Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal

Montreal, Canada

2 - 7 June 2013

Musical Acoustics

Session 3pMU: Perception and Orchestration Practice

3pMU2. Acoustic and musical features of emotional response to orchestral gestures

Meghan Goodchild*

***Corresponding author's address: McGill University, Montreal, H2T 2G1, QC, Canada, meghan.goodchild@mail.mcgill.ca**

Recent empirical research indicates the impact of prominent changes in instrumentation on the listening experience: several studies suggest that timbral changes evoke music-induced emotions. However, orchestration remains an underdeveloped area of music theory. A model of orchestral gestures defined by changes in instrumentation in terms of time course (gradual or sudden) and direction (addition or reduction) is presented. An exploratory behavioural study that tested the perceptual relevance of orchestral gestures on listeners' continuous ratings of emotional intensity was conducted. We demonstrate a new type of visualization that illustrates the relative textural density of each instrument family over time combined with other time-varying parameters extracted from the signal (loudness, spectral centroid, tempo, and roughness) and calculated from the score (instrumental texture and onset density). In addition to quantitative and qualitative comparison of similar orchestral gestures across pieces, this method is used to study interaction of specific instrumentation changes and other musical parameters. Through discussion of the visualizations, the author highlights relationships between the perceptual and musical/acoustical dimensions and quantifies elements of the temporality of these experiences.

Published by the Acoustical Society of America through the American Institute of Physics

Received 22 Jan 2013; published 2 Jun 2013

INTRODUCTION

Background

Recent empirical studies have investigated musical and acoustical correlates of emotional responses to music; several studies indicate that timbral changes induce strong emotional responses in listeners. Guhn et al. (2007) reported that listeners experienced chills (a sudden arousal accompanied by goose bumps, shivers, or tingles down the spine) to musical passages marked by the alternation or contrast of a solo instrument and the orchestra. Similarly, Panksepp (1995) found that participants experienced chills when a soloist emerged from a dense orchestral texture. Tears, chills, and an increase in heart rate were associated with sudden dynamic or textural changes by Sloboda (1991). This body of work suggests texture and instrumentation profoundly impact the listening experience. Orchestration and timbre have not been theorized in music research to the same extent as other parameters such as harmony, melody, rhythm and formal processes (Slawson, 1985). In general, orchestration manuals provide prescriptions for instrumental combinations and short excerpts to be emulated. Timbre research has primarily focused on acoustical attributes that best correlate with perceptual ratings of similarity and blend; however, these studies most often use solo or pairs of instrument sounds producing a single event (McAdams, 1993; Sandell, 1995; McAdams and Giordano, 2009). Further research is still needed in the pursuit of generalized principles, a clear taxonomy of large-scale orchestral techniques, and a conceptual framework related to their function and perceptual effect.

Aims

This interdisciplinary study benefits from a reciprocal exchange between the fields of music theory and experimental psychology, specifically through perceptual testing of music-theoretical hypotheses and through the use of analytical insights to explain emotional responses to music. The first aim is to model one aspect of the temporal dynamics of the listening experience by researching the musical features in orchestral music that elicit emotional responses. The second aim is to contribute to the development of a theory of orchestral gestures through music-theoretical analyses and findings from the perceptual experiment. Considering emotional response as an indicator of perceptual salience, the goal of the exploratory behavioral experiment was to map how listeners experience the orchestral gestures over time. These data were used to isolate response patterns and investigate their connection to texture, tempo, dynamics and other musical parameters.

METHODS

Model and Stimuli

As a starting point for inquiry, it was hypothesized that musical passages in which timbre changes in a coordinated manner create large-scale *orchestral gestures*. This type of passage is grouped perceptually into a coherent musical unit as an auditory image, which is a psychological representation of a sound entity. Due to the combination of timbre, dynamics, tempo and other musical dimensions, these orchestral gestures often give rise to strong emotional experiences. Four categories were defined by changes in instrumentation and are organized into a two-by-two design based on time course (either gradual or sudden changes) and direction (either additive or reductive changes), shown in Figure 1. Each gesture either adds or removes instruments, gradually or suddenly, as a coordinated change over time.

Twelve musical excerpts were chosen to fit within the four categories (three excerpts per category). The excerpts were categorized by only one overarching orchestral gesture (as described above). Drawn from the orchestral repertoire of the late nineteenth and early twentieth century (e.g., works by Mahler, Bruckner, Debussy, Strauss), the excerpts were 1-3 minutes in duration in order to provide a listening context. Well-known excerpts were avoided to minimize the chance of prior associations with the music.

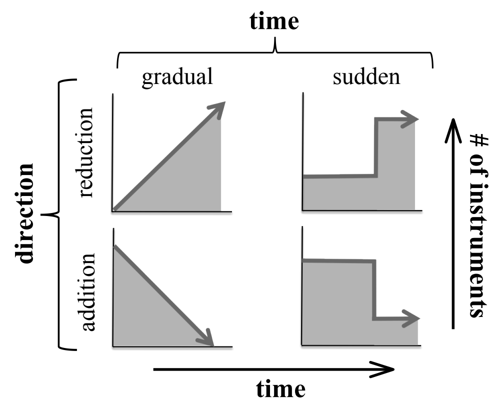

FIGURE 1. Model of Orchestral Gestures.

Hypotheses

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 his regression study, he reported that approximately 60% of the variation in average arousal ratings could be explained by changes in loudness and tempo, but there were inconsistent predictions for the role of texture and timbre.

Gabrielsson and Lindström (2010) reported that a predictable relationship exists between parameters of loudness and timbre on the one hand and emotional intensity on the other; in their study, listeners identified that an increase in loudness and upper harmonics indicates higher emotional intensity.

Conversely, Panksepp (1995) proposed that a timbral reduction (e.g., a solo instrument emerging out of the orchestra) produces a heightened emotional response because the textural change mimics an evolutionary feeling of social loss.

Based on these results, it was predicted that listeners would respond to the orchestral gestures, but that there would not be a direct correspondence between emotional intensity and the shape of the gesture in terms of instrumental texture.

Materials and Procedure

In group-listening sessions, 45 participants (22 musicians and 23 non-musicians) listened to the excerpts in a randomized order and continuously moved a slider to indicate the buildup and decay of the intensity of their emotional responses. Slider responses were measured simultaneously using a MIDI-based acquisition hardware. The slider was equipped with an elastic band to provide force-feedback, which we have found to be an intuitive method of rating that obviates the need for visual attention to the rating device. Participants also completed questionnaires outlining their specific subjective experiences (chills, tears, and other reactions) and their familiarity and preference for each excerpt.

Data Analysis

The emotional intensity slider data were resampled at 2 Hz and normalized to a range between 0 and 1 for each subject across excerpts. Due to the interest in the between-excerpts variation, normalization reduces the inter-participant variability due to different ranges of rating. The slider ratings were averaged for musician and non-musician groups.

Four performance-based features (loudness, spectral centroid, roughness, and tempo) and two score-based features (instrumental texture and onset density) were analyzed based on their importance in similar research (Schubert 2004) and in descriptions of orchestral gestures in treatises (Adler 2002).

Loudness, spectral centroid and roughness were coded with PsySound 3 (Cabrera et al., 2007) using the 16-bit, 44.1 kHz audio WAV files obtained from commercial recordings. For loudness, an auditory filter model (Moore and

Glasberg, 1997) measured the intensity of steady state sounds in sones. The spectral centroid, associated with timbral brightness, was extracted in Hz. A psychoacoustical roughness measure (aspers) was also used (Daniel and Weber, 1997).

Tempo, instrumental texture and onset density were manually coded. Using Sonic Visualiser (Cannam et al., 2010), the inter-onset interval (IOI) was marked by the author, verified by a musician colleague, and then converted into the tempo time series in beats per minute (bpm). Automated methods of beat extraction were unsuccessful due to the textural density of the orchestral sound files. Following Schubert's (2004) procedure, texture was objectively coded as the number of independent voices (or parts) at each beat; however, this parameter was expanded to include information about the contributions of each instrumental family: strings, woodwinds, brass, percussion, and organ/harp. Onset density was measured as the number of onsets in each instrumental family per beat.

First, the data were resampled at 2 Hz. Next, a low-pass Butterworth digital filter with a cutoff frequency at 0.2 Hz was applied to smooth the data of the loudness, spectral centroid, roughness, and tempo parameters. The data for instrumental texture and onset density were not smoothed due to the discrete nature of these values. Loudness, spectral centroid and roughness data were normalized between 0 and 1 for each excerpt due to the expansive range of values and in order to inspect the changes within an excerpt and to compare across excerpts.

RESULTS

Retrospective Ratings

Based on the feedback provided in the questionnaires, participants reported that they experienced strong emotional responses to the orchestral excerpts. Friedman non-parametric tests compared ratings of familiarity and preference as recorded by participants on a seven-point Likert scale and the number of reported chills for the four experimental conditions. There were no significant differences found for the ratings of preference among the categories [$\chi^2(3) = 6.10$, $p = 0.11$]. However, there was a significant difference found for the ratings of familiarity depending on condition [$\chi^2(3) = 22.04$, $p < 0.001$]. The post-hoc Wilcoxon signed rank test with Bonferroni adjustment applied (significance level set at $p < 0.0083$) revealed that the gradual addition category had a significantly higher number of chills reported compared to the gradual reduction ($Z = –3.74$, $p < 0.001$), sudden addition ($Z = –3.67$, $p < 0.001$) and sudden reduction ($Z = –3.54$, $p < 0.001$) categories. There were no significant differences among the latter three categories.

There was a significant difference in the number of chills depending on condition [$\chi^2(3) = 22.05$, $p < 0.001$]. The post-hoc test revealed that the sudden addition category had significantly higher number of chills reported compared to the gradual reduction ($Z = –3.22$, $p = 0.001$) and sudden reduction ($Z = –2.76$, $p = 0.006$) categories; the gradual additive category almost reached significance ($Z = –2.54$, $p = 0.011$). Again, there were no significant differences among the latter three categories.

Emotional Intensity Ratings and Visualizations

Each of the orchestral gesture categories featured distinct emotional response profiles. For the gradual addition category, participants' ratings steadily increased until the end of the excerpt along with musical and acoustical features, particularly instrumental texture and loudness. The reductive categories, both gradual and sudden, exhibited a lingering effect of persistent high emotional intensity despite decreases in the parameters representing acoustical and musical properties.

The emotional intensity ratings and the musical feature variables were graphed together for visual comparison. Figure 2 contains a visualization for the sudden-addition category: an excerpt from Vaughan Williams's *A London Symphony*, 1st movement, measures 8-37. The texture parameter is displayed as a stacked bar graph with the instrument family indicated by colour. The musicians' and non-musicians' mean emotional intensity ratings are superimposed and the remaining musical feature variables (onset density, tempo, loudness, spectral centroid, and roughness) are stacked above for ease of reference.

The visualization indicates that none of the variables directly correspond to the slider profiles. The emotional intensity ratings resemble an arch shape until 130 seconds. In contrast, the texture variable builds linearly in richness, while the onset density, loudness and spectral centroid remain quite low.

Two phrase endings are quite dramatic, shown by the dips in tempo, spectral centroid, and loudness at approximately 32 seconds and 102 seconds; the latter instance corresponds to a local peak in emotional intensity for

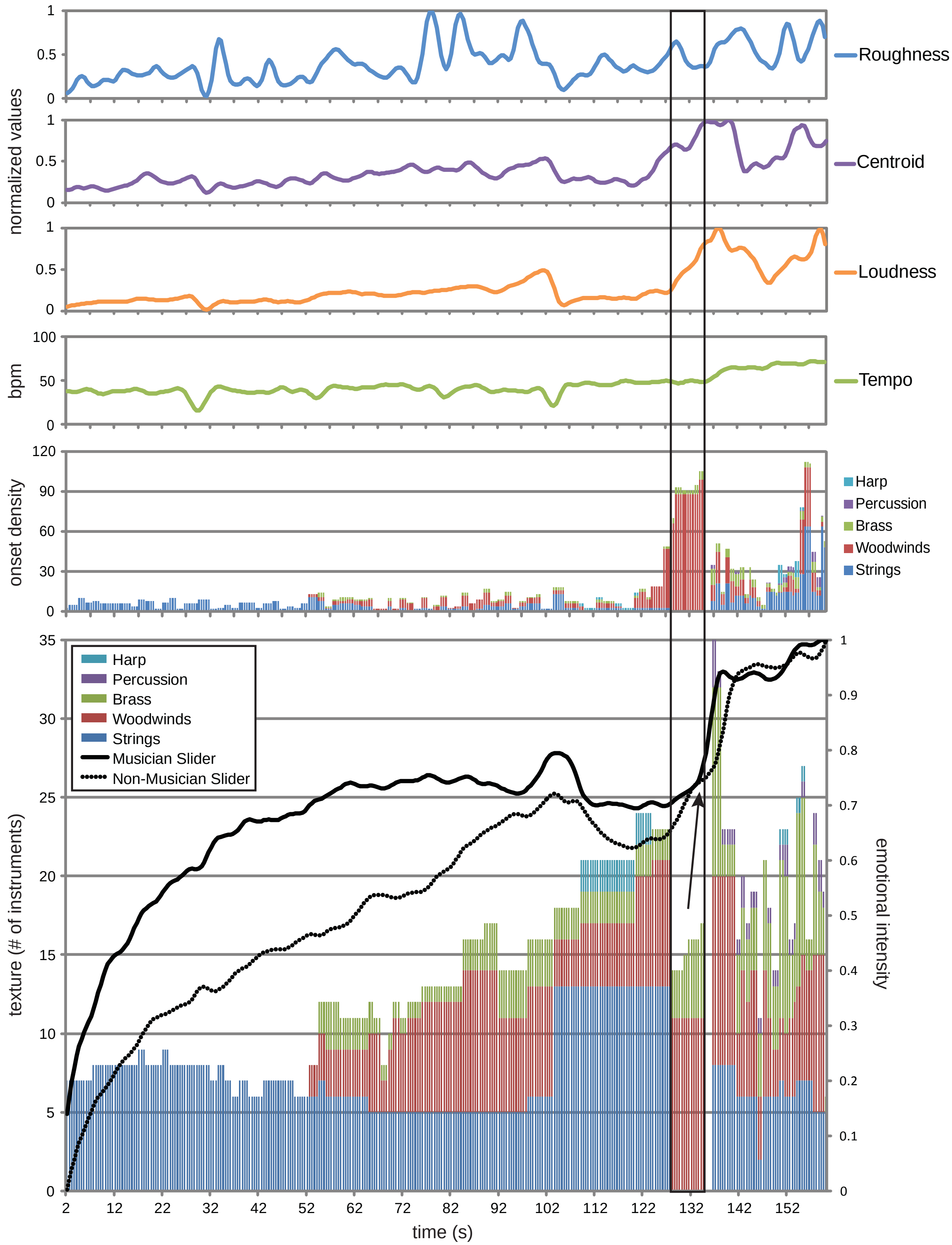

FIGURE 2. Vaughan Williams's *A London Symphony*, 1st movement, mm. 8-37.

both musicians and non-musicians. After this point, the musicians' and non-musicians' ratings dip slightly, while the texture parameter thickens and the spectral centroid and loudness parameters plummet.

Beginning at about 125 seconds (shown with a box), the strings suddenly disappear from the texture, but the woodwinds and brass expand texturally with a sharp increase in loudness, spectral centroid, and onset density. This "anticipatory signal" prepares the listener for the sudden blast. Responding to this indication, the musicians increase their slider ratings in advance of the sudden addition (shown with an arrow). For the non-musicians, however, a similar spike occurs after the blast, a delay of several seconds. Interestingly, the peak emotional intensity values do not occur until the very end of the excerpt and not directly at the sudden addition. This may be related to the way the texture, onset density, loudness and spectral centroid parameters dip intermittently, while the tempo continues to climb to end of the excerpt.

CONCLUSION AND DISCUSSION

This exploratory study demonstrates the importance of researching orchestral gestures and the effect of timbral changes on the listening experience. Participants reported strong emotional responses to the orchestral gestures under investigation, as evidenced by their retrospective ratings of chills experienced and their continuous ratings of emotional intensity. Although there was a significant difference in the number of chills experienced for the sudden addition category, the lack of differences in preference ratings among the categories suggests that the four orchestral gestures are perceptually relevant.

The significant difference in familiarity ratings between the gradual addition category and all other categories appears to be modulated by high ratings for the excerpt from Symphony 1, movement 3 by Mahler. This reflects recognition of the "Frère Jacques" folk melody in the minor mode, but not necessarily an identification of the piece itself. When the familiarity ratings for this piece are removed from the analysis, there are no significant differences found among the categories [$\chi^2(3) = 2.00$, $p = 0.57$].

The visualizations reveal that there is not a direct one-to-one correspondence between the emotional intensity ratings and acoustical/musical features; therefore, a linear combination of the parameters cannot account for the majority of the variance of the emotional response ratings. The direction and magnitude of the acoustical and musical features vary throughout the excerpt and interact with the musical context.

There were unexpected differences between the ratings of the musicians and non-musicians for the sudden categories (additive and reductive). Perhaps due to their exposure to orchestral music, the musicians recognize compositional signals and anticipate the sudden changes. As a result, the musicians' emotional responses are often heightened before the onset of the dramatic event such as a sudden addition. Since their familiarity ratings were low, this suggests that the musicians anticipated musical events based on their knowledge of the repertoire and not specific memories of the excerpt. The musicians may also respond to harmonic and structural cues in the music that have not been coded in this study. The non-musicians' ratings appear to be more directly related to changes in loudness and instrumental texture, trailing behind changes in these features.

The visualizations are a crucial analytical tool for studying orchestration, since they allow for both quantitative and qualitative comparison of excerpts within and across gestural categories. The ability to view the timbral evolution by instrumental family facilitates the analyses of orchestration changes and their interaction with other acoustical and musical elements. In this way, traditional score-based music-theoretical analyses could be combined with and enhanced by acoustical information, including spectral properties of the signal.

In future work, other musical-feature overlays, such as phrase structure and harmonic tension, will be included in order to better understand the interaction between acoustical and musical parameters. An investigation into the role of performance features, such as timing and dynamics, is also needed. Using the visualizations presented in this study to compare specific recordings would be particularly useful in this regard.

ACKNOWLEDGMENTS

This work was supported by a SSHRC Doctoral Fellowship and a CIRMMT Student Award to Meghan Goodchild, a SSHRC Standard Grant to Stephen McAdams and Jonathan Wild, and Stephen McAdams' Canada Research Chair. We would like to thank Bennett K. Smith for programming the experimental interface.

REFERENCES

Adler, S. (2002). *The Study of Orchestration*. New York: W.W. Norton and Company Inc.
Alluri, V., and Toivianinen, P. (2009). "Exploring perceptual and acoustical correlates of polyphonic timbre." *Music Perception* 27(3): 223-241.
Cannam, C., Landone, C., & Sandler, M. (2010). Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files. *Proceedings of the ACM Multimedia 2010 International Conference* (pp. 1467-1468). Firenze, Italy.
Cabrera, D., Ferguson, S., & Schubert, E. (2007). PsySound3: software for acoustical and psychoacoustical analysis of sound recordings*. Proceedings of the 13th International Conference on Auditory Display* (pp. 356-363). Montreal, Canada
Carse, A. (1964). *The History of Orchestration.* New York: Dover Publications Inc.
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.
Guhn, M., Hamm, A., and Zentner, M. (2007). Physiological and Music-Acoustic Correlates of the Chill Response. *Music Perception,* 24(5), 473-483.
Kennan, K. & Grantham, D. (2002). *The Technique of Orchestration.* Upper Saddle River, N.J.: Prentice Hall.
Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception* 13(3), 401-432.
Matthews, P., editor. (2006). *Orchestration: An Anthology of Writings*. New York: Routedge.
McAdams, S. (1993). "Recognition of sound sources and events." In S. McAdams & E. Bigand (Eds.), *Thinking in Sound: The Cognitive Psychology of Human Audition* (pp. 146-198). Oxford: Oxford University Press.
McAdams, S., and Giordano, B. L. (2009). "The perception of musical timbre." In S. Hallam, I. Cross & M. Thaut (Eds.), *Oxford Handbook of Music Psychology* (pp. 72-80). New York: Oxford University Press.
Moore, B., Glasberg, B., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45(4), 224-40.
Panksepp, J. (1995). The emotional sources of chills induced by music. *Music Perception,* 13(2), 171-207.
Poulin-Charronnat, B., Bigand, E., Lalitte, P., Madurell, F., Vieillard, S., and McAdams, S. (2004). "Effects of a change in instrumentation on the recognition of musical materials." *Music Perception* 22: 239-263.
Sandell, G. J. (1995). Roles for spectral centroid and other factors in determining "blended" instrument pairings in orchestration. *Music Perception,* 13(2), 209-246.
Schubert, E. (1999). Measurement and time series analysis of emotion in music. PhD Dissertation. University of New South Wales.
Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, 21(4), 561-585.
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, 110-120.