How Does Parametric Reverberation Change the Space of Instrument Emotional Characteristics?

Ronald K. Mo
Department of Computer Science
and Engineering, HKUST
ronmo@cse.ust.hk

Richard H.Y. So
Department of Industrial Engineering
and Logistics Management, HKUST
rhyso@ust.hk

Andrew Horner
Department of Computer Science
and Engineering, HKUST
horner@cse.ust.hk

ABSTRACT

Previous research has shown that the distinctive emotional characteristics in musical instruments can be significantly changed with reverberation. This paper considers whether these changes in character are relatively uniform or instrument-dependent. We compared eight sustained instrument tones with different amounts and lengths of simple parametric reverberation over eight emotional characteristics. The results show a remarkable consistency in listener rankings of the instruments for each of the different types of reverberation, with strong correlations ranging from 90 to 95%. These results indicate that the underlying instrument space does not change much with reverberation in terms of emotional characteristics, and that each instrument has a particular footprint of emotional characteristics. Among the tones we tested, the instruments cluster into two fairly distinctive groups: those where the positive energetic emotional characteristics are strong, and those where the low-arousal characteristics are strong. The saxophone is an outlier, and is somewhat strong for most emotional characteristics. In terms of applications, the relatively consistent rankings of emotional characteristics between the instruments certainly helps each instrument retain its identity in different reverberation environments, and suggests possible future work in instrument identification.

1. INTRODUCTION

Researchers have found that different instruments have different timbral and emotional characteristics. By changing the pitch, dynamics, and other aspects of the performance, the timbre and emotional characteristics also change (see [1] for a detailed summary). These characteristics are further modified by the performance environment - by the amount and length of reverberation in the space [2, 3, 4, 5], which smears the temporal and spectral envelopes and changes the emotional character of the sound. The same idea holds when artificial reverberation is added as a post-process. For example, concert hall reverberation can bring out emotional characteristics such as Mysterious or Heroic from the original studio recording, or the recording engineer and musicians might use a dry sound to emphasize its Comic character [1, 6].

While reverberation can strengthen or deemphasize particular emotional characteristics, does it change the underlying instrument space? In other words, when reverberation changes the emotional characteristics of the instruments, does it change them uniformly or some instruments more than others? If we compare the instruments in terms of the emotional characteristic Heroic for example and rank them, is the ranking about the same for different amounts and lengths of reverberation? These questions have not been investigated previously to our knowledge, even though they have some implications in timbre and music emotion research. This paper will address these issues by comparing eight sustaining instruments over eight emotional characteristics with different amounts and lengths of reverberation.

This work also has implications for music emotion research of single musical instrument tones. Most of the sample libraries contain tones with light reverberation, and there are only a limited number of anechoic samples available. Most timbre and music emotion studies of single instrument tones do not explicitly state whether the tones are anechoic or with light reverberation, and assume that it does not matter too much. Is this a safe assumption? If reverberation changes the emotional characteristics of instruments uniformly in about the same way, then the relative space of emotional characteristics between the instruments stays about the same with different reverberations. In this situation, we can use the numerous samples that have light reverberation to compare instruments in terms of their emotional characteristics and expect about the same relative characteristics if they had been recorded in an anechoic chamber or a hall with different reverberation. On the other hand, if the change of emotional characteristics is instrument-dependent with reverberation, then the situation is more complicated. It would indicate a strong dependence on the type of reverberation, and suggests the limited applicability of studies of single instrument tones only to tones with similar types of reverberation. In this case, it would also suggest the need for more anechoic sample libraries.

2. METHODOLOGY

2.1 Overview

For this investigation, we used a relatively simple parametric reverberation model to measure the emotional characteristics of instruments for two of the most important reverberation parameters: reverberation length and amount. Our test had listeners compare eight instrument tones over eight emotional categories for each type of reverberation. The basic stimuli consisted of eight sustained wind and bowed string instrument sounds without reverberation: bassoon (bs), clarinet (cl), flute (fl), horn (hn), oboe (ob), saxophone (sx), trumpet (tp), and violin (vn). They were obtained from the *University of Iowa Musical Instrument Samples* [7]. Since pitch, duration, and loudness are potential factors in emotional characteristics, these instrument sounds were further synthesized. They had fundamental frequencies close to Eb4 (311.1 Hz) and were analyzed using a phase-vocoder algorithm [8]. The sounds were resynthesized by additive synthesis at exactly 311.1 Hz, and equalized the total duration to 1.0s. The sounds were also equalized in loudness by manual adjustment.

In addition to the resynthesized anechoic sounds, we compared sounds with reverberation lengths of 1s and 2s. We used the reverberation generator provided by *Cool Edit* [9]. We uses 80% for the amount of reverberation corresponding to the back of the hall, and we approximated the front of the hall with 20%.

33 subjects with no hearing problem were hired to take the listening test. All subjects were fluent in English. They compared the stimuli in paired comparisons for eight emotional categories: Happy, Sad, Heroic, Scary, Comic, Shy, Romantic, and Mysterious. Some choices of emotional characteristics are fairly universal and occur in many previous studies roughly corresponding to the four quadrants of the Valence-Arousal plane [10]. For this study, we used the same categories we have used in our previous research on musical instruments [1, 6, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21].

During each trial, subjects heard a pair of instrument sounds from the same type of reverberation and were prompted to choose which more strongly aroused a given emotional category. Since each trial was a single paired comparison requiring minimal memory from the subjects, subjects did not need to remember all of the tones, just the two in each comparison. Each combination of two different instruments tones were presented for each of the five reverberation types and eight emotional categories, and the listening test totaled $C_2^8 \times 5 \times 8 = 1120$ trials. For each instrument, the overall trial presentation order was randomized. For the two sounds A and B, they heard AB where the order of A and B was random for each comparison (but if they heard AB, they did not hear BA later). Before the first trial, subjects read online definitions of the emotional categories from the *Cambridge Academic Content Dictionary* [22]. The listening test took about 2 hours.

3. RESULTS

Based on the listening test data, we derived scale values using the Bradley-Terry-Luce (BTL) statistical model [23, 24]. Figures 1 to 3 show the BTL scale values and the corresponding 95% confidence intervals for anechoic, small hall front, and large hall back respectively. For each graph, the BTL scale values for the eight instruments sum up to 1. The BTL value for each instrument is the probability that listeners will choose that instrument when considering a certain reverberation type and emotion category. For example, if all eight instruments were judged equally Happy, the BTL scale values would be 1/8 = 0.125.

Though there are certainly differences between Figures 1 - 3, overall they are remarkably similar to one another. For example, the trumpet was consistently ranked highest for Heroic, while the clarinet was ranked highest for Sad. Going further, the trumpet ranked the highest for all reverberation types for Happy, Heroic, and Comic, while the clarinet was highest for Sad, Shy, Romantic, and Mysterious (except a close second for Small Hall Front), and the flute and violin shared top-rankings for Scary. Heroic consistently had the widest range among all reverberation types, and Scary the narrowest.

We wanted to determine the number of times each instrument was significantly greater than the other seven instruments for each reverberation type and emotional characteristic. Paired t-tests were used to analyze the voting data. Table 1 shows the number of times each instrument was significantly greater than the other seven instruments over all five reverberation types for each emotional characteristic. The maximum possible value is 35 and the minimum possible value is 0. For example, for Heroic the trumpet was statistically significantly greater than all the other seven instruments for four reverberation types and six for Large Hall Back, so its value is 34. The maximum value for each emotional characteristic is shown in bold and shaded. Table 1 makes it obvious that the trumpet was ranked the highest for Happy, Heroic, and Comic, the clarinet for Sad, Shy, Romantic, and Mysterious, and the flute for Scary.

We also wanted to determine how similar were the significant differences we found above and the BTL data for the different reverberation types. Therefore, we ran correlations for both of these as well as for the voting data. In all cases, the correlations were statistically significant (at the $p < 0.0001$ level) and very strong, ranging from 90 to 95%, indicating a near-linear relationship and a very high level of agreement. In particular, Table 2 shows Pearson and Spearman correlation between the different reverberation types based on the voting data, since it is the most precise and direct measure of correlation in the sense that it is correlation of the original data and not correlation of statistics based on the original data.

Let's take a look at the question of the consistency of the listeners during this long 2-hour listening test. As further evidence that the 33 subjects were giving sincere and attentive responses, if they had been giving random responses at the end of the test due to fatigue, it would have decreased the number of significant differences in Table 1, making the footprints less clear and less consistent. As it turned out, they were very consistent, suggesting listeners remained reasonably attentive. We don't claim that they were perfect, but the 90 - 95% correlation in Table 2 indicates that listeners were amazingly consistent.

4. DISCUSSION

Previous work has shown that reverberation can greatly change these characteristics [1, 3, 4, 6]. And while these emotional characteristics can be greatly changed with reverberation, the results in this paper have shown that they are changed uniformly in about the same way for different instruments. In other words, the underlying instrument space does not change much with reverberation in terms of

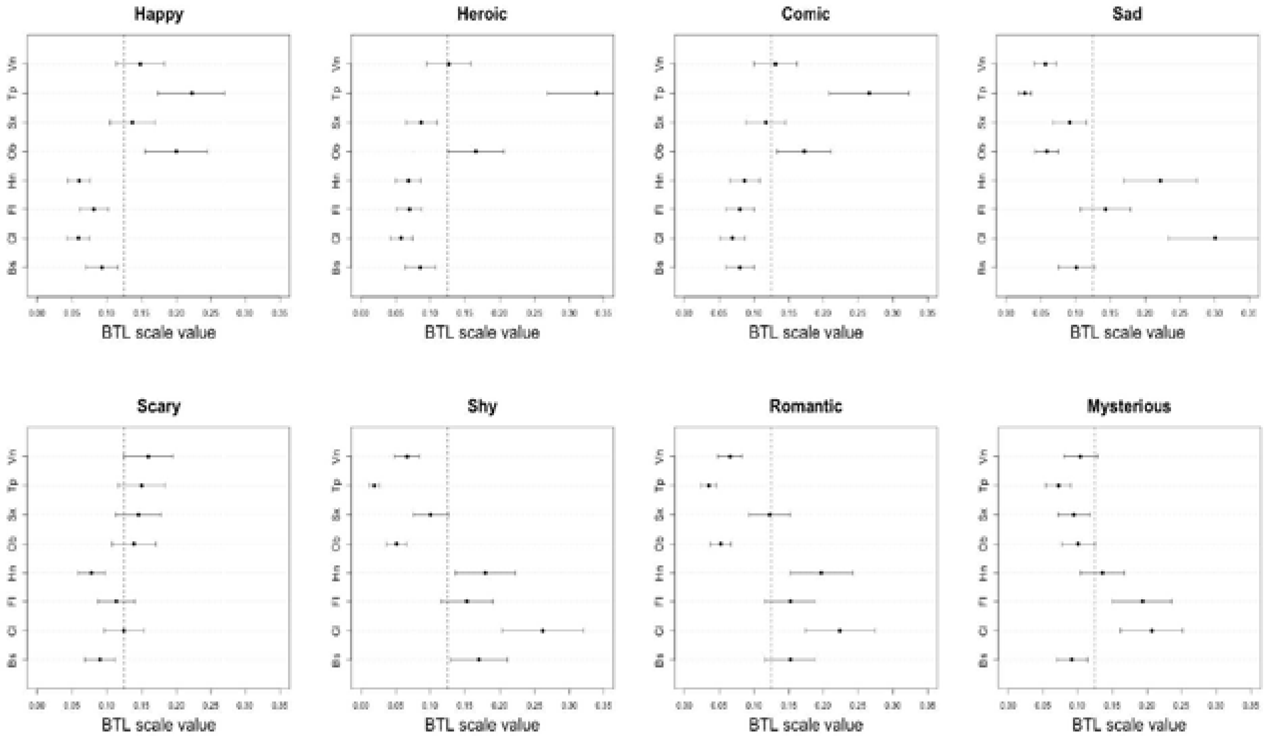

Figure 1. BTL scale values and the corresponding 95% confidence intervals of the anechoic instrument sounds for each emotional characteristic.

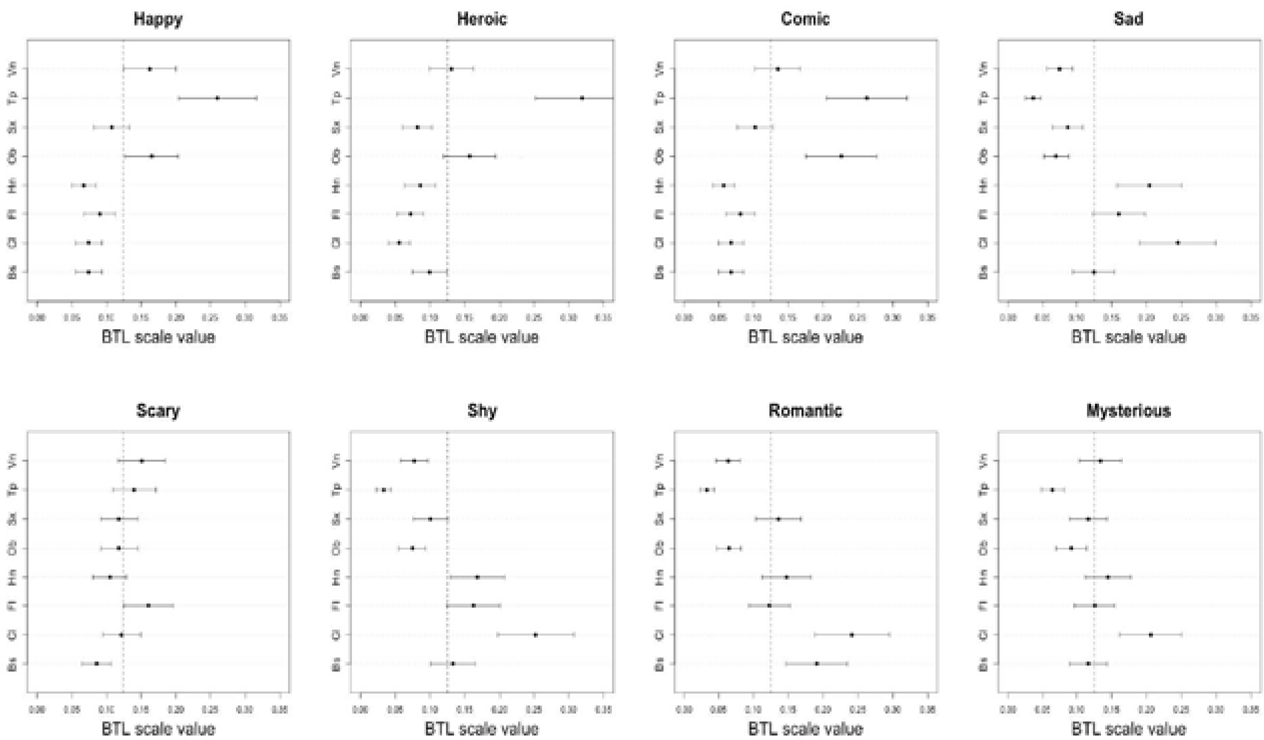

Figure 2. BTL scale values and the corresponding 95% confidence intervals of the instrument sounds with Small Hall Front reverberation for each emotional characteristic.

emotional characteristics. For example, added reverberation might bring out characteristics such as Mysterious or Heroic, but in a uniform way for the instruments, and not some more than others. There seems to be a relatively consistent ranking of emotional characteristics between the instruments that holds with different reverberation amounts and lengths, at least for simple parametric reverberation.

This uniformity is contrasting to our previous study [1, 6], where distinct and significant changes occurred in every instrument and emotional characteristic with different types of reverberation. The strong distinct changes found in our first study led us to expect some instrument-dependencies

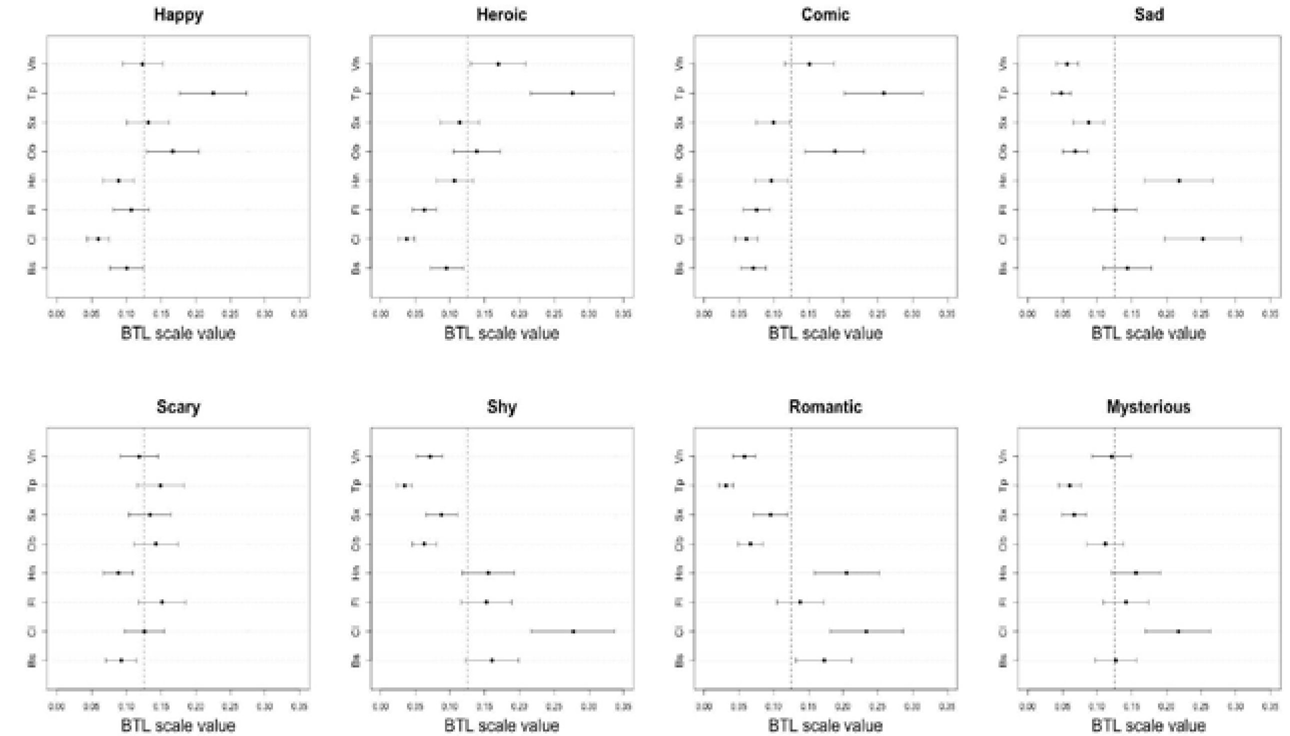

Figure 3. BTL scale values and the corresponding 95% confidence intervals of the instrument sounds with Large Hall Back reverberation for each emotional characteristic.

in this study, which used exactly the same tones. But, the two studies are from contrasting perspectives. In our first study, tones with different types of reverberation were compared for each instrument and emotional characteristic, allowing us to identify which reverberation types heightened each emotional characteristic for each instrument. In this study, tones from different instruments were compared for each reverberation type and emotional characteristic, allowing us to rank the instruments for each reverberation type and emotional characteristic. There is no contradiction in their results: reverberation distinctly changes the character of the sound, but does so in a uniform way across the instruments. It makes sense that reverberation changes the character uniformly across the instruments: if it were not uniform, then performers in orchestras and chamber groups would not be able to practice in small rehearsal rooms in a reliable way if reverberation affected the character in an instrument-dependent way. Musicians would need to carefully rehearse in the performance venue, not just to get used to the hall, but to adjust their blends and balances differently for each different venue.

	Bs	Cl	Fl	Hn	Ob	Sx	Tp	Vn
Happy	5	0	1	0	21	10	**26**	17
Heroic	6	0	2	4	21	5	**34**	12
Comic	1	0	2	1	25	8	**31**	18
Sad	17	**27**	20	26	4	10	0	4
Scary	0	1	**10**	0	1	2	1	3
Shy	16	**33**	21	16	5	9	0	5
Romantic	18	**25**	15	18	3	14	0	4
Mysterious	4	**26**	12	6	3	2	0	4

Table 1. How often each instrument was statistically significantly greater than the others over the five reverberation types.

Types	Pearson	Spearman
Anechoic & Small Hall Front	0.944	0.930
Anechoic & Small Hall Back	0.924	0.922
Anechoic & Large Hall Front	0.946	0.941
Anechoic & Large Hall Back	0.934	0.932
Small Hall Front & Small Hall Back	0.946	0.944
Small Hall Front & Large Hall Front	0.950	0.944
Small Hall Front & Large Hall Back	0.927	0.917
Small Hall Back & Large Hall Front	0.930	0.922
Small Hall Back & Large Hall Back	0.902	0.897
Large Hall Front & Large Hall Back	0.935	0.922

Table 2. Pearson and Spearman correlation between the different reverberation types based on the listener voting data.

More broadly, perhaps the relatively consistent ranking of emotional characteristics between the instruments is what allows each instrument to identify each instrument regardless of room reverberation, or at least helps. Perhaps each instrument has a characteristic footprint, that varies with pitch and dynamic level, which makes it identifiable.

In addition to the above results, where do the footprints appear in our data? The columns of Table 1 represent the overall footprints of the emotional characteristics for each instrument, that is, the instruments clustered into two fairly distinct groups: those where the positive energetic emotional characteristics were strong (e.g., oboe, trumpet, violin), and those where the low-arousal characteristics were strong (e.g., bassoon, clarinet, flute, horn). The saxophone was an outlier, and was uniquely somewhat strong for most emotional characteristics. Looking in more detail, the oboe, trumpet, and violin had similar footprints, but the trumpet's footprints were deeper for Happy, Heroic, and Comic than the other two instruments. In the same way, the clarinet and horn had similar footprints, though the clarinet was deeper especially for Shy and Mysterious. The flute also had a similar footprint to the clarinet and horn, but was deeper for Scary. The bassoon was similar to the horn except deeper for Happy, less for Sad. The saxophone had the most even distribution, with medium values for most emotional categories.

As a disclaimer, probably the footprint for each instrument varies depending on its pitch and dynamics as well as other factors of each particular tone. Anyway. The relatively consistent rankings of emotional characteristics between the instruments certainly helps explain why listeners can identify each instrument in different reverberation environments. This raises an interesting question about instrument identification: When listeners identify an instrument, are they identifying its unique sound, timbre, relative emotional characteristics, or a combination of these? This is a potential area for further work.

This work also has implications for music emotion research of single musical instrument tones, where most studies do not explicitly state whether the tones are anechoic or with light reverberation, and assume it does not matter too much. The results of this study suggest that this is a somewhat safe assumption if the relative emotional characteristics between instruments are the main consideration. We can therefore use the numerous samples with light reverberation to compare instruments in terms of their emotional characteristics and expect about the same relative characteristics if they had been recorded in an anechoic chamber or a hall with different reverberation. Of course in other situations it really can make a difference. Since reverberation smears the temporal and spectral envelopes, it changes the timbre of the sound. Similarly, reverberation can greatly change the emotional characteristics of the sound. If changes in timbre or absolute emotional characteristics are the main consideration of the study, reverberation can indeed make a difference, and should be handled with caution and appropriate disclaimers should be included. In any case, it is useful to know which situations are relatively safe and which can be problematic.

Another great area for further work would be in the parameterization of the temporal and spectral envelope smearing of reverberation. With different amounts and lengths of reverberation, how much change can we expect in the temporal and spectral envelopes? Will it be uniform among different instruments as we found here, or instrument-dependent? To our knowledge, the temporal and spectral envelope smearing effects have not been parameterized in detail.

5. REFERENCES

[1] R. Mo, B. Wu, and A. Horner, "The Effects of Reverberation on the Emotional Characteristics of Musical Instruments," *J. Audio Eng. Soc*, vol. 63, no. 12, pp. 966–979, 2016. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=18055

[2] L. Cremer, H. A. Müller, and T. J. Schaultz, *Principles and Applications of Room Acoustics*. Applied Science New York, 1982, vol. 1.

[3] D. Västfjäll, P. Larsson, and M. Kleiner, "Emotion and auditory virtual environments: affect-based judgments of music reproduced with virtual reverberation times," *CyberPsychology & Behavior*, vol. 5, no. 1, pp. 19–32, 2002.

[4] A. Tajadura-Jiménez, P. Larsson, A. Väljamäe, D. Västfjäll, and M. Kleiner, "When room size matters: acoustic influences on emotional responses to sounds," *Emotion*, vol. 10, no. 3, pp. 416–422, 2010.

[5] F. Rumsey, "Reverberation and How to Remove It," *J. Audio Eng. Soc*, vol. 64, no. 4, pp. 262–266, 2016. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=18133

[6] R. Mo, B. Wu, and A. Horner, "The Effects of Reverberation Time and Amount on the Emotional Characteristics," in *42nd International Computer Music Conference, Utrecht, The Netherlands*, 2016, p. 12.

[7] "University of Iowa Musical Instrument Samples," *University of Iowa*, 2004, http://theremin.music.uiowa.edu/MIS.html.

[8] J. W. Beauchamp, "Analysis and synthesis of musical instrument sounds," in *Analysis, Synthesis, and Perception of musical sounds*. Springer, 2007, pp. 1–89.

[9] "Cool Edit," *Adobe Systems*, 2000, https://creative.adobe.com/products/audition.

[10] P. N. Juslin and J. Sloboda, *Handbook of music and emotion: Theory, research, applications*. Oxford University Press, 1993.

[11] B. Wu, C. Wun, C. Lee, and A. Horner, "Investigating Correlation between Musical Timbres and Emotions," in *International Society for Music Information Retrieval Conference (ISMIR), Curitiba, Brazil*, 2013, pp. 415–420.

[12] B. Wu, C. Lee, and A. Horner, "The Correspondence of Music Emotion and Timbre in Sustained Musical Instrument Tones," *Journal of the Audio Engineering Society*, vol. 62, no. 10, pp. 663–675, 2014.

[13] B. Wu, A. Horner, and C. Lee, "Emotional Predisposition of Musical Instrument Timbres with Static Spectra," in *International Society for Music Information Retrieval Conference (ISMIR), Taipei, Taiwan*, vol. 253-258, Nov 2014.

[14] ——, "Musical Timbre and Emotion: The Identification of Salient Timbral Features in Sustained Musical Instrument Tones Equalized in Attack Time and Spectral Centroid," in *International Computer Music Conference (ICMC), Athens, Greece*, 14-20 Sept 2014, pp. 928–934.

[15] C.-j. Chau, B. Wu, and A. Horner, "Timbre Features and Music Emotion in Plucked String, Mallet Percussion, and Keyboard Tones," in *International Computer Music Conference (ICMC), Athens, Greece*, 14-20 Sept 2014, pp. 982–989.

[16] ——, “The Emotional Characteristics and Timbre of Nonsustaining Instrument Sounds,” *Journal of the Audio Engineering Society*, vol. 63, no. 4, pp. 228–244, 2015.

[17] C.-j. Chau and A. Horner, “The Effect of Pitch and Dynamics on the Emotional Characteristics of Piano Sounds,” 2015.

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

[19] C.-j. Chau, R. Mo, and A. Horner, “The Emotional Characteristics of Piano Sounds with Different Pitch and Dynamics,” *Journal of the Audio Engineering Society*, vol. 64, no. 11, pp. 918–932, 2016.

[20] R. Mo, G. L. Choi, C. Lee, and A. Horner, “The Effects of MP3 Compression on Emotional Characteristics,” in *42nd International Computer Music Conference, Utrecht, The Netherlands*, 2016, p. 411.

[21] ——, “The Effects of MP3 Compression on Perceived Emotional Characteristics in Musical Instruments,” *Journal of the Audio Engineering Society*, vol. 64, no. 11, pp. 858–867, 2016.

[22] Cambridge Academic Content Dictionary. [Online]. Available: http://dictionary.cambridge.org/dictionary/american-english

[23] R. A. Bradley, “14 Paired comparisons: Some basic procedures and examples,” *Nonparametric Methods*, vol. 4, pp. 299–326, 1984.

[24] 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.