

# The Emotional Characteristics of Bowed String Instruments with Different Pitch and Dynamics

CHUCK-JEE CHAU<sup>1</sup>, SAMUEL J. M. GILBERT<sup>2</sup>, RONALD MO<sup>1</sup>, AND ANDREW HORNER<sup>1</sup>,<sup>1</sup> *AES Member*  
(chuckjee@cse.ust.hk) (s.j.m.gilbert@ncl.ac.uk) (ronmo@cse.ust.hk) (horner@cs.ust.hk)

<sup>1</sup>The Hong Kong University of Science and Technology, Hong Kong

<sup>2</sup>Newcastle University, Newcastle, UK

Previous research has shown that different musical instrument sounds have strong emotional characteristics. This paper investigates how emotional characteristics vary with pitch and dynamics within the bowed string instrument family. We conducted listening tests to compare the effects of pitch and dynamics on the violin, viola, cello, and double bass. Listeners compared the sounds pairwise over 10 emotional categories. Results showed that the emotional characteristics Happy, Heroic, Romantic, Comic, and Calm generally increased with pitch but decreased at the highest pitches. Angry and Sad generally decreased with pitch. Scary was strong in the extreme low and high registers, while Shy and Mysterious were unaffected by pitch. For dynamics, the results showed that Heroic, Comic, and Angry were stronger for loud notes, while Romantic, Calm, Shy, Sad, and the high register for Happy were stronger for soft notes. Scary and Mysterious were unaffected by dynamics. The results also showed significant differences between different bowed string instruments on notes of the same pitch and dynamic level. These results help quantify our understanding of the relative emotional characteristics of the strings. They provide audio engineers and musicians with suggestions for emphasizing emotional characteristics of the bowed strings in sound recordings and performances.

## 0 INTRODUCTION

Music emotion has been a hot topic in recent years with many studies on music emotion recognition systems [1–13] and other applications [14–22]. One strand of music emotion research has focused on the various connections between timbre and music emotion.

In particular, a number of recent studies have found that different musical instruments have strong emotional characteristics [23–30]. For example, among sustained instruments the violin was found to be stronger in the characteristics Happy, Heroic, and Comic than the horn [24].

These studies have focused on a single common pitch, usually a note just above middle C, so that as many treble and bass clef instruments can be compared against one another as possible. Such a comparison provides a practical and useful point of reference when comparing the spectral profiles and emotional characteristics of the instruments. But it is also valuable to see how the instruments vary in their spectral and emotional characteristics with different pitches and dynamic levels. Several studies have shown that pitch and dynamic levels can change perceived aspects of the sound in musical excerpts [31–33], speech [34, 35], and isolated musical instrument tones [23].

Most relevant to the current study, we recently studied how the piano's emotional characteristics changed with pitch and dynamics from C1 to C8 over *piano*, *mezzo*, and *forte* dynamic levels [36, 37]. In that study we found that the emotional characteristics Happy, Romantic, Comic, Calm, Mysterious, and Shy generally increased in pitch in an arching shape that decreased at the highest pitches. The characteristics Heroic, Angry, and Sad basically decreased in pitch. Comic was strongest in the mid-register. Scary had a U-shape that was strongest in the extreme low and high registers. In terms of dynamics on the piano, the characteristics Heroic, Comic, Angry, and Scary were stronger for loud notes, while Romantic, Calm, Mysterious, Shy, and Sad were stronger for soft notes. Surprisingly, Happy was not affected by dynamics.

In a similar way to that previous study, this paper considers how the emotional characteristics of the bowed string instruments differ with pitch and dynamics. While various studies have investigated the timbre of the violin or bowed strings, considering factors such as perceived pitch [38], vibrato [38], and openness [39], to our knowledge none has considered the emotional characteristics of the bowed strings for different pitches and dynamic levels. We will compare the pitches C1–C7 at *piano* and *forte* dynamic levels.

We are curious to see how the bowed string results compare to those of the piano. Will the emotional characteristics follow similar patterns or be completely different? We will determine which emotional characteristics increase with pitch and which decrease. Are any strongest in the mid-register? Are any relatively unaffected by pitch? Are the curves smoothly varying or are there some with strong isolated peaks?

We expect some instrument-dependent differences between the piano and bowed strings, but there might be some strong similarities as well especially for dynamics. After all, it seems like Calm would generally be soft in dynamics regardless of instrument, but there could be surprises for other emotional characteristics. In any case, we will determine which emotional characteristics are strong for soft notes, and which for loud notes. Perhaps some emotional characteristics will be relatively unaffected by dynamics. And perhaps there are some emotional characteristics with differences in dynamics just for isolated parts of the pitch range (e.g., high notes but not low notes).

We are also especially curious to see the differences in the individual bowed string instruments (e.g., violin and viola) at the same pitch and dynamic levels. The bowed strings are mostly resized versions of the violin making them more uniform in timbre when compared to the brass and wind instrument families that have more distinct differences. Nevertheless, there are distinct timbral and perceptual differences between the violin, viola, and cello at for example C4 or C5, and these differences are bound to show up as differences in their emotional characteristics as well.

Overall, this study will help quantify the emotional effects of pitch and dynamics in the bowed strings. The results will provide possible suggestions for musicians in orchestration, performers in blending and balancing instruments, and audio engineers in recording and mixing bowed string instruments.

## 1 EXPERIMENT METHODOLOGY

We conducted listening tests to compare the effects of pitch and dynamics on the emotional characteristics of individual bowed string instrument sounds. We tested the violin, viola, cello, and double bass at three or four different pitches, and at both *forte* (loud) and *piano* (soft) dynamic levels. We compared the sounds pairwise over 10 emotional categories (Happy, Heroic, Romantic, Comic, Calm, Mysterious, Shy, Angry, Scary, and Sad) to determine the effects of pitch and dynamics.

For this investigation we used short sounds isolated from musical context in order to isolate the effects of pitch and dynamics. We also correlated the emotion ratings with several pitch, dynamic, and spectral features in order to better understand how their emotional characteristics are related to timbre in the bowed strings.

### 1.1 Stimuli

The experiment used sounds from the four main instruments in the Western bowed string family: violin (Vn),

viola (Va), cello (Vc), and double bass (Cb). The sounds were obtained from the Prosonus sample library [40]. The sounds presented were approximately 0.9 s in length. For each comparison, the first sound was played, followed by 0.2 s of silence, and then the second sound. Thus the total for one comparison was 2 s.

We chose this duration as it was long enough to allow listeners to hear a representative portion of the sound, but not too long or else the overall length of the listening test would become too long. In our previous study [29] we found that emotional characteristics were clear for very short sounds of 0.25 s duration, but only for mid-register pitches. To allow listeners to hear the details of the attack and early decay well for the lowest pitches, we could not use a duration too short, and 1 s tones were the best compromise in length. We felt that listeners would judge 1 s or 2 s sounds (or longer) with similar results. Previous studies of Peretz et al. [14], Krumhansl [41], and Filipic et al. [42] also confirmed that listeners could already discriminate emotions or even identify the artist by using very short musical excerpts as short as 0.25 s. The sounds were so short that factors such as rhythm, melody, and chord progression were largely excluded.

The pitches for each instrument were as follows:

- Vn: C4, C5, C6, C7
- Va: C3, C4, C5
- Vc: C2, C3, C4, C5
- Cb: C1, C2, C3

The sounds were all Cs of different octaves so as to avoid other musical intervals influencing the emotional characteristics of the sounds. Each note also had two dynamic levels, corresponding to *forte* (*f*) and *piano* (*p*)—loud and soft. The total number of sounds was 28 (14 notes  $\times$  2 dynamic levels).

The samples at the two dynamic levels were provided by the Prosonus sample library, which was judged by the performers and sound engineers involved in the production. We confirmed that they were consistent and reasonable and did not make further adjustments in the amplitude.

The instrument sounds were analyzed using a phase-vocoder algorithm, where bin frequencies were aligned with harmonics [43]. Temporal equalization was carried out in the frequency domain, identifying attacks and decays by inspection of the time-domain amplitude-vs.-time envelopes. These envelopes were reinterpolated to achieve a standardized attack time of 0.07 s, sustain time of 0.36 s, and decay time of 0.43 s for all sounds. These values were chosen based on the average attack and decay times of the original sounds. As different attack and decay times are known to affect the emotional responses of subjects [23], equalizing avoids this potential factor. The stimuli were resynthesized from the time-varying harmonic data using the standard method of time-varying additive sine wave synthesis (oscillator method) [43] with frequency deviations set to zero. We verified that the resynthesized sounds were representative of the original sounds and free from audible artifacts. We also confirmed that the spectral features were also

Table 1. The 10 chosen emotional categories and related music expression markings commonly used by classical composers.

Emotional Category	Commonly-used Italian musical expression markings
Happy	<i>allegro, gustoso, gioioso, giocoso, contento</i>
Heroic	<i>eroico, grandioso, epico</i>
Romantic	<i>romantico, affetto, affettuoso, passionato</i>
Comic	<i>capriccio, ridicolosamente, spiritoso, comico, buffo</i>
Calm	<i>calmato, tranquillo, pacato, placabile, sereno</i>
Mysterious	<i>misterioso, misteriosamente</i>
Shy	<i>timido, riservato, timoroso</i>
Angry	<i>adirato, stizzito, furioso, feroce, irato</i>
Scary	<i>sinistro, terribile, allarmante, feroce, furioso</i>
Sad	<i>dolore, lacrimoso, lagrimoso, mesto, triste</i>

retained intact using the phase-vocoder algorithm. All sounds were recorded and sampled at 44100 Hz with 16-bit resolution and played back using the D/A converter with 24-bit resolution at the original sampling rate.

## 1.2 Emotional Categories

The subjects compared the stimuli in terms of ten emotional categories: Happy, Heroic, Romantic, Comic, Calm, Mysterious, Shy, Angry, Scary, and Sad.

We selected the same categories we have used in previous studies [37, 44–48]. Composers often use these terms in tempo and expression markings in their scores. We chose to use simple English emotional categories so that they would be familiar and self-apparent to subjects rather than Italian music expression markings traditionally used by classical composers to specify the character of the music. The chosen emotional categories and related Italian expression markings [49–52] are listed in Table 1. We include a group of 10 emotional categories that are similar to the 8 adjective groups of Hevner [53]. We chose the four main categories that are commonly used to represent the four quadrants of the Valence–Arousal plane [54] (Happy, Sad, Angry, and Calm). The other categories are distinctly different from these four, but frequently occur in music expression markings. Other researchers have also used some of these (or related) emotional categories [55–57]. These emotional categories also provide easy comparison with the results in [24–30, 36, 37, 47, 58].

The 10 emotional categories were considered separately. For example, the stimuli were rated for how Happy they were relative to one another within each category. Ratings were not compared across categories.

## 1.3 Test Procedure

Twenty-three subjects were hired to take the listening test. All subjects were fluent in English. They were all undergraduate students at the Hong Kong University of Science and Technology where all courses are taught in

Table 2. The dictionary definitions of the emotional categories used in our experiment.

Emotional Category	Definition [59]
Happy	Glad, pleased
Heroic	Exhibiting or marked by courage and daring
Romantic	Making someone think of love
Comic	Causing laughter or amusement
Calm	A quiet and peaceful state or condition
Mysterious	Exciting wonder, curiosity, or surprise while baffling efforts to comprehend or identify
Shy	Disposed to avoid a person or thing
Angry	Having a strong feeling of being upset or annoyed
Scary	Causing fright
Sad	Affected with or expressive of grief or unhappiness

English. All stated that they had no known hearing impairments. Subjects did not have highly-trained ears (e.g., recording engineers, professional musicians, or music conservatory students) but were average attentive listeners.

The subjects were seated in a quiet room with 39 dB SPL background noise level. The noise level was further reduced with headphones. Sound signals were presented through Sony MDR-7506 headphones. We felt that basic-level professional headphones were adequate in representing the bowed string sounds for this test as the sounds were readily distinguishable. A big advantage of the Sony MDR-7506 headphones is their relative comfort in a relatively long listening test such as this one, especially for subjects not used to tight-fitting studio headphones. The volume on all computers were calibrated manually so that the C4 *forte* violin tone sounded at the same moderate loudness level as judged by the authors.

The subjects were provided with an instruction sheet containing definitions of the 10 emotional categories from the *Cambridge Academic Content Dictionary* [59]. The dictionary definitions we used in our experiment are shown in Table 2.

Every subject made pairwise comparisons on a computer among all the 28 combinations of instruments, pitches, and dynamics for each emotional category. During each trial, subjects heard a pair of sounds of different instruments/pitches/dynamics and were prompted to choose the sound that represented the given emotional category more strongly. Each trial was a single paired comparison requiring minimal memory from the subjects. In other words, subjects did not need to remember all of the tones, just the two in each comparison. One big advantage of using paired comparisons of emotional categories is that it allows faster decision-making by the subjects. Paired comparison is also a simple decision and is easier than absolute rating.

Fig. 1 shows a screenshot of the listening test interface. Each combination of sounds was presented once for each emotional category, and the listening test totaled  $\binom{28}{2} \text{ combinations} \times 10 \text{ emotional categories} = 3780$  trials. For each emotional category, the overall trial



Fig. 1. Paired comparison listening test interface.

presentation order was randomized (i.e., all the Happy comparisons were first in a random order, then all the Sad comparisons were second, and so on). However, the emotional categories were presented in order to avoid confusing and fatiguing the subjects. As with any listening test, there can be learning at the beginning and fatigue at the end. For this test, there were 10 random test trials at the start of the test that were not used in calculations to minimize the effects of learning.

Altogether the listening test took about 2–3 hours over several sessions. There were also forced short breaks of 5 minutes after every 30 minutes to help minimize listener fatigue and maintain consistency.

## 2 RESULTS

We ranked the sounds by the number of positive votes received for each emotional category, deriving scale values using the Bradley–Terry–Luce (BTL) statistical model [60, 61]. The BTL values for each emotional category sum to 1. The BTL value given to a sound is the probability that listeners will choose that sound when considering a given emotional category. For example, if all 28 sounds (14 notes  $\times$  2 dynamic levels) were considered equally Happy, the BTL scale values would be  $1/28 \approx 0.0357$ . The corresponding 95% confidence intervals were derived using Bradley’s method [61]. Fig. 2 shows graphs for the BTL scale values and the corresponding 95% confidence intervals for each emotional category and instrument.

At first glance, one notices that the individual instrument lines are similar and together outline an overall trend for the bowed strings for each emotional characteristic. There are some distinctive outliers such as the *forte* cello at C5.

In terms of pitch, several of the emotional categories in Fig. 2 had a similar arching shape, including Happy, Heroic, Romantic, Comic, and Calm, peaking at C5 (or C6). Sad was also arching but peaked at C2. Both Sad and Angry could be characterized as mostly decreasing with pitch. Scary was uniquely U-shaped with peaks at the extreme high and low pitches. Both Shy and Mysterious were flatter with relatively little change in pitch.

In terms of dynamics, the emotional characteristics Heroic, Comic, and Angry were stronger for loud notes, while Romantic, Calm, Shy, and Sad were stronger for soft notes. The characteristics Mysterious, Scary, and most of Happy (except the high register) were relatively unaffected by dynamics.

### 2.1 The Effects of Pitch and Dynamics

The curves for most of the categories in Fig. 2 showed clear trends. For example, Heroic had a strong arch for *forte* and a gentler arch for *piano*.

To more precisely quantify these trends, we wanted to determine whether the effects of pitch and dynamics were significant for the bowed string tones. An ANOVA analysis is the usual way to accomplish this, but ANOVA requires independent variables. For the bowed strings, pitch and dynamics are independent but instrument is not since the violin for example only ranges from C4 to C7 and does not include tones from C1 to C3. Alternatively, we can treat the bowed string family as a single instrument for the purposes of this analysis and select representative tones for each pitch. For example, at C5 we can select the violin as the most representative instrument of the string family, leaving aside the viola and cello. Using this idea, we constructed our “most representative” bowed strings using double bass for C1, cello for C2–C4, and violin for C5–C7 (see Fig. 3). In a way, this make sense since the violin and cello are much more common as solo instruments than the viola and double bass and in this sense most representative of the bowed strings. But, to be fair to the viola and double bass, we also constructed a second “less common” representative consisting of double bass for C1–C2, viola for C3–C5, and violin for C6–C7 (see Fig. 4). The values in Figs. 3 and 4 are simply the extracted values from Fig. 2. The overall shapes in Figs. 3 and 4 are basically similar. We can then run ANOVA on both representatives for the bowed strings and compare the results.

For each emotional category, a two-way ANOVA with replication was performed to test the effects of pitch and dynamics. Since we wanted to consider each listener’s preferences individually, rather than the composite BTL value, ANOVA was performed on the number of times each listener preferred each tone compared to the others category-by-category. So, for example if listener #1 preferred the violin–C5–*forte* to all the other tones for Happy, then listener #1’s violin–C5–*forte* value would be 27 since there were 27 other tones. And, if listener #1 preferred all the other tones to the double bass–C1–*piano*, the corresponding value would be 0. The other tones’ values would be somewhere between 0 and 27.

As a preliminary step to the ANOVA, a Shapiro–Wilk test was performed to check whether the listener preference data were normally distributed. Table 7 in the Appendix shows the result. The degree of freedom for each tone and category was 28, representing the 28 listeners. About 90% of the data were normally distributed. Since the vast majority of the data were normally distributed, we performed a two-way ANOVA. Sphericity was violated with a small epsilon (see



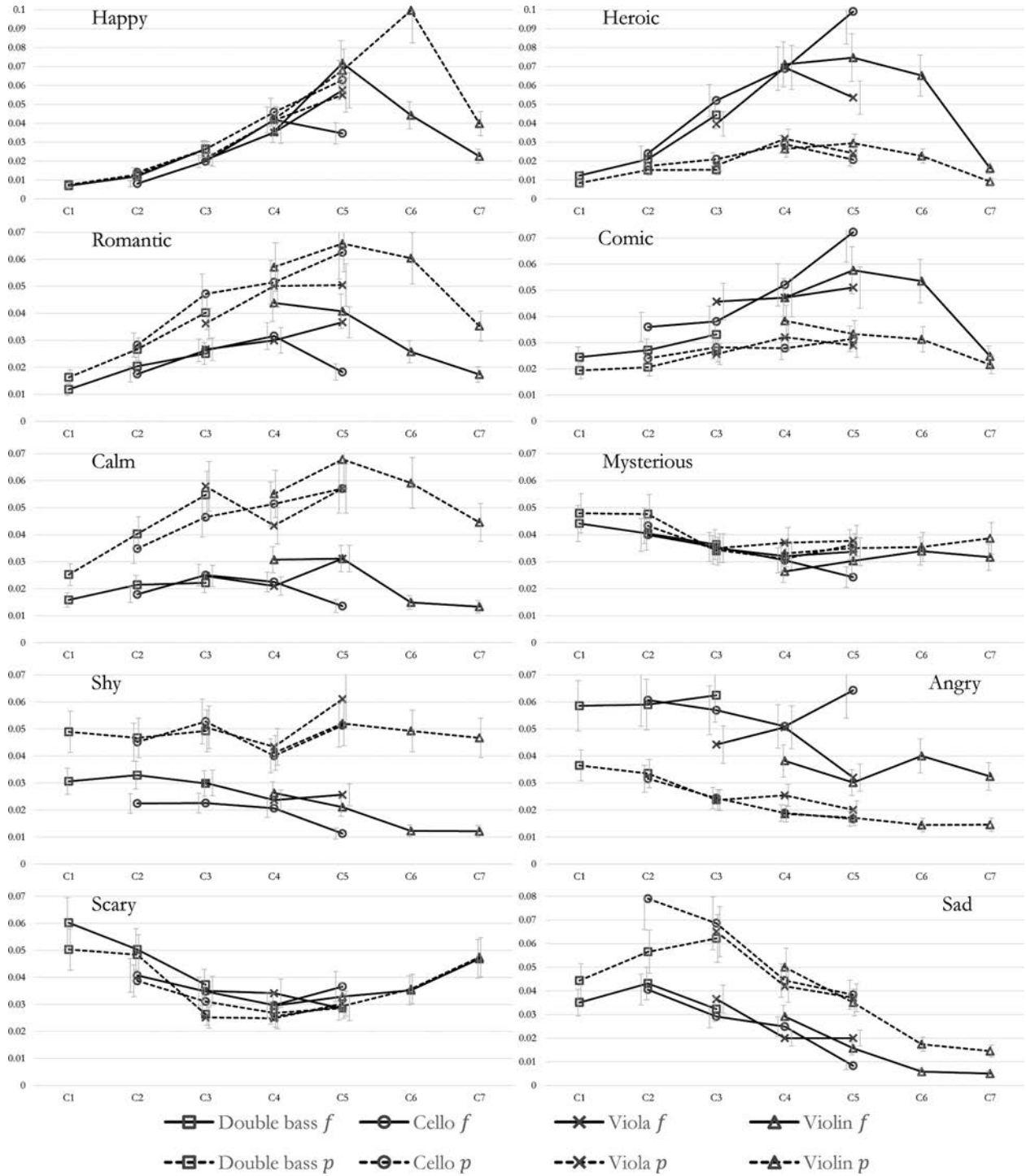


Fig. 2. Emotional characteristics of bowed string sounds based on the BTL scale values and the corresponding 95% confidence intervals. (*f* = forte, *p* = piano)

Table 8 in the Appendix), thus we applied the Greenhouse–Geisser correction. Table 3 shows the corrected ANOVA results.

Table 3 confirms that the most representative and less common instruments are largely in agreement. This means that the viola and double bass had about the same collective effects on pitch and dynamics as the violin and cello. In particular, the effect of pitch and dynamics were both significant for 7 or 8 of the 10 categories at the  $p < 0.05$  level.

Mysterious was basically flat and noisy and not significant for either pitch or dynamics. Shy was not significant for pitch since the *piano* curve was basically flat across pitch. Scary was not significant for dynamics since the *piano* and *forte* curves nearly overlapped.

For pitch, Scary was significant for the less common strings, but not for the most representative strings, indicating some instrument-dependence. Similarly, Happy was also instrument-dependent and significant for dynamics

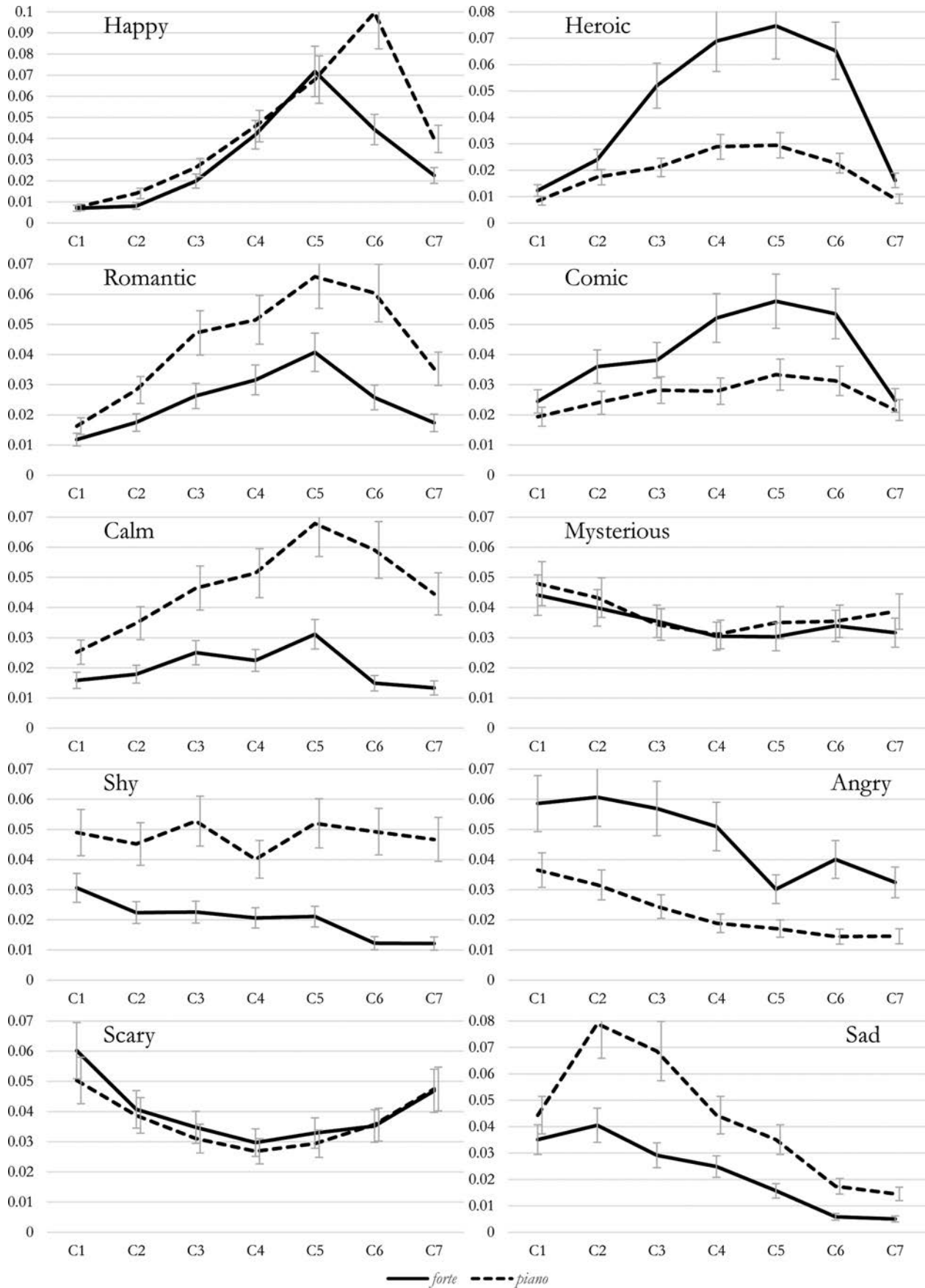


Fig. 3. Emotional characteristics of the “more representative” bowed string sounds based on the BTL scale values and the corresponding 95% confidence intervals.

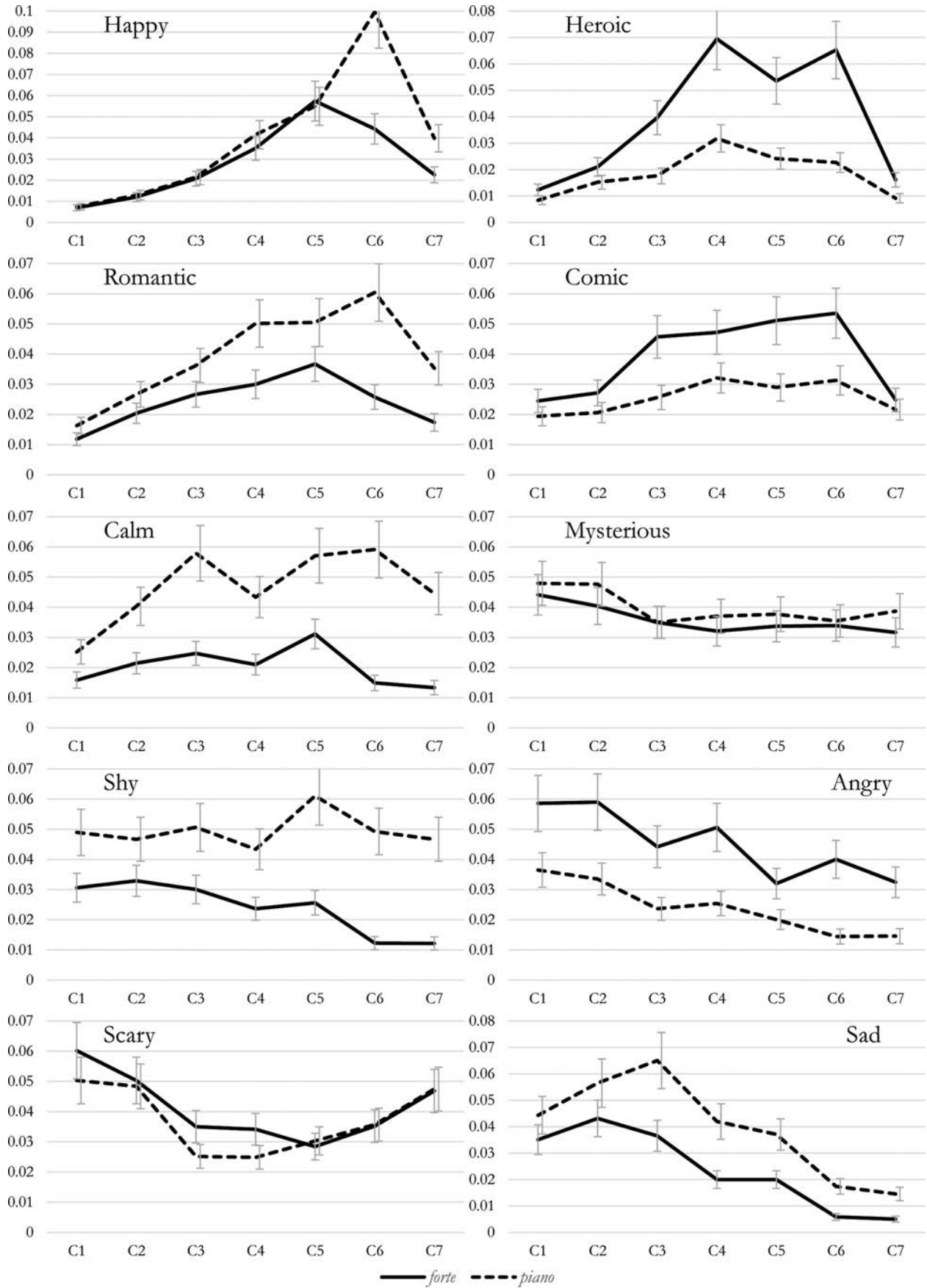


Fig. 4. Emotional characteristics of the “less common” bowed string sounds based on the BTL scale values and the corresponding 95% confidence intervals.

Table 3. *p*-values from the two-way ANOVA for the effects of pitch and dynamics. Values that were significant (*p* < 0.05) are shown in bold and shaded in grey.

	<i>Most Representative</i>		<i>Less Common</i>	
	Pitch	Dynamics	Pitch	Dynamics
Happy	<b>0.000</b>	<b>0.025</b>	<b>0.000</b>	0.069
Heroic	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
Romantic	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
Comic	<b>0.037</b>	<b>0.004</b>	<b>0.045</b>	<b>0.001</b>
Calm	<b>0.001</b>	<b>0.000</b>	<b>0.008</b>	<b>0.000</b>
Mysterious	0.467	0.388	0.512	0.217
Shy	0.318	<b>0.000</b>	0.107	<b>0.000</b>
Angry	<b>0.017</b>	<b>0.000</b>	<b>0.026</b>	<b>0.000</b>
Scary	0.072	0.508	<b>0.039</b>	0.255
Sad	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>

with the most representative strings but not with the less common strings. If you look carefully, comparing the category Happy in Figs. 3 and 4, there are significant differences at C2 and C3 for the most representative but not the less common strings.

### 2.2 Differences between the Individual Instruments

We identified differences between individual bowed string instruments by calculating BTL differences between instruments at the same pitch and dynamic level. Most of the 200 possible pairs (10 shared pitches between the 4 instruments × 2 dynamic levels × 10 emotional categories) were not significantly different. However, there were a number of exceptions. Table 4 lists the biggest BTL differences ordered from largest to smallest. The cello and C5 *forte* appear in 12 of the 13 largest pairs listed in Table 4. We will mention several notable examples. For C5 *forte* notes, the cello was regarded as much more Heroic, Angry, and Comic and less Happy, Romantic, and Calm than the violin and viola. For C2, the cello was considered more Sad than the double bass. The violin was ranked much more Heroic and Romantic than the viola at C5 or C4 *forte*, and less Happy.

Table 4. Biggest BTL differences between different bowed string instruments at the same pitch and dynamic level, ordered from largest to smallest.

Emotional Categories	Instruments	Sound	ΔBTL
Heroic	Vc > Va	C5f	0.0455
Happy	Vn > Vc	C5f	0.0371
Angry	Vc > Vn	C5f	0.0341
Angry	Vc > Va	C5f	0.0323
Heroic	Vn > Vc	C5f	0.0244
Happy	Va > Vc	C5f	0.0228
Romantic	Vn > Vc	C5f	0.0225
Sad	Vc > Cb	C2p	0.0225
Comic	Vc > Va	C5f	0.0212
Heroic	Vn > Va	C5f	0.0211
Romantic	Va > Vc	C5f	0.0184
Calm	Vn > Vc	C5f	0.0176
Calm	Va > Vc	C5f	0.0176
Romantic	Vn > Va	C5p	0.0154
Comic	Vc > Vn	C5f	0.0146
Shy	Va > Vc	C5f	0.0143
Happy	Va > Vn	C5f	0.0143
Romantic	Vn > Va	C4f	0.0139
Happy	Vn > Va	C5p	0.0130

As an alternative, more general perspective, Fig. 5 shows the percentage of cases where each instrument was significantly greater than other instruments for each category at the same pitch and dynamic level. So, for example the violin was significantly greater than the viola and cello about 40% of the time for Romantic. Based on the figure, we see that the violin was significantly greater than the other instruments for Romantic, Calm, and Sad. The double bass was significantly greater than the other instruments for Happy, Shy, and Angry.

Fig. 6 shows the total number of cases that an instrument was significantly greater than another instrument for each pitch. It only includes C2 to C5 since only the double bass is at C1 and only the violin is at C6 and C7. There were almost twice as many cases at C5 compared to all the other pitches together. This indicates that C5 is a hotspot for differentiating the individual string instruments.

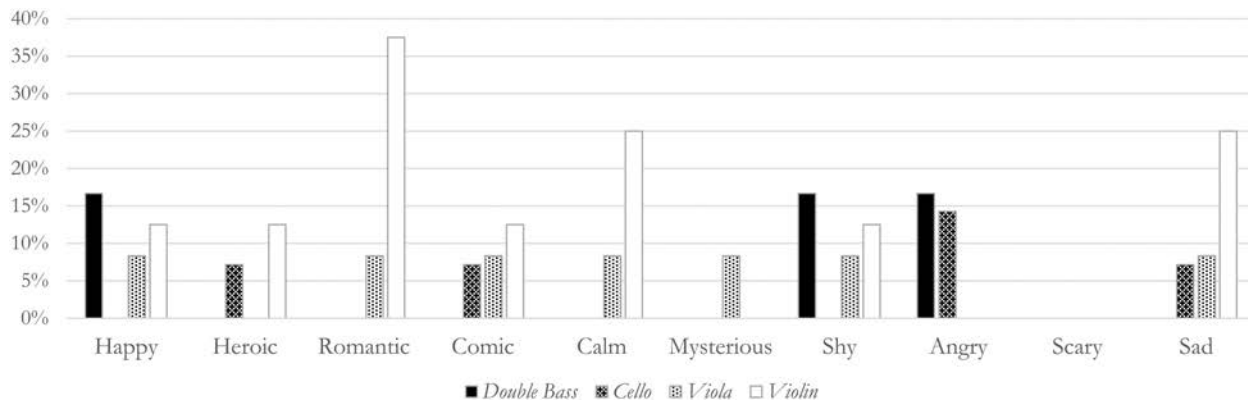


Fig. 5. Percentage of cases where each instrument was significantly greater than other instruments at the same pitch and dynamic level.



Table 5. Pearson correlation between emotional categories and features, with significant correlation values shown in bold (for  $p < 0.05$ ).

Features	Emotional Categories										No. of emotional categories with significant correlations
	Happy	Heroic	Romantic	Comic	Calm	Mysterious	Shy	Angry	Scary	Sad	
Log of Fundamental Frequency	<b>0.68</b>	0.27	<b>0.41</b>	0.32	0.19	<b>-0.57</b>	-0.15	<b>-0.42</b>	-0.31	<b>-0.68</b>	5
Peak RMS Amplitude (dB)	<b>-0.52</b>	<b>0.51</b>	<b>-0.59</b>	<b>0.40</b>	<b>-0.75</b>	-0.19	<b>-0.76</b>	<b>0.73</b>	0.25	-0.26	7
Spectral Centroid	<b>-0.66</b>	-0.37	<b>-0.54</b>	-0.35	-0.30	<b>0.68</b>	0.11	<b>0.38</b>	<b>0.54</b>	<b>0.46</b>	6
Spectral Centroid Deviation	<b>-0.68</b>	<b>-0.39</b>	<b>-0.55</b>	<b>-0.40</b>	-0.31	<b>0.70</b>	0.08	<b>0.43</b>	<b>0.53</b>	<b>0.47</b>	8
Spectral Incoherence	-0.17	-0.15	-0.36	-0.04	-0.33	<b>0.45</b>	-0.11	0.35	<b>0.55</b>	-0.10	2
Spectral Irregularity	0.12	-0.23	0.27	-0.23	0.37	-0.21	0.30	<b>-0.38</b>	-0.21	0.12	1
Tristimulus T1 (harmonic 1)	0.23	0.13	0.12	0.04	-0.03	-0.33	-0.24	-0.23	0.00	<b>-0.45</b>	1
Tristimulus T2 (harmonics 2-4)	<b>0.47</b>	0.02	<b>0.42</b>	0.16	<b>0.38</b>	-0.06	0.31	-0.19	<b>-0.38</b>	0.11	4
Tristimulus T3 (harmonics 5+)	<b>-0.59</b>	-0.18	<b>-0.43</b>	-0.15	-0.23	<b>0.45</b>	0.09	<b>0.40</b>	0.26	<b>0.48</b>	5
No. of features with significant correlations	6	2	6	2	2	5	1	6	4	5	

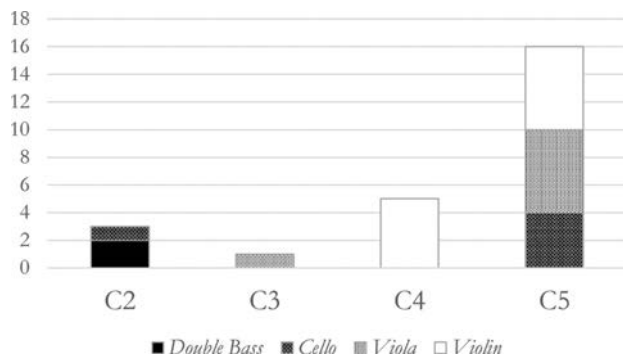


Fig. 6. Total number of cases that an instrument was significantly greater than another instrument for each pitch.

### 2.3 Correlation between Emotional Characteristics and Pitch, Dynamic, and Spectral Features

Pitch, dynamic, and spectral features for each of the sounds are given in Tables 9 and 10 in the Appendix (the spectral features are described in detail in [29, 62, 63]). Correlations between these features and the BTL values in Fig. 2 are shown in Table 5. Peak RMS Amplitude (dB) showed significant correlations for most emotional categories, confirming the importance of dynamics in emotional characteristics. Mysterious and Scary did not show correlation for the amplitude feature and showed less dynamic sensitivity in Fig. 2.

The pitch feature Log of Fundamental Frequency was also significant for half of the emotional categories. At the same time, 7 or 8 of the 10 emotional categories showed responses clearly varying in pitch in Fig. 3, but many were non-monotonic and distinctly arched or U-shaped (e.g., Heroic, Romantic, Comic, Calm, Scary in Fig. 2). Since

correlation is calculated linearly, the correspondence of pitch was underestimated by correlation. Happy was the most linear emotional characteristic in its relationship to pitch, especially between C2 and about C5 (similarly for Sad from C3 to C6).

The correlation values for Spectral Centroid (and its deviation) and Log of Fundamental Frequency in Table 5 are very similar, but with a change of sign, suggesting the strong inverse dependence of brightness on pitch. The correlation between these features based on the values in Tables 9 and 10 was  $-0.97$  ( $p < 10^{-5}$ ), indicating a very strong nearly-linear inverse correlation.

Aside from Spectral Centroid (and its deviation) other spectral features did not have as many significant correlations. Since pitch and dynamics were such strong factors, listeners probably did not focus on other spectral features as much.

Partial correlation analysis was also done to study the effects of pitch and dynamics. Table 6 shows the correlation between emotional categories and features when the effects of pitch and dynamics were removed. Interestingly, Spectral Centroid (and its deviation) remained correlated with more than half of the emotional categories. Spectral Incoherence increased in its number of significant differences indicating its relative importance after pitch and dynamics. This indicates that listeners frequently used spectral features as a secondary criterion to differentiate emotional characteristics in tones from different instruments but with the same pitch and dynamic levels. On the other hand, Tristimulus T3 greatly decreased indicating its relative dependence on pitch and dynamics.

The biggest changes among emotional characteristics were for Happy and Mysterious, which greatly decreased in the number of significant features from Table 5 to Table 6 when the effects of pitch and dynamics were removed. This

Table 6. Pearson partial correlation between emotional categories and features with the effects of pitch and dynamics removed. Significant correlation values shown in bold (for  $p < 0.05$ ).

Features	Emotional Categories										No. of emotional categories with significant correlations
	Happy	Heroic	Romantic	Comic	Calm	Mysterious	Shy	Angry	Scary	Sad	
Spectral Centroid	-0.34	-0.19	<b>-0.65</b>	-0.05	<b>-0.69</b>	<b>0.40</b>	<b>-0.46</b>	0.26	<b>0.65</b>	<b>-0.70</b>	6
Spectral Centroid Deviation	-0.38	-0.30	<b>-0.60</b>	-0.22	<b>-0.59</b>	<b>0.50</b>	<b>-0.42</b>	0.36	<b>0.58</b>	<b>-0.48</b>	6
Spectral Incoherence	-0.02	0.05	<b>-0.45</b>	0.20	<b>-0.65</b>	0.30	<b>-0.63</b>	<b>0.54</b>	<b>0.55</b>	<b>-0.76</b>	6
Spectral Irregularity	0.04	-0.24	0.23	-0.25	<b>0.41</b>	-0.28	<b>0.41</b>	<b>-0.42</b>	-0.18	0.21	3
Tristimulus T1 (harmonic 1)	-0.26	<b>-0.50</b>	0.01	<b>-0.64</b>	0.21	0.34	<b>0.40</b>	-0.36	0.23	<b>0.41</b>	4
Tristimulus T2 (harmonics 2-4)	<b>0.55</b>	<b>0.45</b>	0.29	<b>0.56</b>	0.09	-0.36	-0.16	0.15	-0.37	-0.22	3
Tristimulus T3 (harmonics 5+)	-0.2	0.22	-0.32	0.32	-0.38	-0.11	<b>-0.39</b>	0.34	0.05	-0.34	1
No. of features with significant correlations	1	2	3	2	4	2	6	2	3	4	

indicates Happy and Mysterious were strongly affected by pitch and dynamics and less affected by spectral features than other emotional categories.

Conversely, Shy greatly increased in the number of significant features from 1 to 6 when the effects of pitch and dynamics were removed, revealing its relative sensitivity to spectral features. Since pitch was not a significant factor for Shy, this indicates that listeners were keying on Spectral Centroid and other spectral fluctuations as secondary criteria to dynamics when making their judgments. For example, when listeners compared two tones both at the same dynamic level, they used spectral features to differentiate how Shy they were.

Looking at the the particular categories and features in Table 6, Spectral Centroid, Centroid Deviation, and Spectral Incoherence were strongly negatively correlated with Romantic, Calm, Shy, and Sad indicating sounds that were less bright and had fewer spectral fluctuations were considered more Romantic, Calm, Shy, and Sad. Sounds that were bright with more spectral fluctuations were considered more Scary and Angry. The Tristimulus features indicate that sounds with weaker fundamental frequencies relative to harmonics 2-4 were considered more Heroic and Comic for the bowed strings.

### 3 DISCUSSION

The main goal of our work was to determine how the emotional characteristics of bowed string instruments vary with pitch and dynamics. With respect to the original motivating questions of this paper, from Fig. 2 and Table 3 we can observe the following regarding pitch in the bowed strings:

- Eight out of ten emotional categories showed significant effects due to pitch.
- Happy, Heroic, Romantic, Comic, and Calm generally increased with pitch, but decreased at the highest

itches. They were distinctly arched and peaked at C5 (or C6).

- Angry and Sad generally decreased with pitch, though in different ways. Sad decreased after an initial increase in the lowest pitches of the double bass, while Angry started at a high level.
- Heroic, Comic, Calm, and Romantic were strongest in the mid-register, C4 to C5, and weaker in the highest and lowest registers.
- Shy and Mysterious were not significantly affected by pitch.
- Scary was strong in the lowest and highest registers, and weak in the mid-register between C3 and C6.

Regarding dynamics in the bowed strings:

- Eight of ten categories showed significant effects due to dynamics.
- Heroic, Comic, and Angry were stronger for loud notes.
- Romantic, Calm, Shy, and Sad were stronger for soft notes. Soft notes were also stronger in the high register for Happy.
- Surprisingly, Mysterious and Scary were not affected by dynamics.
- The high register had the widest gap between dynamics for Happy, Calm, and Shy. The middle register had the widest gap between dynamics for Heroic, Romantic, Comic, and Sad. The gap between dynamics was somewhat uniform across different registers for Angry.

Overall, the results showed that pitch generally had a similar effect on emotional categories with similar Valence. The high-Valence characteristics Happy, Heroic, Comic, and Calm had broadly similar shapes in Fig. 2 (mostly-increasing and arching), while the low-Valence characteristics Angry and Sad were decreasing. The middle-Valence characteristics, Mysterious and Shy, were unaffected by

pitch. Scary was the biggest exception, increasing with pitch in the high register rather than decreasing like the other low-Valence characteristics Angry and Sad. Dynamics had a similar effect on emotional categories with similar Arousal, though there were more exceptions. The high-Arousal characteristics Heroic, Comic, and Angry were strongest for loud notes, while the low-Arousal characteristics Calm, Shy, Romantic and Sad were strongest for soft notes. However, the high-Arousal categories Happy and Scary were relatively unaffected by dynamics. It seems strings could be terrifyingly Scary when loud, or suspensefully Scary when soft.

We expect some of these results are rather specific to bowed string instruments while others might be specific to sustaining instruments. Further research will uncover these specific differences. We would expect more differences compared to the non-sustaining instruments. Nevertheless, we were curious to see how the bowed string results compared to those we found in our previous study of another instrument, the piano.

The bowed strings results in Fig. 2 show some remarkable similarities to the piano results in Chau et al. [37]. But since bowed string instruments have a continuous tone while the piano does not, there exist expectable differences. For pitch, seven of ten emotional characteristics basically agreed in their overall trends (Heroic, Mysterious, and Shy were different). Heroic was most different with a decreasing trend for the piano and a mostly increasing arching shape for the strings. For dynamics, the similarities were even more striking, and all ten of the emotional categories basically agreed. It is not surprising that Heroic, Comic, and Angry would be loud for both piano and bowed strings, while Calm, Shy, and Sad would be soft. However, it is interesting that Happy, Mysterious, and Scary would be unaffected or relatively less affected by dynamics in both piano and bowed strings. The correlation between the bowed strings and piano BTL data was 0.60 ( $p < 10^{-14}$ ), indicating a strong correlation.

We suspect that the agreement in dynamics is probably fairly instrument-independent since categories such as Shy and Calm are inherently soft by nature. Pitch is almost certainly more instrument-dependent, since each instrument has its own particular pitch range and timbre. Further work with other instruments can help put these ideas on firmer footing.

As a disclaimer, musical features such as intention and articulation were not involved in the experiment. The results are therefore only for a generic context-free situation, and that these effects would be further modulated by context-dependent musical features such as melody, harmony, and phrasing.

The above results can give suggestions to musicians in orchestration, performers in blending and balancing instruments, and recording engineers in mixing recordings and live performances. Emotional characteristics can be manipulated in a recording, performance, or composition by emphasizing instruments, pitches, and dynamics that are comparatively stronger in representing these characteristics. The results confirm some existing com-

mon practices for emotional emphasis (e.g., using low double basses and high violins together for Scary passages). However, they also identify some less-commonly understood characteristics of the bowed strings, such as the Angry and Comic qualities of the high cello at loud dynamics.

## 4 ACKNOWLEDGMENTS

Thanks very much to the anonymous reviewers for their careful and insightful comments to improve the clarity, presentation, and analysis in the the paper.

## 5 REFERENCES

- [1] L. Lu, D. Liu, and H.-J. Zhang, "Automatic Mood Detection and Tracking of Music Audio Signals," *IEEE Trans. Audio Speech Lang. Process.*, vol. 14, no. 1, pp. 5–18 (2006), doi:<https://doi.org/10.1109/TSA.2005.860344>.
- [2] J. Skowronek, M. F. McKinney, and S. Van De Par, "A Demonstrator for Automatic Music Mood Estimation," *Proc. Int. Soc. Music Inform. Retrieval Conf. (ISMIR)*, pp. 345–346 (2007).
- [3] R. O. Gjerdingen and D. Perrott, "Scanning the Dial: The Rapid Recognition of Music Genres," *J. New Music Research*, vol. 37, no. 2, pp. 93–100 (2008), doi:<https://doi.org/10.1080/09298210802479268>.
- [4] Y.-H. Yang, Y.-C. Lin, Y.-F. Su, and H. H. Chen, "A Regression Approach to Music Emotion Recognition," *IEEE Trans. Audio Speech Lang. Process.*, vol. 16, no. 2, pp. 448–457 (2008), doi:<https://doi.org/10.1109/TASL.2007.911513>.
- [5] Y. Hu, X. Chen, and D. Yang, "Lyric-Based Song Emotion Detection with Affective Lexicon and Fuzzy Clustering Method," *Proc. Int. Soc. Music Inform. Retrieval Conf. (ISMIR)*, pp. 123–128 (2009).
- [6] C. Laurier, M. Sordo, J. Serra, and P. Herrera, "Music Mood Representations from Social Tags," *Proc. Int. Soc. Music Inform. Retrieval Conf. (ISMIR)*, pp. 381–386 (2009).
- [7] B. Kostek, "Content-Based Approach to Automatic Recommendation of Music," presented at the *131st Convention of the Audio Engineering Society* (2011 Oct.), convention paper 8505.
- [8] R. Panda and R. P. Paiva, "Using Support Vector Machines for Automatic Mood Tracking in Audio Music," presented at the *130th Convention of the Audio Engineering Society* (2011 May), convention paper 8378.
- [9] C. Lee, A. Horner, and J. Beauchamp, "Impact of MP3-Compression on Timbre Space of Sustained Musical Instrument Tones," *J. Acoust. Soc. Amer.*, vol. 131, no. 4, p. 3433 (2012), doi:<https://doi.org/10.1121/1.4708878>.
- [10] B. Den Brinker, R. Van Dither, and J. Skowronek, "Expressed Music Mood Classification Compared with Valence and Arousal Ratings," *EURASIP J. Audio, Speech, and Music Processing*, vol. 2012, no. 1, pp. 1–14 (2012), doi:<https://doi.org/10.1186/1687-4722-2012-24>.
- [11] B. Kostek and M. Plewa, "Parametrisation and Correlation Analysis Applied to Music

Mood Classification,” *Int. J. Computational Intelligence Studies*, vol. 2, no. 1, pp. 4–25 (2013), doi:<https://doi.org/10.1504/IJCISTUDIES.2013.054734>.

[12] C. Baume, “Evaluation of Acoustic Features for Music Emotion Recognition,” presented at the *134th Convention of the Audio Engineering Society* (2013 May), convention paper 8811.

[13] P. Saari, T. Eerola, G. Fazekas, M. Barthelet, O. Lartillot, and M. B. Sandler, “The Role of Audio and Tags in Music Mood Prediction: A Study Using Semantic Layer Projection,” *Proc. Int. Soc. Music Inform. Retrieval Conf. (ISMIR)*, pp. 201–206 (2013).

[14] I. Peretz, L. Gagnon, and B. Bouchard, “Music and Emotion: Perceptual Determinants, Immediacy, and Isolation after Brain Damage,” *Cognition*, vol. 68, no. 2, pp. 111–141 (1998), doi:[https://doi.org/10.1016/S0010-0277\(98\)00043-2](https://doi.org/10.1016/S0010-0277(98)00043-2).

[15] P. N. Juslin, “Cue Utilization in Communication of Emotion in Music Performance: Relating Performance to Perception,” *J. Experimental Psychology Human Perception and Performance*, vol. 26, no. 6, pp. 1797–1813 (2000), doi:<https://doi.org/10.1037/0096-1523.26.6.1797>.

[16] G. Tzanetakis and P. Cook, “Musical Genre Classification of Audio Signals,” *IEEE Trans. Speech Audio Process.*, vol. 10, no. 5, pp. 293–302 (2002), doi:<https://doi.org/10.1109/TSA.2002.800560>.

[17] W. Ellermeier, M. Mader, and P. Daniel, “Scaling the Unpleasantness of Sounds According to the BTL Model: Ratio-Scale Representation and Psychoacoustical Analysis,” *Acta Acustica united with Acustica*, vol. 90, no. 1, pp. 101–107 (2004).

[18] M. Leman, V. Vermeulen, L. De Voogdt, D. Moeulants, and M. Lesaffre, “Prediction of Musical Affect Using a Combination of Acoustic Structural Cues,” *J. New Music Research*, vol. 34, no. 1, pp. 39–67 (2005), doi:<https://doi.org/10.1080/09298210500123978>.

[19] E. Asutay, D. Västfjäll, A. Tajadura-Jiménez, A. Genell, P. Bergman, and M. Kleiner, “Emoacoustics: A Study of the Psychoacoustical and Psychological Dimensions of Emotional Sound Design,” *J. Audio Eng. Soc.*, vol. 60, pp. 21–28 (2012 Jan./Feb.).

[20] S. Lui, “Generate Expressive Music from Picture with a Handmade Multi-Touch Music Table,” *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, pp. 374–377 (2015).

[21] G. Leslie, R. Picard, and S. Lui, “An EEG and Motion Capture Based Expressive Music Interface for Affective Neurofeedback,” *Proc. 1st Int. BCMI Workshop* (2015).

[22] K. Trochidis and S. Lui, “Modeling Affective Responses to Music Using Audio Signal Analysis and Physiology,” in *Music, Mind, and Embodiment*, pp. 346–357 (Springer, 2016), doi:[https://doi.org/10.1007/978-3-319-46282-0\\_22](https://doi.org/10.1007/978-3-319-46282-0_22).

[23] T. Eerola, R. Ferrer, and V. Alluri, “Timbre and Affect Dimensions: Evidence from Affect and Similarity Ratings and Acoustic Correlates of Isolated Instrument Sounds,” *Music Perception*, vol. 30, no. 1, pp. 49–70 (2012), doi:<https://doi.org/10.1525/mp.2012.30.1.49>.

[24] B. Wu, S. Wun, C. Lee, and A. Horner, “Spectral Correlates in Emotion Labeling of Sustained Musical Instrument Tones,” *Proc. 14th Int. Soc. Music Inform. Retrieval Conf. (ISMIR)*, pp. 415–420 (2013 November 4–8).

[25] B. Wu, A. Horner, and C. Lee, “Musical Timbre and Emotion: The Identification of Salient Timbral Features in Sustained Musical Instrument Tones Equalized in Attack Time and Spectral Centroid,” *Proc. 40th Int. Comp. Music Conf. (ICMC)*, pp. 928–934 (2014).

[26] C.-j. Chau, B. Wu, and A. Horner, “Timbre Features and Music Emotion in Plucked String, Mallet Percussion, and Keyboard Tones,” *Proc. 40th Int. Comp. Music Conf. (ICMC)*, pp. 982–989 (2014).

[27] B. Wu, A. Horner, and C. Lee, “Emotional Predisposition of Musical Instrument Timbres with Static Spectra,” *Proc. 15th Int. Soc. Music Inform. Retrieval Conf. (ISMIR)*, pp. 253–258 (2014 Nov.).

[28] B. Wu, A. Horner, and C. Lee, “The Correspondence of Music Emotion and Timbre in Sustained Musical Instrument Sounds,” *J. Audio Eng. Soc.*, vol. 62, pp. 663–675 (2014 Oct.), doi:<https://doi.org/10.17743/jaes.2014.0037>.

[29] C.-j. Chau, B. Wu, and A. Horner, “The Emotional Characteristics and Timbre of Nonsustaining Instrument Sounds,” *J. Audio Eng. Soc.*, vol. 63, pp. 228–244 (2015 Apr.), doi:<https://doi.org/10.17743/jaes.2015.0016>.

[30] C.-j. Chau, B. Wu, and A. Horner, “The Effects of Early-Release on Emotion Characteristics and Timbre in Non-Sustaining Musical Instrument Tones,” *Proc. 41st Int. Comp. Music Conf. (ICMC)*, pp. 138–141 (2015).

[31] S. B. Kamenetsky, D. S. Hill, and S. E. Trehub, “Effect of Tempo and Dynamics on the Perception of Emotion in Music,” *Psychology of Music*, vol. 25, pp. 149–160 (1997), doi:<https://doi.org/10.1177/0305735697252005>.

[32] C. L. Krumhansl, “An Exploratory Study of Musical Emotions and Psychophysiology,” *Canadian J. Experimental Psychology/Revue canadienne de psychologie expérimentale*, vol. 51, no. 4, pp. 336–353 (1997), doi:<https://doi.org/10.1037/1196-1961.51.4.336>.

[33] D. Huron, D. Kinney, and K. Precoda, “Relation of Pitch Height to Perception of Dominance/Submissiveness in Musical Passages,” *Music Perception*, vol. 10, no. 1, pp. 83–92 (2000), doi:<https://doi.org/10.2307/40285540>.

[34] S. Lui, “A Preliminary Analysis of the Continuous Axis Value of the Three-Dimensional PAD Speech Emotional State Model,” *Proc. 16th Int. Conf. on Digital Audio Effects (DAFx)* (2013).

[35] C. Lee, S. Lui, and C. So, “Visualization of Time-Varying Joint Development of Pitch and Dynamics for Speech Emotion Recognition,” *J. Acoust. Soc. of Amer.*, vol. 135, no. 4, pp. 2422–2422 (2014), doi:<https://doi.org/10.1121/1.4878044>.

[36] C.-j. Chau and A. Horner, “The Effects of Pitch and Dynamics on the Emotional Characteristics of Piano Sounds,” *Proc. 41st Int. Comp. Music Conf. (ICMC)*, pp. 372–375 (2015).

[37] C.-j. Chau, R. Mo, and A. Horner, “The Emotional Characteristics of Piano Sounds with Different Pitch and



- Dynamics,” *J. Audio Eng. Soc.*, vol. 64, pp. 918–932 (2016 Nov.), doi:<https://doi.org/10.17743/jaes.2016.0049>.
- [38] J. C. Brown and K. V. Vaughn, “Pitch Center of Stringed Instrument Vibrato Tones,” *J. Acoust. Soc. of Amer.*, vol. 100, no. 3, pp. 1728–1735 (1996), doi:<https://doi.org/10.1121/1.416070>.
- [39] C. L. Krumhansl, “Topic in Music: An Empirical Study of Memorability, Openness, and Emotion in Mozart’s String Quintet in C Major and Beethoven’s String Quartet in A Minor,” *Music Perception*, vol. 16, no. 1, pp. 119–134 (Fall 1998), doi:<https://doi.org/10.2307/40285781>.
- [40] J. Rothstein, “ProSonus Studio Reference Disk and Sample Library Compact Disks,” *Comp. Music Journal*, vol. 4, no. 13, pp. 92–93 (1989).
- [41] C. L. Krumhansl, “Plink: ‘Thin Slices’ of Music,” *Musical Perception*, vol. 7, pp. 337–354 (2010), doi:<https://doi.org/10.1525/mp.2010.27.5.337>.
- [42] S. Filipic, B. Tillmann, and E. Bigand, “Judging Familiarity and Emotion from Very Brief Musical Excerpts,” *Psychonomic Bulletin & Review*, vol. 17, no. 3, pp. 335–341 (2010), doi:<https://doi.org/10.3758/PBR.17.3.335>.
- [43] J. W. Beauchamp, “Analysis and Synthesis of Musical Instrument Sounds,” in *Analysis, Synthesis, and Perception of Musical Sounds*, pp. 1–89 (Springer, 2007), doi:[https://doi.org/10.1007/978-0-387-32576-7\\_1](https://doi.org/10.1007/978-0-387-32576-7_1).
- [44] R. Mo, B. Wu, and A. Horner, “The Effects of Reverberation on the Emotional Characteristics of Musical Instruments,” *J. Audio Eng. Soc.*, vol. 63, pp. 966–979 (2015 Dec.), doi:<https://doi.org/10.17743/jaes.2015.0082>.
- [45] R. Mo, G. L. Choi, C. Lee, and A. Horner, “The Effects of MP3 Compression on Emotional Characteristics,” *Proc. 42th Int. Comp. Music Conf. (ICMC)*, pp. 411–416 (2016).
- [46] R. Mo, G. L. Choi, C. Lee, and A. Horner, “The Effects of MP3 Compression on Perceived Emotional Characteristics in Musical Instruments,” *J. Audio Eng. Soc.*, vol. 64, pp. 858–867 (2016 Nov.), doi:<https://doi.org/10.17743/jaes.2016.0031>.
- [47] C.-j. Chau and A. Horner, “The Emotional Characteristics of Mallet Percussion Instruments with Different Pitches and Mallet Hardness,” *Proc. 42th Int. Comp. Music Conf. (ICMC)*, pp. 401–404 (2016).
- [48] R. Mo, R. H. Y. So, and A. Horner, “An Investigation into How Reverberation Effects the Space of Instrument Emotional Characteristics,” *J. Audio Eng. Soc.*, vol. 64, pp. 988–1002 (2016 Dec.), doi:<https://doi.org/10.17743/jaes.2016.0054>.
- [49] M. Kennedy and K. Joyce Bourne, *The Oxford Dictionary of Music* (Oxford University Press, 2012).
- [50] Dolmetsch Organisation, “Dolmetsch Online - Music Dictionary,” URL <http://www.dolmetsch.com/musictheorydefs.htm>.
- [51] Classical.dj, “Classical Musical Terms,” URL [http://classical.dj/musical\\_terms.html](http://classical.dj/musical_terms.html).
- [52] Connect For Education Inc., “OnMusic Dictionary,” URL <http://dictionary.onmusic.org/>.
- [53] K. Hevner, “Experimental Studies of the Elements of Expression in Music,” *Amer. J. Psych.*, vol. 48, no. 2, pp. 246–268 (1936 Apr.), doi:<https://doi.org/10.2307/1415746>.
- [54] J. A. Russell, “A Circumplex Model of Affect,” *J. Personality and Social Psychology*, vol. 39, no. 6, p. 1161 (1980), doi:<https://doi.org/10.1037/h0077714>.
- [55] K. R. Scherer and J. S. Oshinsky, “Cue Utilization in Emotion Attribution from Auditory Stimuli,” *Motivation and Emotion*, vol. 1, no. 4, pp. 331–346 (1977), doi:<https://doi.org/10.1007/BF00992539>.
- [56] M. Zentner, D. Grandjean, and K. R. Scherer, “Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement,” *Emotion*, vol. 8, no. 4, p. 494 (2008), doi:<https://doi.org/10.1037/1528-3542.8.4.494>.
- [57] J. C. Hailstone, R. Omar, S. M. Henley, C. Frost, M. G. Keyword, and J. D. Warren, “It’s Not What You Play, It’s How You Play It: Timbre Affects Perception of Emotion in Music,” *Quarterly J. Experimental Psychology*, vol. 62, no. 11, pp. 2141–2155 (2009), doi:<https://doi.org/10.1080/17470210902765957>.
- [58] S. J. M. Gilburt, C.-j. Chau, and A. Horner, “The Effects of Pitch and Dynamics on the Emotional Characteristics of Bowed String Instruments,” *Proc. 42th Int. Comp. Music Conf. (ICMC)*, pp. 405–410 (2016).
- [59] Cambridge University Press, “Cambridge Academic Content Dictionary,” URL <http://dictionary.cambridge.org/dictionary/american-english>.
- [60] F. Wickelmaier and C. Schmid, “A Matlab Function to Estimate Choice Model Parameters from Paired-Comparison Data,” *Behavior Research Methods, Instruments, and Computers*, vol. 36, no. 1, pp. 29–40 (2004), doi:<https://doi.org/10.3758/BF03195547>.
- [61] R. A. Bradley, “Paired Comparisons: Some Basic Procedures and Examples,” *Nonparametric Methods*, vol. 4, pp. 299–326 (1984), doi:[https://doi.org/10.1016/S0169-7161\(84\)04016-5](https://doi.org/10.1016/S0169-7161(84)04016-5).
- [62] J. W. Beauchamp and A. Horner, “Error Metrics for Predicting Discrimination of Original and Spectrally Altered Musical Instrument Sounds,” *J. Acoust. Soc. of Amer.*, vol. 114, no. 4, pp. 2325–2325 (2003), doi:<https://doi.org/10.1121/1.4781005>.
- [63] A. B. Horner, J. W. Beauchamp, and R. H. So, “A Search for Best Error Metrics to Predict Discrimination of Original and Spectrally Altered Musical Instrument Sounds,” *J. Audio Eng. Soc.*, vol. 54, pp. 140–156 (2006 Mar.).

APPENDIX

Table 7. Results of a Shapiro–Wilk test to check the normality of the pairwise voting data. Entries in bold and shaded in grey were normally distributed.

	Happy	Heroic	Romantic	Comic	Calm	Mysterious	Shy	Angry	Scary	Sad
CbC1f	<b>0.224</b>	<b>0.436</b>	<b>0.063</b>	<b>0.381</b>	<b>0.425</b>	<b>0.622</b>	<b>0.159</b>	<b>0.345</b>	<b>0.431</b>	<b>0.542</b>
CbC1p	0.023	<b>0.533</b>	<b>0.171</b>	<b>0.176</b>	<b>0.675</b>	<b>0.523</b>	<b>0.480</b>	<b>0.452</b>	<b>0.233</b>	<b>0.620</b>
CbC2f	<b>0.076</b>	<b>0.806</b>	<b>0.954</b>	<b>0.116</b>	<b>0.659</b>	<b>0.318</b>	<b>0.800</b>	<b>0.876</b>	<b>0.917</b>	<b>0.576</b>
CbC2p	<b>0.296</b>	<b>0.282</b>	<b>0.589</b>	<b>0.718</b>	<b>0.151</b>	<b>0.458</b>	<b>0.107</b>	<b>0.318</b>	<b>0.050</b>	<b>0.412</b>
CbC3f	<b>0.466</b>	<b>0.795</b>	<b>0.809</b>	<b>0.723</b>	<b>0.484</b>	<b>0.746</b>	<b>0.336</b>	<b>0.788</b>	<b>0.054</b>	<b>0.868</b>
CbC3p	<b>0.684</b>	<b>0.281</b>	<b>0.087</b>	<b>0.430</b>	<b>0.259</b>	<b>0.555</b>	<b>0.363</b>	<b>0.314</b>	<b>0.461</b>	<b>0.561</b>
VcC2f	0.046	<b>0.310</b>	<b>0.824</b>	<b>0.673</b>	<b>0.129</b>	<b>0.207</b>	<b>0.120</b>	<b>0.308</b>	<b>0.878</b>	<b>0.188</b>
VcC2p	<b>0.486</b>	0.007	<b>0.412</b>	<b>0.625</b>	<b>0.330</b>	<b>0.134</b>	<b>0.316</b>	<b>0.975</b>	<b>0.801</b>	<b>0.630</b>
VcC3f	<b>0.125</b>	<b>0.178</b>	<b>0.119</b>	<b>0.248</b>	<b>0.772</b>	<b>0.601</b>	<b>0.881</b>	<b>0.813</b>	0.045	<b>0.351</b>
VcC3p	0.013	<b>0.839</b>	<b>0.448</b>	<b>0.340</b>	<b>0.293</b>	<b>0.626</b>	0.031	<b>0.062</b>	<b>0.431</b>	<b>0.084</b>
VcC4f	<b>0.563</b>	<b>0.184</b>	<b>0.308</b>	<b>0.571</b>	<b>0.348</b>	<b>0.773</b>	0.031	<b>0.476</b>	<b>0.511</b>	0.018
VcC4p	<b>0.512</b>	<b>0.439</b>	<b>0.294</b>	<b>0.472</b>	<b>0.923</b>	<b>0.688</b>	<b>0.708</b>	<b>0.179</b>	<b>0.294</b>	<b>0.112</b>
VcC5f	<b>0.157</b>	<b>0.060</b>	<b>0.390</b>	<b>0.445</b>	<b>0.367</b>	<b>0.256</b>	0.002	<b>0.282</b>	<b>0.512</b>	0.005
VcC5p	<b>0.775</b>	<b>0.264</b>	<b>0.218</b>	<b>0.073</b>	0.021	<b>0.342</b>	0.038	0.046	<b>0.714</b>	<b>0.286</b>
VaC3f	<b>0.283</b>	<b>0.856</b>	<b>0.451</b>	<b>0.180</b>	<b>0.521</b>	<b>0.155</b>	<b>0.407</b>	<b>0.849</b>	<b>0.413</b>	<b>0.665</b>
VaC3p	<b>0.211</b>	<b>0.476</b>	<b>0.205</b>	<b>0.958</b>	<b>0.600</b>	<b>0.268</b>	0.011	<b>0.099</b>	<b>0.518</b>	<b>0.642</b>
VaC4f	<b>0.814</b>	<b>0.636</b>	<b>0.425</b>	<b>0.906</b>	<b>0.065</b>	<b>0.199</b>	0.003	<b>0.343</b>	<b>0.331</b>	<b>0.552</b>
VaC4p	<b>0.879</b>	<b>0.521</b>	<b>0.183</b>	<b>0.882</b>	<b>0.098</b>	<b>0.799</b>	<b>0.072</b>	<b>0.128</b>	<b>0.077</b>	<b>0.686</b>
VaC5f	<b>0.407</b>	0.026	0.042	<b>0.136</b>	0.007	<b>0.583</b>	<b>0.336</b>	0.009	<b>0.902</b>	<b>0.190</b>
VaC5p	<b>0.113</b>	<b>0.506</b>	<b>0.143</b>	<b>0.153</b>	<b>0.126</b>	<b>0.150</b>	<b>0.419</b>	<b>0.595</b>	<b>0.402</b>	<b>0.900</b>
VnC4f	<b>0.363</b>	<b>0.511</b>	<b>0.620</b>	<b>0.138</b>	<b>0.348</b>	<b>0.981</b>	<b>0.603</b>	<b>0.508</b>	<b>0.705</b>	0.011
VnC4p	<b>0.903</b>	<b>0.639</b>	<b>0.513</b>	<b>0.320</b>	<b>0.958</b>	<b>0.098</b>	<b>0.948</b>	<b>0.156</b>	<b>0.333</b>	<b>0.183</b>
VnC5f	<b>0.070</b>	<b>0.143</b>	<b>0.784</b>	<b>0.173</b>	0.028	<b>0.161</b>	<b>0.294</b>	<b>0.342</b>	<b>0.137</b>	<b>0.078</b>
VnC5p	<b>0.560</b>	<b>0.216</b>	<b>0.538</b>	<b>0.700</b>	<b>0.389</b>	<b>0.836</b>	0.032	<b>0.440</b>	<b>0.594</b>	<b>0.116</b>
VnC6f	<b>0.065</b>	<b>0.053</b>	<b>0.211</b>	<b>0.396</b>	<b>0.130</b>	<b>0.192</b>	0.005	<b>0.398</b>	<b>0.462</b>	0.002
VnC6p	0.012	<b>0.570</b>	<b>0.840</b>	<b>0.242</b>	0.007	<b>0.999</b>	<b>0.185</b>	<b>0.361</b>	<b>0.617</b>	0.013
VnC7f	0.001	<b>0.063</b>	<b>0.314</b>	<b>0.737</b>	<b>0.158</b>	<b>0.400</b>	0.003	<b>0.669</b>	<b>0.359</b>	0.000
VnC7p	<b>0.341</b>	0.011	<b>0.763</b>	<b>0.406</b>	<b>0.422</b>	<b>0.726</b>	0.006	<b>0.267</b>	<b>0.717</b>	<b>0.101</b>

Table 8. Results of the Mauchly’s Sphericity Test on the data for ANOVA.

	<i>Most Representative</i>		<i>Less Common</i>	
	Sig.	Epsilon	Sig.	Epsilon
Happy	<b>0.000</b>	0.450	<b>0.000</b>	0.420
Heroic	<b>0.000</b>	0.509	<b>0.000</b>	0.392
Romantic	<b>0.000</b>	0.522	<b>0.000</b>	0.436
Comic	<b>0.000</b>	0.387	<b>0.000</b>	0.417
Calm	<b>0.000</b>	0.516	<b>0.000</b>	0.424
Mysterious	<b>0.000</b>	0.330	<b>0.000</b>	0.270
Shy	<b>0.000</b>	0.485	<b>0.000</b>	0.382
Angry	<b>0.000</b>	0.250	<b>0.000</b>	0.307
Scary	<b>0.000</b>	0.399	<b>0.000</b>	0.394
Sad	<b>0.000</b>	0.480	<b>0.000</b>	0.416

Table 9. Pitch, dynamic, and spectral features of the sounds of the double bass (Cb) and cello (Vc).

Features Sounds	CbC1f	CbC1p	CbC2f	CbC2p	CbC3f	CbC3p	VcC2f	VcC2p	VcC3f	VcC3p	VcC4f	VcC4p	VcC5f	VcC5p
Log of Fundamental Frequency	5.03	5.03	6.03	6.03	7.03	7.03	6.03	6.03	7.03	7.03	8.03	8.03	9.03	9.03
Peak RMS Amplitude (dB)	-1.82	-9.93	-0.95	-9.81	-0.38	-14.94	-6.03	-12.71	-3.49	-15.84	0.00	-7.64	-8.85	-19.80
Spectral Centroid	15.88	16.97	7.75	7.88	6.02	5.70	12.11	11.92	7.54	7.89	4.47	3.82	5.25	4.46
Spectral Centroid Deviation	2.63	4.55	2.79	2.33	1.52	1.40	4.30	2.56	1.69	1.44	1.06	0.75	0.57	0.56
Spectral Incoherence	0.29	0.25	0.18	0.14	0.10	0.08	0.34	0.15	0.14	0.15	0.11	0.06	0.20	0.14
Spectral Irregularity	0.11	0.15	0.05	0.05	0.09	0.13	0.05	0.07	0.12	0.12	0.17	0.20	0.08	0.13
Tristimulus T1 (harmonic 1)	0.01	0.01	0.30	0.30	0.33	0.30	0.01	0.02	0.14	0.11	0.42	0.48	0.12	0.16
Tristimulus T2 (harmonics 2-4)	0.42	0.41	0.31	0.28	0.44	0.51	0.42	0.49	0.34	0.36	0.27	0.27	0.49	0.56
Tristimulus T3 (harmonics 5+)	0.57	0.58	0.39	0.41	0.24	0.20	0.57	0.50	0.52	0.53	0.31	0.25	0.39	0.28

Table 10. Pitch, dynamic, and spectral features of the sounds of the viola (Va) and violin (Vn).

Features Sounds	VaC3f	VaC3p	VaC4f	VaC4p	VaC5f	VaC5p	VnC4f	VnC4p	VnC5f	VnC5p	VnC6f	VnC6p	VnC7f	VnC7p
Log of Fundamental Frequency	7.03	7.03	8.03	8.03	9.03	9.03	8.03	8.03	9.03	9.03	10.03	10.03	11.03	11.03
Peak RMS Amplitude (dB)	-0.95	-11.32	-2.40	-11.40	-11.47	-22.81	-1.55	-7.65	-0.19	-16.73	-4.47	-18.88	-5.93	-16.57
Spectral Centroid	8.38	7.35	5.11	4.48	4.17	4.41	4.09	3.51	2.76	3.95	2.59	2.87	1.84	1.67
Spectral Centroid Deviation	1.01	1.31	0.80	0.71	0.45	0.62	0.61	0.60	0.33	0.46	0.30	0.31	0.17	0.15
Spectral Incoherence	0.09	0.04	0.11	0.10	0.15	0.19	0.09	0.08	0.09	0.16	0.20	0.21	0.09	0.10
Spectral Irregularity	0.17	0.17	0.04	0.04	0.09	0.13	0.16	0.16	0.05	0.13	0.02	0.10	0.15	0.15
Tristimulus T1 (harmonic 1)	0.00	0.00	0.26	0.28	0.20	0.17	0.29	0.30	0.49	0.24	0.41	0.20	0.70	0.69
Tristimulus T2 (harmonics 2-4)	0.41	0.47	0.38	0.41	0.51	0.51	0.47	0.51	0.33	0.46	0.45	0.66	0.18	0.23
Tristimulus T3 (harmonics 5+)	0.58	0.53	0.36	0.31	0.28	0.33	0.24	0.19	0.18	0.30	0.14	0.13	0.11	0.08

**THE AUTHORS**



Chuck-jee Chau



Samuel J. M. Gilbert



Ronald Mo



Andrew Horner

Chuck-jee Chau is a Ph.D. student in the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology (HKUST). His research focuses on timbre analysis and music emotion. During his master studies he developed the timbre visualization tool *pvan+* for phase vocoder analysis. He obtained his BEng in computer engineering from the Chinese University of Hong Kong (CUHK) with a minor in music. Besides computer music research, he is also a versatile collaborative pianist and mallet percussionist active in chamber music performances.

Samuel J. M. Gilbert is an undergraduate student in the School of Computing Science at Newcastle University in the United Kingdom. He is studying for a MComp integrated master's degree, which also included study abroad at the Hong Kong University of Science and Technology in 2014/15. His academic interests include software engineering and development, computer security, and computer music. Aside from academia, Sam is an accomplished

orchestral cellist and choral singer, having performed in renowned venues in the UK and across Europe.

Ronald Mo is pursuing his Ph.D. in the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology. His research interests include timbre of musical instruments, music emotion recognition, and digital signal processing. He received his B.Eng. of computer science and M.Phil. of computer science and engineering from the Hong Kong University of Science and Technology in 2007 and 2015 respectively.

Andrew Horner is a professor in the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology. His research interests include music analysis and synthesis, timbre of musical instruments, and music emotion. He received his Ph.D. in computer science from the University of Illinois at Urbana-Champaign.