.4%) experienced at least one chill response.

A Friedman nonparametric test compared the median number of reported chills among the four experimental conditions and a statistically significant difference was found, $\chi^2(3) = 22.1$, $p < .001$. Post-hoc Wilcoxon signed-rank tests with Bonferroni adjustment for multiple tests revealed that the sudden addition type had significantly more chills reported compared to the gradual reduction ($Z = -3.2$, $p = .001$), sudden reduction ($Z = -2.8$, $p = .006$), and gradual addition ($Z = -2.5$, $p = .01$). There were no significant differences among the other gestural types.

Emotional intensity ratings

Median

Medians were calculated from the normalized continuous rating data for each participant for each of the 12 excerpts. A grand mean value across excerpts within each category was then calculated for each participant. The means and 95% confidence intervals are shown in Figure 2. The reductive gestures are rated with higher emotional intensity than additive ones, but the effect of time course appears to depend on direction and participant group. A mixed-design ANOVA was performed to determine whether repeated-measures orchestral gesture parameters of time course (addition vs. reduction) and direction (addition vs. reduction), as well as the between-subjects factor of musical training, had an effect on participants' median emotional intensity ratings. In addition to a significant main effect of direction, $F(1, 43) = 22.0$, $p = .001$, $\eta_p^2 = 0.29$, a significant three-way interaction was found, $F(1, 43) = 5.80$, $p = .01$, $\eta_p^2 = 0.14$. To determine the nature of the interaction, two tests of simple interactions were performed on the ratings of the musicians and nonmusicians separately, with time course and direction as repeated-measures factors. A significant main effect of direction was found for the musicians, $F(1, 22) = 13.7$, $p =$

.001, $\eta_p^2 = 0.38$, and nonmusicians, $F(1,21) = 6.17, p = .02, \eta_p^2 = 0.23$. Post-hoc analyses with Bonferroni adjustment revealed that the median ratings were lower for the addition type compared to the reduction type for musicians ($M = 0.92$, 95% CI [0.41, 0.14], $p = .001$) and nonmusicians ($M = 0.85$, 95% CI [0.01, 0.16], $p = .02$). Given that the main effect of direction (additive and reductive) was preserved for both musicians and nonmusicians in the tests of simple interactions, the interaction appears to be modulated by a difference in ratings between the musicians and nonmusicians for the time course (gradual and sudden): musicians rated gradual reductive excerpts higher than gradual additive excerpts, whereas nonmusicians rated sudden reductive higher than sudden additive.

[Insert Figure 2]

Interquartile range

In order to investigate the variability of ratings, we calculated the interquartile range as a measure of statistical dispersion for the time series of emotional intensity ratings for each participant for each of the 12 excerpts. A grand mean value for each category was then calculated. A mixed-design ANOVA was performed to determine whether time course (gradual and sudden), direction (addition and reduction), and musical training had an effect on the variability of emotional intensity ratings. A significant main effect of direction was found, $F(1, 43) = 27.7, p = .001, \eta_p^2 = 0.39$. Post-hoc analysis with a Bonferroni adjustment revealed that the variability for the additive type ($M = 0.27$) was significantly higher than for the reduction type ($M = 0.20$), $p = .001$. No significant effect of training was found, indicating that the variability of the musicians' and nonmusicians' ratings were similar. Furthermore, the effect of

time course was not significant, suggesting that the variability was not different between gradual and sudden gestures.

Emotional intensity profiles

We graphed the average emotional intensity ratings for the musicians and nonmusicians to examine the evolution of emotional responses over time (see Figures 3 and 4). Colour versions of Figures 3 and 4 can be found in the supplemental material. The shaded clouds around the musicians' and nonmusicians' curves represent ± 1 standard error of the mean (Campbell, 2010).

The emotional intensity ratings of the gradual addition type (Figure 3) generally follow the gestural shape of the orchestration changes with an increasing trajectory towards the end of the excerpt. The ratings of the gradual reduction gestures, in contrast, reach a plateau and do not completely diminish at the end of the excerpt. Given the overlapping areas of the standard error, there are very few differences between the ratings of two groups for the gradual excerpts. Exceptions include excerpt E02, in which the nonmusicians' ratings are higher in the middle section and excerpt E04, in which the musicians' ratings feature higher emotional intensity in the second half of the excerpt.

Similar to the gradual addition type, the responses to the sudden addition gestures (Figure 4) increase towards the end of the excerpt, but feature a steeper slope surrounding the moment of textural change (shown by the vertical dotted line). The musicians' ratings are initially higher for excerpt E07 and generally higher for E09. The sudden reduction gestures share a similar characteristic with those of the gradual reduction type in that they both present a plateau of high emotional intensity ratings after the reduction of forces. This plateau is more pronounced for the musicians in E10.

[Insert Figures 3 and 4]

Time series regression analysis

Given the difficulty in interpreting a single measure of central tendency of a time-varying behavioural measure, we employed a time-series analysis approach to explore the contribution of musical feature variables on continuous ratings of emotional intensity. Schubert's (1999, 2004) quantitative modeling procedure employs the ordinary least squares approach to linear regression and takes into account response delay and serial correlation. We conducted two linear regression models of the musicians' and nonmusicians' emotional responses. One excerpt was chosen as the most representative example (in terms of the shape of the texture parameter) for each orchestral gesture type: E01 for gradual addition, E04 for gradual reduction, E07 for sudden addition, and E10 for sudden reduction.

Four musical features were used as predictors of the emotional response ratings in the regression analysis: loudness, spectral centroid, tempo, and instrumental texture (sum of the number of active instruments). We limited the number of predictors given that the likelihood of a type I error is increased with a greater number of variables included in the model.

To minimize serial correlation, we applied a first-order difference transformation to all of the variables, which produces a gradient time series indicating the amount of change in the variable for each sample point. To address response delays, we duplicated each musical feature variable at lags of 0, 1, 2, 3, and 4 seconds, thereby generating 20 different predictors (four variables at five lags).

We adapted Schubert's (1999) syntax in SPSS Statistics (IBM, Armonk, NY) to conduct a stepwise ordinary-least squares (OLS) regression in order to determine appropriate predictors.

Next, the residual was diagnosed for serial correlation using an autocorrelation function. Given that all of the residuals were serially correlated, we conducted a first-order autoregressive adjustment (AR1) that consists of a linear combination of the previous error term. Finally, we analyzed the residual of this combination using the Durbin-Watson statistic and an autocorrelation function to determine if the effects of serial correlation had been removed.

Table 4 reports the first-order autoregression models using a stepwise method for the emotional intensity ratings of the musicians and nonmusicians. All but one of the autoregressive models have significant coefficients, but only a small portion (from 1% to 17%) of the variance was explained. The one exception is the model for the nonmusicians' average response to E01 (gradual addition), which revealed no significant results for any of the features. Nearly all of the musical feature coefficients were positive, indicating changes in emotional intensity occur in the same direction as changes in the other variables; however, the model for nonmusicians in E10 (sudden reduction) revealed that loudness (lagged at 2 seconds) decreased as emotional intensity increased.

The Durbin-Watson statistic values were close to 0 (revealing positive serial correlation), indicating that the difference transformation and the autoregressive adjustment were not able to remove completely the effects of serial correlation. All of the autoregressive coefficients (AR1) were close to 1, which suggests that the emotional intensity ratings at a given moment were mainly determined by the preceding context, rather than changes in musical features (Gottman, 1981; Schubert, 2004). Therefore, the interpretation of the models needs to proceed with caution due to the presence of serial correlation.

[Insert Table 4]

Similar to Schubert's findings, tempo and loudness were the main musical features that were entered into the models. Loudness appeared in almost every model; in fact, loudness at various lags was the only musical feature included in the models for the gradual gesture types (E01 and E04). Tempo was included in the models for the sudden types (E07 and E10). For the features related to timbre, the texture predictor was not entered into any of the models and spectral centroid was only included in the models for E10.

For the gradual gesture types (E01 and E04), we speculate that the timbral parameters (instrumental texture and spectral centroid) and loudness were collinear, given that the level and spectral extent, both of which contribute to loudness, are directly affected by the number and variety of instruments. In regression models, collinearity is an issue when predictor variables are highly correlated. It can be diagnosed by calculating the tolerance, which is a reflection of the value of fit ($1-R^2$) (Schubert, 1999). All of the models' tolerance estimates were high (greater than 0.5), which indicates that there were no issues with excessive collinearity.

To investigate whether there were strong associations among the variables, Table 5 reports the Pearson correlation coefficients for the relationships between the loudness predictors entered into the regression models and texture and spectral centroid at various lags. Given that the coefficients ranged between .14 and .53, there were no strong relationships among the variables. However, there were some weak to moderate relationships between loudness and texture/spectral centroid for the models of E01 and E04.

[Insert Table 5]

Profiles of emotional intensity and musical parameters

A visualization tool was developed in order to facilitate the analysis of the emotion profiles in response to instrumentation changes and their interactions with score-based and performance-based musical parameters. The tool provides an exploratory method of analysis to complement other types of analytical approaches (e.g., time series regression analysis, music-theoretical analyses) by presenting a synoptic view of the excerpt, graphing together for visual comparison the mean emotional intensity ratings¹ and the musical feature variables. Although the visualizations cannot be used for inferential analytical techniques to test statistical hypotheses, they can be used for qualitative and quantitative interpretation of musical parameters and for discovering patterns by compiling diverse types of data in a meaningful way. The visualizations for E01, E04, E07, and E10 are presented as Figures 5-8, respectively. The colour visualizations for all twelve excerpts can be found in the supplemental material.

Gradual addition

The gradual addition examples build a large-scale orchestral crescendo by adding instruments progressively, contributing to a sense of orchestral growth. In excerpt E01 (Mahler Symphony 1, III, mm. 1-32), Mahler systematically adds part after part in canon, resulting in a gradual thickening of instrumental texture, and intensification in onset density, loudness, and spectral centroid (see Figure 5). The emotional intensity ratings of the musicians and nonmusicians begin with a steep incline that arches towards the end of the excerpt. Although the number of instruments remains small at the onset, the deliberate, slow-building orchestral swell likely contributes to the accumulation of emotional intensity.

[Insert Figure 5]

Gradual reduction

The gradual reduction gestures feature the progressive removal of instruments in a process of abatement. Excerpt E04 (Mussorgsky/Ravel, *Pictures at an Exhibition*, “Bydlo,” mm. 39-64) exemplifies this process of gradual lessening with the coordination of musical processes (see Figure 6): the instrumental parts gradually fade out, along with a reduction of the onset density, loudness, and spectral centroid. The ambitus narrows to a single instrument of the double bass.

The emotional intensity ratings do not mirror the reduction of musical features. Despite the evident concluding musical function, the responses exhibit responses of persistent high emotional intensity. This sustained emotional lingering effect is the reverse of what one might expect as a response to a receding environmental source (Hall & Moore, 2003).

[Insert Figure 6]

Sudden addition

The sudden addition category examples create a dramatic turning point in the ongoing musical trajectory. As shown in Figure 7, the sudden addition of E07 (Vaughan Williams, *London Symphony*, I, mm. 8-53) uses an anticipatory signal (shown with a bracket) before the sudden textural change, which likely prepares the listener for the upcoming dramatic event. The signal involves ascending arpeggios with instruments entering in successively higher registers, causing a marked increase in spectral centroid, loudness, tempo, and onset density. The use of silence after the anticipatory signal emphasizes the sudden addition of instrumental forces.

Musicians respond to the anticipatory signal by sharply increasing their slider ratings to reach their peak in advance of the sudden addition. For the nonmusicians, however, a similar spike occurs after the sudden addition, a delay of several seconds, indicating more of a reaction to the sudden increase in instrumental forces and loudness. The peak emotional intensity values do not occur until toward the end of the excerpt and not directly at the sudden addition. This may be related to the way the musical features slightly decrease and build further, expressively waxing and waning after the textural change.

[Insert Figure 7]

Sudden reduction

In the sudden reduction gestures, the moment of textural change marks a structural rupture in the dramatic trajectory of the excerpt. The main feature is the abrupt reduction from full forces to a subgroup of the orchestra and the sudden drop in loudness. As shown in Figure 8, excerpt E10 (Holst, *The Planets*, “Uranus,” mm. 193-236) features a wide spread between the high and low contour variables throughout, which emphasizes change from thick to thin texture after the drop-off. The spectral centroid varies based on the differing instrumental combinations. The tempo plummets before the textural change and continues to be held back, creating the effect of time being suspended.

The musicians’ and nonmusicians’ ratings demonstrate a sustained lingering effect of high emotional intensity that goes against the large reduction in musical forces. The high emotional intensity may be linked to the combination of the surprising textural change followed by the tense atmosphere in the reduced texture produced by the dissonant sustained string chords and ominous harp melody. The musicians’ ratings generally increase in intensity throughout;

however, the nonmusicians' ratings peak before the drop-off and trail behind the changes in loudness and tempo. Even the nonmusicians' ratings reach a plateau and do not completely plummet. As with the sudden additive excerpts, this result could suggest that nonmusicians are exhibiting more of a bottom-up approach that responds to surface features of the music than are musicians, as discussed by Ferguson, Schubert, and Dean (2011).

[Insert Figure 8]

Discussion

Retrospective ratings

Familiarity of the excerpts was low in general for both the musicians and nonmusicians. The musicians were significantly more familiar with excerpts E04, E05, E09, and E12. This familiarity may assist in explaining the higher emotional intensity ratings for E04 and E09, but not for excerpts E07 and E10, which may suggest that familiarity does not appear to consistently modulate the emotional intensity ratings.

As expected, participants experienced chills across all experimental excerpts, indicating that orchestral gestures are capable of inducing strong emotional responses. However, the number of chills was significantly higher for the sudden addition gestures compared to the other gestural types, including sudden reductions. This finding differs from the results of other studies that associated chills with sudden or unexpected changes in texture and loudness in either direction (Grewe et al., 2007; Guhn et al., 2007; Sloboda, 1991) and more closely aligns with the acoustical correlate of a loudness increase (Huron & Margulis, 2010). A study design that allows listeners to pinpoint chill onsets would provide more insight into the musical mechanisms.

Emotional intensity ratings

Based on the numerous findings in the literature that associated additive orchestration techniques (e.g., orchestral crescendo and sudden addition of full orchestral forces) with emotions of high intensity, we predicted that the emotional intensity ratings would be higher for the additive gestures compared to the reductive gestures. However, the ANOVA results indicated the opposite effect: the median emotional intensity ratings were higher for reductive gestures compared to additive ones. We also predicted that the variability, as measured by the interquartile range, would be higher in sudden excerpts, in which a dramatic change in orchestral texture would likely cause a distinct spike in emotional intensity. The ANOVA results revealed that direction, not time course, showed a distinction: the additive gestures had higher response variability compared to the reductive gestures.

Examination of the visualizations provides insight into these findings. For both gradual and sudden reductions, the emotional intensity ratings increased initially at the beginning of the excerpt followed by relatively little decrease in intensity over the remainder of the excerpt. This plateau of elevated emotional intensity resulted in higher median ratings and low variability. In contrast, the ratings generally climb over the course of the excerpt to the climax for the gradual and sudden addition gestures, resulting in lower median ratings and higher variability. This analysis highlights the importance of considering the evolution of emotional intensity ratings in context. Measures of central tendency are not able to capture the full picture, particularly when studying responses to excerpts with dynamic musical processes.

Time series regression analysis

Similar to Schubert's (2004) findings, the results of the exploratory time series regression indicated that loudness and tempo were the main musical features that were included in the models. For the gradual gestures, changes in emotional intensity occurred in the same direction as changes in loudness. For sudden gestures, tempo was included in the models in addition to loudness. Spectral centroid was only included for the sudden reduction type.

Although the predictors related to orchestration and timbre (instrumental texture and spectral centroid) were largely excluded from the models, this was not due to excessive collinearity with loudness. Given that there were low to moderate associations between these variables, the addition of texture and spectral centroid in the model likely did not increase the predictive power above that of loudness in the stepwise regression method. Therefore, the analysis does not necessarily indicate that timbral musical features do not play a role, but that loudness and tempo are the parameters that best explain the variance of the emotional responses in this exploratory study. Recompositions of the excerpts or specifically composed pieces that control for the different parameters could further investigate their individual contributions to emotional responses in future.

Despite the first order difference transformation and autoregressive adjustment, the effects of serial correlation were not completely removed in all of the models. Inspection of the visualizations reveals that there is not a direct one-to-one correspondence between the emotional intensity ratings and the individual musical feature variables employed. In his regression study, Schubert selected four pieces that generally capture a specific mood represented within the four quadrants of the two-dimensional emotion space of valence and arousal. In contrast, we chose pieces that feature and are characterized by large-scale expressive changes over time. The

direction and relative magnitude of the effect of each variable likely vary throughout the excerpt and may interact with the musical context created by other parameters that have not yet been included as predictors (e.g., harmony).

The low goodness-of-fit values for the sudden gestures are likely related to the variability and abrupt changes of the responses and musical features. We suspect that it may be useful to consider various transforms of the variables in future work to overcome issues with the non-linearity of the relations between the musical features and the behavioural data, but this approach is beyond the scope of the present study. Future work could also use the approach of Dubnov, McAdams and Reynolds (2006) to investigate the extent to which the emotional intensity profiles are associated with signal properties from the acoustical recordings, relating to overall sound colour or texture.

Another limitation of the regression study is related to the use of mean ratings of emotional intensity, which cannot account for response variability across participants. Activity analysis (Upham & McAdams, 2015) could be used in future to accompany this study to assess response coordination within and across the musician and nonmusician groups, as well as to pinpoint statistically significant moments of local activity for increasing or decreasing emotional intensity. This type of analysis would assist in disentangling instances where participants' responses are coordinated and where they diverge, and whether these instances relate to particular moments within the shaping of orchestral gestures as shown on the visualizations.

Visualizations

Each of the orchestral gesture types featured distinct emotional response profiles. For the gradual addition type, participants' ratings steadily climbed towards the end of the excerpt along

with increases in musical features, particularly instrumental texture and loudness. For the sudden addition type, the emotional intensity ratings increase sharply at the sudden textural change, but peak towards the end of the excerpt, which indicates that the interaction of musical features after the textural change contribute to the climax of the gesture.

The gradual reduction excerpts resulted in a progressive decrease in spectral centroid, loudness, texture, onset density, and ambitus range. The emotional effect of the gradual reduction, however, is not the reverse of the gradual addition. The emotional intensity ratings plateau and remain high even at the end of the excerpt. Similarly, the moment of sudden reduction comprises a structural rupture in the dramatic trajectory, resulting in persistent high emotional intensity despite marked drops in instrumental texture, loudness, and onset density. This lingering effect relates to findings by Gabrielsson (2010), who reported that the effects of strong experiences to music are long-lasting. Prolonged high emotional intensity is likely related to an “afterglow” effect, in which emotional arousal continues after dramatic events and even after the music has stopped (Schubert, 1999, 2012).

The visualizations reveal slight differences between the ratings of the musicians and nonmusicians in relation to the sudden gestures. Due to their greater exposure to orchestral music, the musicians may recognize anticipatory signals and, as a result, their emotional responses were heightened before the onset of the dramatic moment of sudden addition. The musicians may also be more sensitive to factors that are not yet coded, such as harmony. This finding is in line with previous studies, such as Steinbeis, Koelsch, and Sloboda (2006), who investigated the effect of harmonic expectancy violations on brain processing of emotional stimuli. They found that responses did not differ in amplitude between musicians and nonmusicians, but responses were earlier for musicians, suggesting that they had enhanced

processing abilities of harmonic expectancy violations and greater sensitivity to stylistic violations.

The nonmusicians appear to be more directly affected by low-level features such as loudness and tempo changes and trail behind variations in these features, as seen in the delay in the responses compared to musicians in excerpts E07 and E10. Similarly, Farbood's (2012) musical tension study found that musicians were more sensitive than nonmusicians to changes in certain musical features (e.g., harmony). As a result, musicians may have been more sensitive to the dissonant harmony after the sudden reduction in excerpt E10, for example.

We propose that the visualizations containing musical features and emotional responses provide a crucial tool for analysis by moving beyond anecdotal or descriptive accounts of musical processes. The exploratory approach provides a synoptic view of the excerpt, allowing for the investigation of score-based information (instrumental texture, onset density, and contour), performance-based features (loudness, spectral centroid, and tempo), and experimental data (emotional intensity ratings). In future work, other musical-feature overlays, such as phrase structure and harmony, will be included in order to better understand the interaction between orchestral texture and other musical parameters. Additionally, we plan on developing an automated method to create the visualizations from symbolic representations of scores and audio recordings.

The visualizations could be adapted for other analytical contexts. For example, multiple recordings could be consulted in order to examine the role that performance features play on the emotional force of the gestures. Expressive factors, such as performance timing, dynamic changes and instrumental balance, contribute to emotional responses and can highlight certain compositional choices, such as the anticipatory signal before a sudden addition or the surprise of

the moment of the sudden reduction. Using the visualizations for comparison of particular recordings and the resulting emotional responses would be useful in this regard.

This exploratory study provides a foundation for a theory of orchestral gestures that focuses on the role of orchestration and timbre (an underdeveloped research topic), considers the musical context and progression of musical features over time, and incorporates listener experiences. We explored listeners' responses to orchestral gestures in order to gain a better understanding of the evolution of emotional intensity in connection to orchestration changes and other musical features. The four types of orchestral gestures provide a framework to interpret previous empirical findings related to gradual and sudden changes in texture and instrumentation and a starting point for future music theoretical and experimental exploration. The visualization tool offers a qualitative and quantitative method to study score-based features, performance-based features, and emotional responses within a large-scale context, an aspect of music and emotion research that is in need of further study.

Acknowledgements

Color versions of Figures 3-8 and visualizations for all twelve excerpts are available as Supplemental Online Material, which can be found attached to the online version of this article at <http://msx.sagepub.com>. Click on the hyperlink "Supplemental material" to view the additional files.

We would like to thank Emery Schubert for his helpful suggestions on a previous version of the manuscript, Bennett K. Smith for programming the experimental interface, and Ryan Ouckama for data processing. Many thanks to Jonathan Crellin for his assistance in running the

experimental sessions and with the initial data analysis. We also appreciate the support of the Marvin Duchow Music Library staff in locating musical examples.

Preliminary analyses of the experimental data in this chapter were presented and reported in the conference proceedings at the 12th International Conference on Music Perception and Cognition (ICMPC12) in Thessaloniki, Greece in July 2012 and at the 21st International Congress on Acoustics (ICA 2013) in Montreal, Canada in June 2013.

Funding Acknowledgement

This work was supported by a Doctoral Fellowship from the Social Sciences and Humanities Research Council (SSHRC) of Canada, the Archie Malloch Graduate Fellowship from the Institute for the Public Life of Arts and Ideas, and student awards from the Centre for Interdisciplinary Research in Music Media and Technology to Meghan Goodchild, a SSHRC Standard Grant and an FRQSC Team Grant to Stephen McAdams and Jonathan Wild, and a Canada Research Chair awarded to Stephen McAdams.

References

- Adler, S. (2002). *The study of orchestration* (3rd ed.). New York: W.W. Norton and Company Inc.
- Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100, 100–130.
- Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion*, 19(8), 1113–1139.
- Boulez, P. (1987). Timbre and composition – Timbre and language. *Contemporary Music Review*, 2, 161–171.
- Cabrera, D., Ferguson, S., & Schubert, E. (2007). PsySound3: Software for acoustical and psychoacoustical analysis of sound recordings. In *Proceedings of the 13th International Conference on Auditory Display* (pp. 356–363). Montreal, Canada.
- Campbell, R. (2010). shadedErrorBar. Retrieved from <http://www.mathworks.com/matlabcentral/fileexchange/26311-shadederrorbar>
- Cannam, C., Landone, C., & Sandler, M. (2010). Sonic Visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the ACM Multimedia Conference*.
- Dubnov, S., McAdams, S., & Reynolds, R. (2006). Structural and affective aspects of music from statistical audio signal analysis. *Journal of the American Society for Information Science and Technology*, 57(11), 1526–1536.
- Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception*, 30(3), 307–340.
- Farbood, M. M. (2012). A parametric, temporal model of musical tension. *Music Perception*, 29(4), 387–428.
- Ferguson, S., Schubert, E., & Dean, R. T. (2011). Continuous subjective loudness responses to reversals and inversions of a sound recording of an orchestral excerpt. *Musicae Scientiae*, 15(3), 387–401.
- Fléty, E. (2002). AtoMIC Pro: A multiple sensor acquisition device. *Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME 2002)*, 1–6.
- Gabrielsson, A. (2010). Strong experiences with music. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 547–574). New York: Oxford University Press.
- Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367–400). New York: Oxford University Press.
- Goodchild, M. (2016a). Orchestral gestures. *Manuscript Submitted for Publication*.
- Goodchild, M. (2016b). *Orchestral gestures: Music-theoretical perspectives and emotional responses*. PhD dissertation, McGill University, Montreal, Canada.
- Gottman, J. M. (1981). *Time-series analysis: A comprehensive introduction for social scientists*. Cambridge: Cambridge University Press.
- Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Listening to music as a re-creative process: Physiological, Psychological, and Psychoacoustical Correlates of Chills and Strong Emotions. *Music Perception*, 24(3), 297–314.
- Guhn, M., Hamm, A., & Zentner, M. (2007). Physiological and musico-acoustic correlates of the

- chill response. *Music Perception*, 24(5), 473–484.
- Hall, D. A., & Moore, D. R. (2003). Auditory neuroscience: The salience of looming sounds. *Current Biology*, 13(3), R91–R93.
- 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). New York: Oxford University Press.
- Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *The Behavioral and Brain Sciences*, 31(5), 559–75.
- McAdams, S. (2013). Musical timbre perception. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 35–67). San Diego: Academic Press.
- McAdams, S., Vines, B. W., Viellard, S., Smith, B. K., & Reynolds, R. (2004). Influences of large-scale form on continuous ratings in response to a contemporary piece in a live concert setting. *Music Perception*, 22(2), 297–350.
- Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45(4), 224–240.
- Panksepp, J. (1995). The emotional sources of “Chills” induced by music. *Music Perception*, 13(2), 171–207.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- Sandell, G. J. (1995). Roles for spectral centroid and other factors in determining “blended” instrument pairings in orchestration. *Music Perception*, 13(2), 209–46.
- Schubert, E. (1999). *Measurement and time series analysis of emotion in music*. Unpublished Doctoral Dissertation, University of New South Wales, Australia.
- Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, 21(4), 561–585.
- Schubert, E. (2010). Continuous self-report methods. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 223–253). New York: Oxford University Press.
- Schubert, E. (2012). Reliability issues regarding the beginning, middle and end of continuous emotion ratings to music. *Psychology of Music*, 41(3), 350–371.
- Slawson, W. (1985). *Sound color*. Berkeley: University of California Press.
- Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, 19(2), 110–120.
- Smith, B. K. (1995). An environment for psychoacoustic experimentation using the IRCAM musical workstation. In *Proceedings of the Society for Music Perception and Cognition Conference*. Berkeley, CA: University of California.
- Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18(8), 1380–1393.
- Todd, R. L. (1986). Orchestral texture and the art of orchestration. In J. Peyser (Ed.), *The orchestra: Origins and transformations* (pp. 191–226). New York: Charles Scribner’s Sons.
- Upham, F., & McAdams, S. (2015). Activity analysis and coordination in continuous responses to music. *Manuscript Submitted for Publication*.
- Whaley, J., Sloboda, J. A., & Gabrielsson, A. (2009). Peak experiences in music. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 452–461). New York: Oxford University Press.

Endnote

¹The emotional intensity ratings in Figures 5-8 are based on those reported in Figures 3 and 4. However, they have been normalized between 0 and 1 to show the maximum reached for the excerpt. This normalization was performed for clarity of presentation and to focus on the changes in ratings rather than on the absolute values.

Table Captions

Table 1. List of stimuli by gesture type, excerpt number (#), composer, piece, measure numbers (mm.), recording information, and duration (min:s).

Table 2. Median and range for familiarity ratings by musicians and nonmusicians; Mann-Whitney U statistic comparing differences in training, effect size (r), and p -value with Bonferroni correction (* $p < .004$).

Table 3. Median and range for preference ratings by musicians and nonmusicians; Mann-Whitney U statistic comparing differences in training, effect size (r), and p -value with Bonferroni correction (* $p < .004$).

Table 4. Time series first-order autoregression models summary for E01, E04, E07, and E10.

Table 5. Pearson coefficients for cross-correlations with significant p -values between musical feature variables of loudness and timbral parameters (spectral centroid and texture) for the gradual gestural types.

Figure Captions

Figure 1. Four types of orchestral gestures.

Figure 2. Grand mean emotional intensity ratings with 95% confidence intervals for musicians and nonmusicians.

Figure 3. Average emotional intensity ratings for musicians (black line) and nonmusicians (dotted line) for gradual addition and gradual reduction excerpts. Shaded regions indicate ± 1 standard error from the average of musicians and nonmusicians, respectively. Overlapping areas are represented by darker shading.

Figure 4. Average emotional intensity ratings for musicians (black line) and nonmusicians (dotted line) for sudden addition and sudden reduction excerpts. Shaded regions indicate ± 1 standard error from the average of musicians and nonmusicians, respectively. Overlapping areas are represented by darker shading. Vertical dotted line represents the moment of sudden textual change.

Figure 5. Visualization of Mahler, Symphony No. 1, III, mm. 1-32, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings.

Figure 6. Visualization of Mussorgsky/Ravel, *Pictures at an Exhibition*, “Bydlo,” mm. 21-64, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings.

Figure 7. Visualization of Vaughan Williams, *London Symphony*, I, mm. 8-53, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings. The bracket indicates the anticipatory signal and the vertical dotted line indicates the moment of sudden textural change.

Figure 8. Visualization of Holst, *The Planets*, “Uranus,” mm. 193-236, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings. The vertical dotted line indicates the moment of sudden textural change.

Tables

Table 1.

List of stimuli by gesture type, excerpt number (#), composer, piece, measure numbers (mm.), recording information, and duration (min:s).

Type	#	Composer	Piece	mm.	Recording	Duration
Gradual Addition	E01	Gustav Mahler	Symphony 1, III (1888)	1-32	New York Philharmonic, conducted by Leonard Bernstein, Sony Classical, 2001	1:50
	E02	Gustav Mahler	Symphony 2, I (1894)	254-302	New York Philharmonic, conducted by Leonard Bernstein, Sony Classical, 2001	2:28
	E03	Jean Sibelius	Symphony 1, I (1900)	1-96	Lahti Symphony Orchestra, conducted by Osmo Vanska, BIS, 1995	2:53
Gradual Reduction	E04	Modest Mussorgsky (orchestrated by Maurice Ravel)	<i>Pictures at an Exhibition</i> , “Bydlo” ([1874] 1922)	39-64	Berliner Philharmoniker, conducted by Herbert von Karajan, Deutsche Grammophon, 1964	1:59
	E05	Gustav Mahler	Symphony 2, I (1894)	80-116	New York Philharmonic, conducted by Leonard Bernstein, Sony Classical, 2001	2:03
	E06	Richard Strauss	<i>Ein Heldenleben</i> , III (1898)	331-353	London Symphony Orchestra, conducted by Sir John Barbirolli, EMI Classics, 1996	2:09
Sudden Addition	E07	Ralph Vaughan Williams	<i>London Symphony</i> , I (1914)	8-53	Royal Liverpool Philharmonic Orchestra, conducted by Vernon Handley, EMI Classics, 2011	2:42
	E08	Richard Strauss	<i>Tod und Verklärung</i> , II (1889)	360-395	London Symphony Orchestra, conducted by Jascha Horenstein, Chandos Records, 1987	1:35
	E09	Claude Debussy	<i>La Mer</i> , I (1905)	122-141	Berliner Philharmoniker, conducted by Simon Rattle, EMI Classics, 2005	2:02
Sudden Reduction	E10	Gustav Holst	<i>The Planets</i> , “Uranus” (1916)	193-236	London Philharmonic, conducted by Sir Adrian Boult, EMI Classics, 2002	1:19
	E11	Anton Bruckner	Symphony 8, I (1890)	221-270	Wiener Philharmoniker, conducted by Pierre Boulez, Deutsche Grammophon, 2000	1:40
	E12	Antonín Dvořák	Symphony 9, I (1893)	200-286	Houston Symphony, conducted by Christoph Eschenbach, EMI Records/Virgin Classics, 2000	1:26

Table 2.
Median and range for familiarity ratings by musicians and nonmusicians.

		Familiarity Ratings											
		E01	E02	E03	E04	E05	E06	E07	E08	E09	E10	E11	E12
Musician	Mdn	5.0	1.0	1.0	2.0	2.0	1.0	1.0	1.0	4.0	1.0	1.0	5.0
	Range	6.0	6.0	6.0	6.0	6.0	6.0	5.0	3.0	6.0	6.0	3.0	6.0
Nonmusician	Mdn	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Range	6.0	3.0	4.0	5.0	1.0	3.0	5.0	5.0	3.0	4.0	2.0	3.0
Mann-Whitney	<i>U</i>	179	197	207	139	141	216	235	233	136	214	231	93
	<i>r</i>	—	—	—			—		—		—		
	<i>p</i>	.26	.23	.18	-.43	-.45	.16	-.08	.09	-.46	.18	-.11	-.56
		.08	.12	.23	.004*	.002*	.29	.60	.56	.002*	.23	.45	.001*

Note. Mann-Whitney U statistic comparing differences in training, effect size (*r*), and *p*-value with Bonferroni correction (* $p < .004$).

Table 3.
Median and range for preference ratings by musicians and nonmusicians.

		Preference Ratings											
		E01	E02	E03	E04	E05	E06	E07	E08	E09	E10	E11	E12
Musician	Mdn	4.0	4.0	5.0	4.0	5.0	5.0	5.0	4.0	5.0	5.0	4.0	5.0
	Range	5.0	6.0	5.0	6.0	6.0	6.0	6.0	6.0	4.0	6.0	6.0	6.0
Non-Musician	Mdn	5.0	4.5	4.0	4.0	4.0	4.0	5.0	3.5	5.0	2.0	3.5	4.5
	Range	5.0	7.0	5.0	5.0	6.0	5.0	6.0	6.0	5.0	6.0	5.0	6.0
Mann-Whitney	<i>U</i>	217	193	176	175	164	203	213	189	187	94	185	183
	<i>r</i>	-.12	-.14	-.27	-.24	-.31	-.18	-.14	-.16	-.23	-.55	-.24	-.21
	<i>p</i>	.40	.34	.07	.11	.04	.24	.35	.29	.12	.01	.11	.15

Note. Mann-Whitney U statistic comparing differences in training, effect size (*r*), and *p*-value with Bonferroni correction (* $p < .004$).

Table 4.
Time series first-order autoregression models summary for E01, E04, E07, and E10.

Excerpt	Model	Model Fit Summary			Regression Coefficients		
		AR1	Adj. R ²	<i>d</i>	MF	β	<i>t</i>
E01 (gradual addition)	Musicians	.95	.16	.38	Loud_2	.68	6.0**
					Loud_4	.53	6.0**
					Loud	.40	4.6**
	Nonmusicians				/	/	n.s.
E04 (gradual reduction)	Musicians	.97	.09	.43	Loud_2	.23	2.3*
	Nonmusicians	.96	.16	.39	Loud	.47	6.5**
					Loud_2	.34	4.2**
					Loud_4	.14	2.0*
E07 (sudden addition)	Musicians	.94	.01	.27	Temp_2	.12	2.1*
					Loud_4	.12	2.0*
	Nonmusicians	.96	.08	.26	Temp_2	.27	4.7**
					Loud_2	.21	3.5*
E10 (sudden reduction)	Musicians	.93	.14	.40	Cent_1	.39	5.0**
	Nonmusicians	.92	.17	.74	Loud_2	-3.7	-4.6**
					Cent_1	.22	2.3*
					Temp_4	.18	2.2*

Note. AR1 = coefficient for first order autoregressive term. R_{adj}^2 = approximate model fit. *d* = Durbin-Watson statistic. MF = musical feature variable. β = standardized coefficient. *t* = *t* statistic. Loud = loudness. Temp = tempo. Cent = spectral centroid. Number of seconds of lag indicated after each musical feature variable. Significance levels: ** $p < .001$, * $p < .05$

Table 5.

Pearson coefficients for cross-correlations with significant p -values between musical feature variables of loudness and timbral parameters (spectral centroid and texture) for the gradual gestural types.

Excerpt	Model	MF	Cent	Cent_1	Cent_2	Cent_3	Cent_4	Text	Text_1	Text_2	Text_3	Text_4
E01	Mus	Loud	.36**	.50**	.22**							
		Loud_2	-		.48**	.52**	.14*					
			.31**									
		Loud 4										
E04	Mus	Loud 2		.42**	.53**	.46**	.23**		.18*	.26*	.21*	
		Non Loud	.53**	.46**	.23**			.26**	.21*			
		Loud 2		.42**	.53**	.46**	.23**		.18*	.26*	.21*	
		Loud 4				.42**	.53**				.18*	.26*

Note. Mus = musician model. Non = nonmusicians model. MF = musical feature variables. Loud = loudness. Cent = spectral centroid. Text = Texture. Number of seconds of lag indicated after each musical feature variable. Significance levels: ** $p < .001$, * $p < .05$. Blank cells indicate non-significant p -values

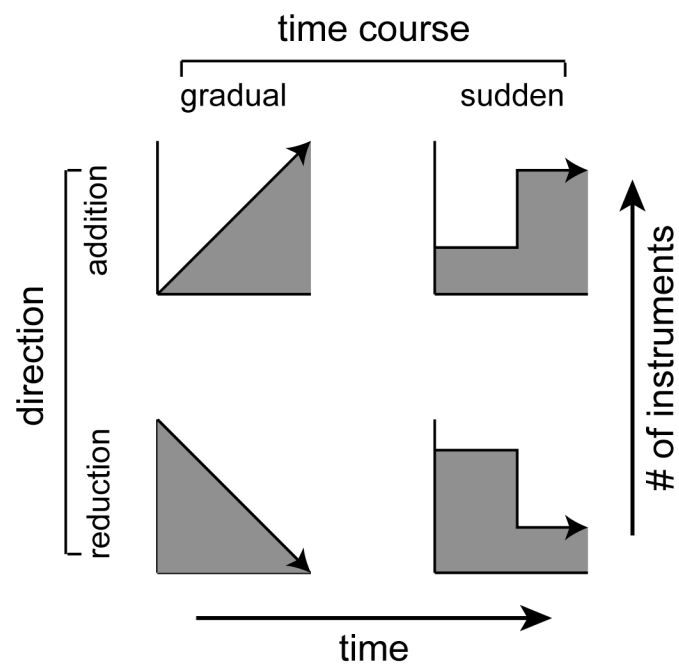
Figures

Figure 1. Four types of orchestral gestures.

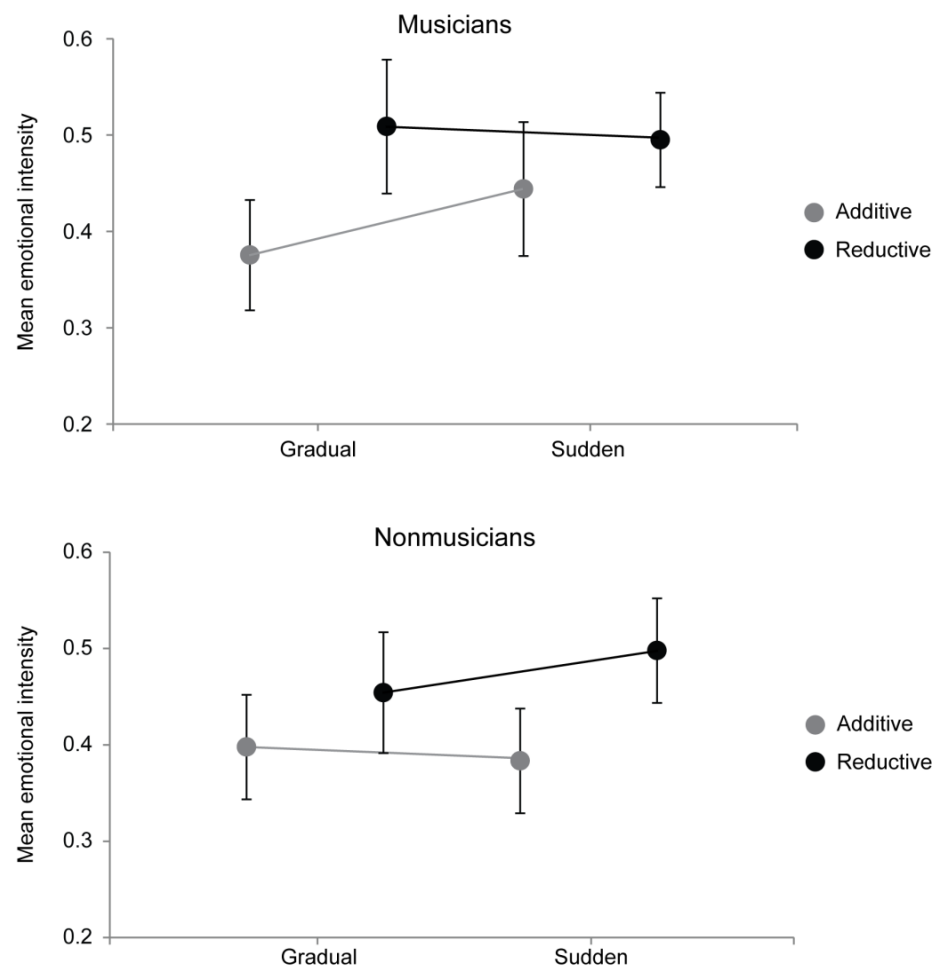


Figure 2. Grand mean emotional intensity ratings with 95% confidence intervals for musicians and nonmusicians.

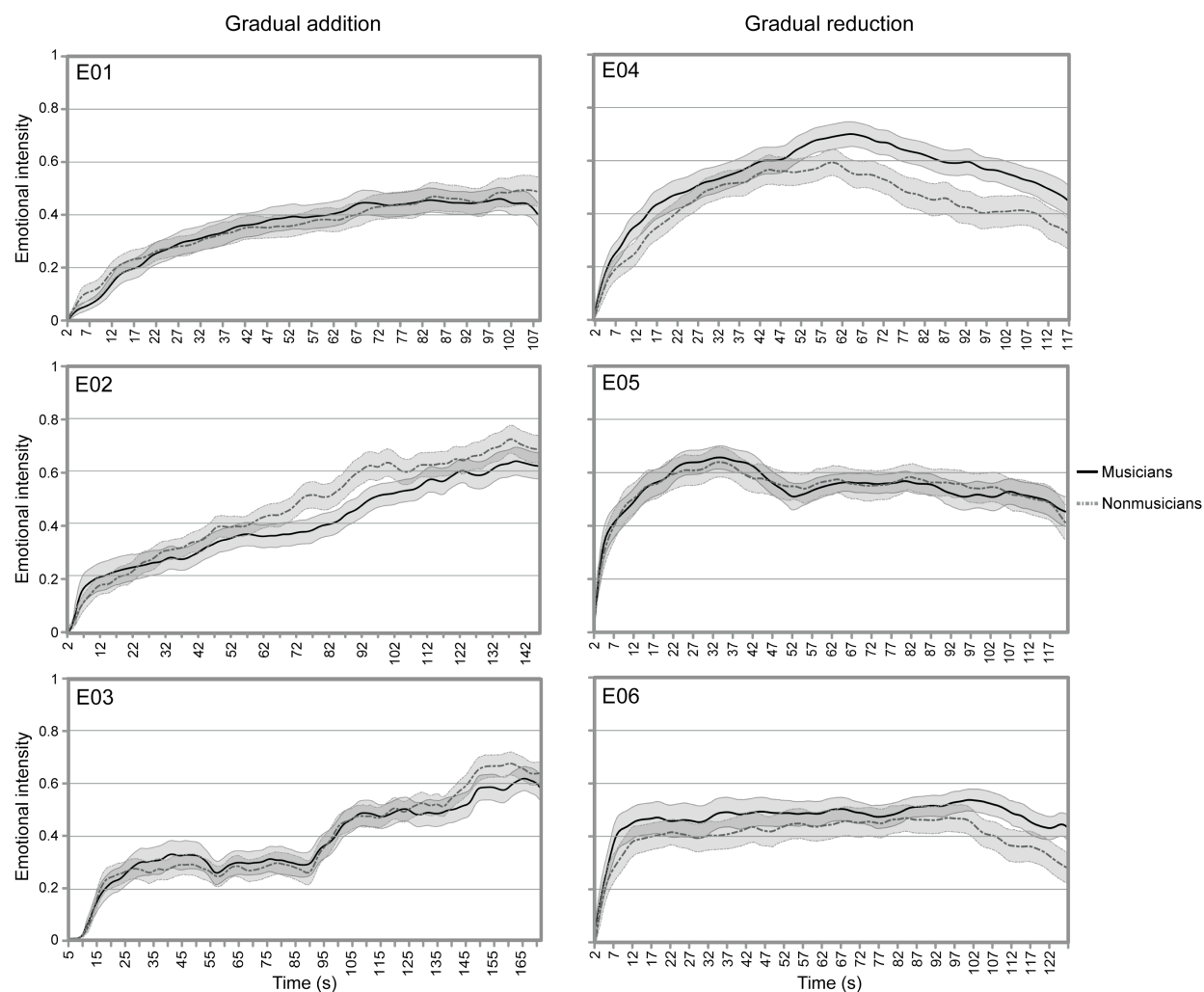


Figure 3. Average emotional intensity ratings for musicians (black line) and nonmusicians (dotted line) for gradual addition and gradual reduction excerpts. Shaded regions indicate ± 1 standard error of the mean for musicians and nonmusicians, respectively. Overlapping areas are represented by darker shading

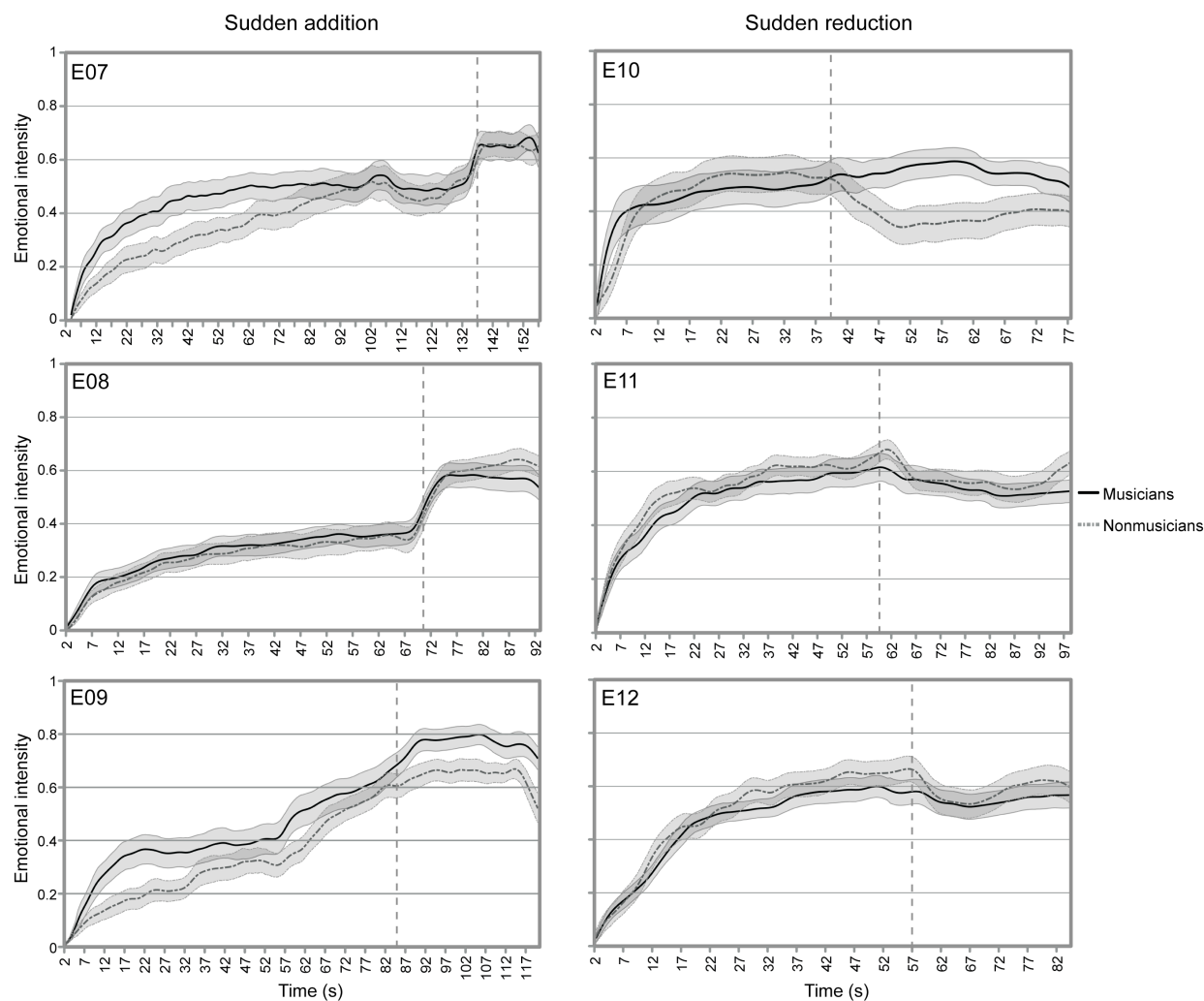


Figure 4. Average emotional intensity ratings for musicians (black line) and nonmusicians (dotted line) for sudden addition and sudden reduction excerpts. Shaded regions indicate ± 1 standard error of the mean for musicians and nonmusicians, respectively. Overlapping areas are represented by darker shading. Vertical dotted line represents the moment of sudden textual change.

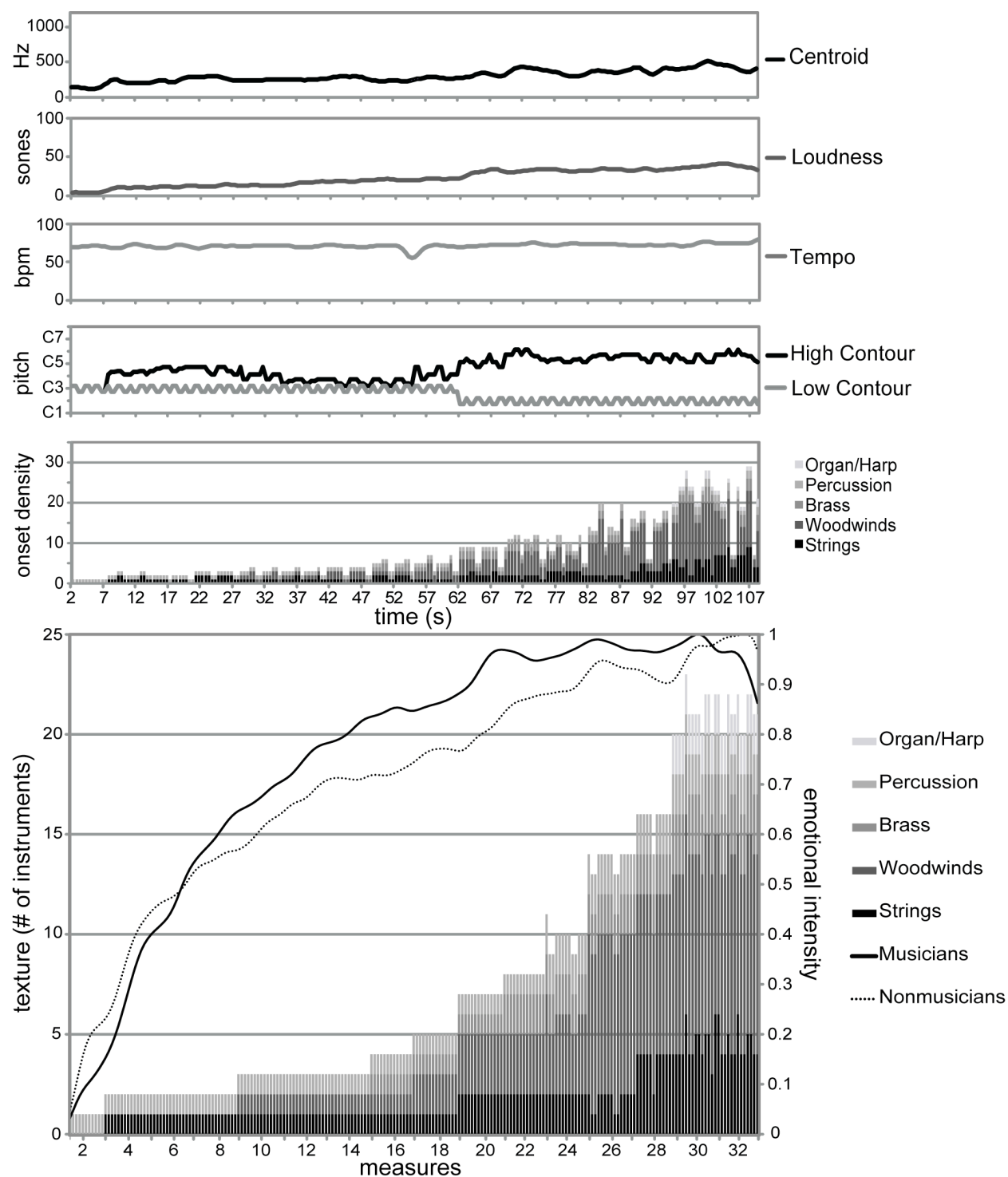


Figure 5. Visualization of Mahler, Symphony No. 1, III, mm. 1-32, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings.

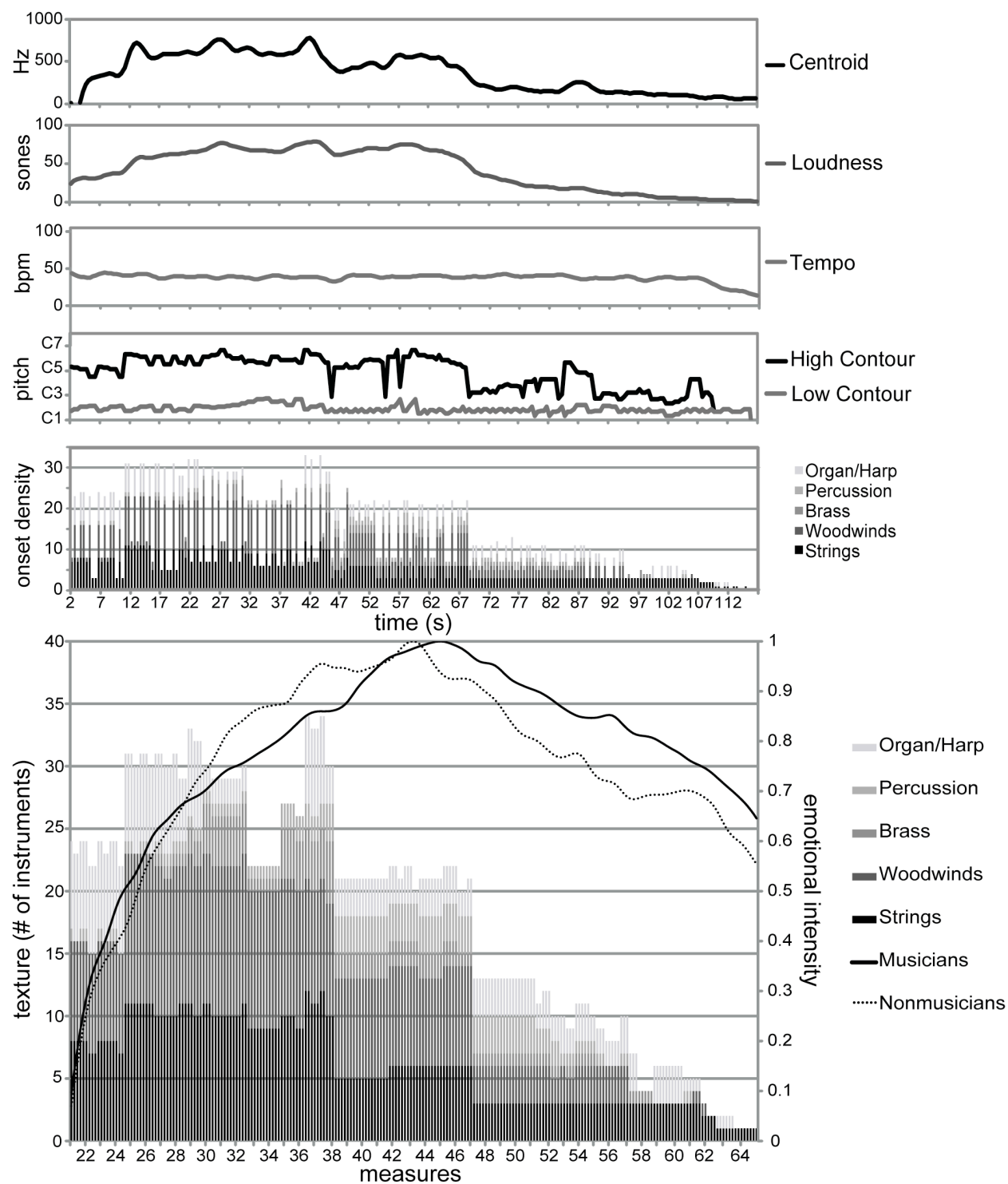


Figure 6. Visualization of Mussorgsky/Ravel, *Pictures at an Exhibition*, “Bydlo,” mm. 21-64, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings.

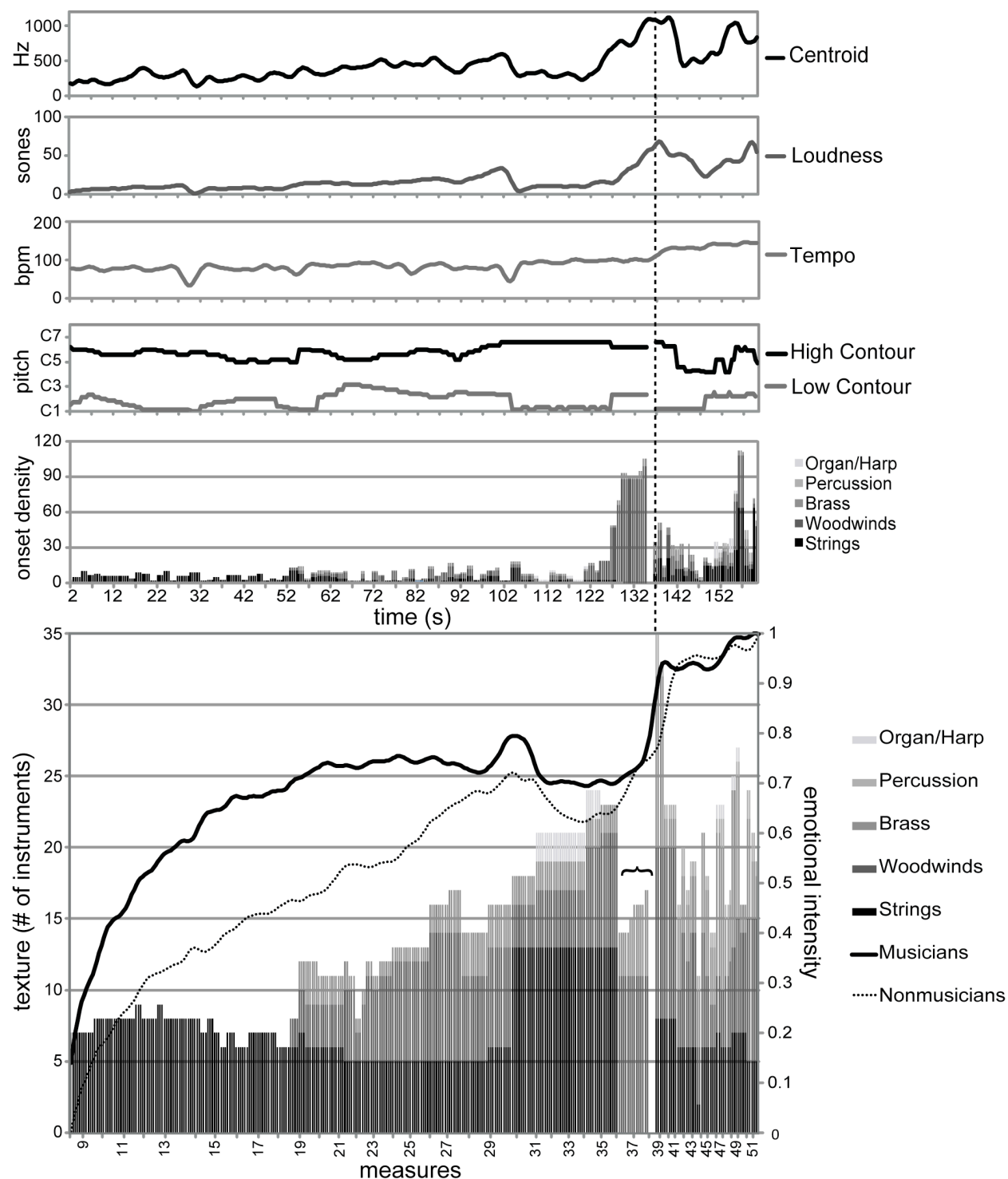


Figure 7. Visualization of Vaughan Williams, *London Symphony*, I, mm. 8-53, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings. The bracket indicates the anticipatory signal and the vertical dotted line indicates the moment of sudden textural change.

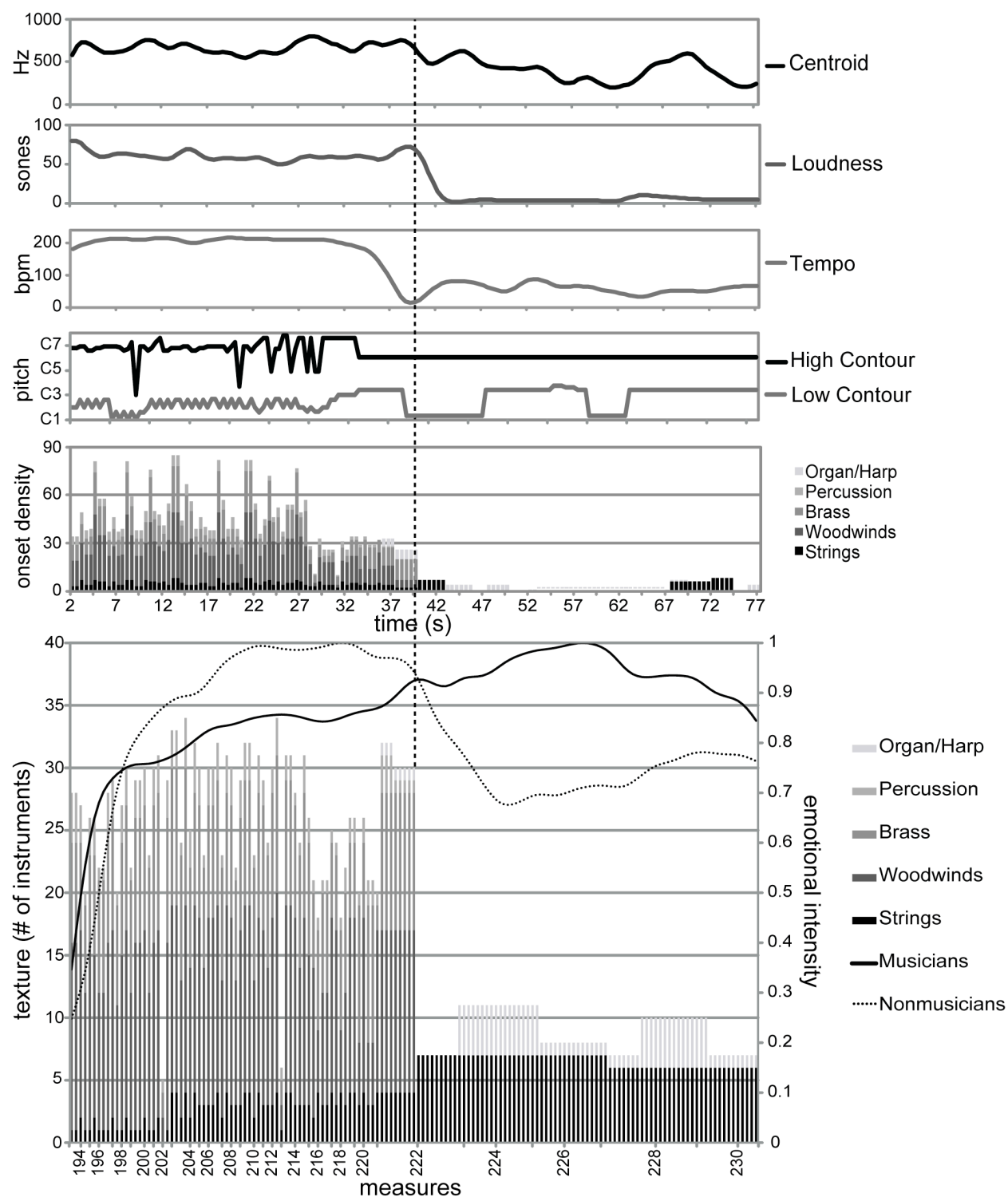


Figure 8. Visualization of Holst, *The Planets*, "Uranus," mm. 193-236, with spectral centroid, loudness, tempo, pitch range, onset density, instrumental texture, and emotional intensity ratings. The vertical dotted line indicates the moment of sudden textural change.