Volume 52 http://acousticalsociety.org/

Acoustics 2023 Sydney
185th Meeting of the Acoustical Society of America

Sydney, Australia

4-8 December 2023

Musical Acoustics: Paper 3aMU9

A comparative analysis of violin and erhu: differences and similarities through statistical analysis of multiple musical excerpts

Wenyi Song, Zeyu Huang and Andrew Brian Horner

Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, School of Engineering, Hong Kong, HONG KONG; wsongak@cse.ust.hk; zhuangbi@connect.ust.hk; horner@cse.ust.hk

Recent work has compared the different emotional characteristics between violin and erhu on the Butterfly Lovers Concerto. In this study, we investigate whether the previous studies' results hold generally in Chinese and Western classical music. Building upon previous research, we hypothesize that the violin conveys more positive emotions, while the erhu is perceived as sadder. We also expect the violin to be better at conveying high-arousal excerpts. To test these hypotheses, 46 subjects were presented with 14 excerpts, with each excerpt represented by at least four different performances of both instruments. For each performance, subjects were asked 4 binary questions: whether it was happy, sad, agitated, and calm. Results show that the erhu is consistently perceived as sadder than the violin, while the violin is perceived as happier, calmer, and more agitated when significant differences appear. Linear regression analysis suggests that instrument is a more significant factor in the emotional perception of low-arousal excerpts than high-arousal ones. Additionally, performance was a less important factor than instrument on affect perception, and made more of a difference on low-arousal than high arousal excerpts. Meaning, there was more variation between different performances of the same except on low-arousal excerpts than high-arousal excerpts.

Published by the Acoustical Society of America

 https://doi.org/10.1121/2.0001817

1. INTRODUCTION

Music emotion studies have garnered significant interest in the field of psychology over the past few decades. These studies aim to investigate the intricate relationship between human affect and music, particularly how various musical elements influence emotional responses. Previous research has predominantly focused on examining the structural features of music, including tempo, mode, loudness, melody, rhythm, and timbre, to understand their impact on emotional perception [1, 2]. Notably, timbre has emerged as a crucial element in shaping the emotional characteristics of musical instruments [3].

Several studies have explored the connection between the timbre of musical instruments and different emotional attributes. These studies have investigated the effects of diverse instruments, pitches, and dynamics on emotional perception [4–17]. While Western instruments have been the primary focus of research, there has been a growing interest in Chinese traditional instruments [18–23]. However, researchers in this area still have limited access to comprehensive datasets [24–28].

This paper aims to contribute to the existing body of knowledge by conducting a comparative analysis of the emotional expression exhibited by the erhu and violin—two popular Chinese and Western bowed-string instruments, respectively. We strive to provide a comprehensive examination of the emotional characteristics of these instruments by assessing their performance quality at a higher level. Our study extends previous research by incorporating a diverse range of Chinese and Western musical pieces to determine the generalizability of the results across different pieces.

The remaining sections of this paper are organized as follows. Section 2 provides a review of related works in the field. Section 3 describes the experimental methodology employed to compare the emotional expression of the erhu and violin. Section 4 presents the results derived from our statistical analysis. Finally, Section 5 concludes the paper by discussing the implications of our findings and suggesting directions for future research.

2. RELATED WORK

The timbre of musical instruments plays a vital role in conveying emotional characteristics in music. Over the years, several studies have extensively investigated the connection between instrument timbre and various emotional dimensions [3–10]. Additionally, researchers have delved into the impact of different instruments, pitches, and dynamics on emotional perception [11–17]. This literature review presents a comprehensive overview of the existing research on the emotional characteristics of musical instruments, with a specific focus on Western and Chinese traditional instruments.

A. EMOTIONAL CHARACTERISTICS OF MUSICAL INSTRUMENTS

Numerous studies have examined the emotional attributes of single notes produced by Western classical instruments, including the piano, pitched percussion, woodwinds, brass, and bowed strings [11–17]. Within the string family, previous research has compared the emotional characteristics of the same note played at different octaves and dynamics on the violin, viola, cello, and double bass [13, 14].

While Western instruments have received considerable attention, there has been a growing interest in Chinese traditional instruments. For instance, Yang conducted a comparative study on the emotional expression of the violin and erhu, focusing on performance style vibrato in a Chinese music piece [18]. Lee conducted tests on *the Butterfly Lovers Concerto*, revealing distinct emotional conveyance between the erhu and violin [19, 20].

B. AVAILABILITY OF DATASETS

Despite the increasing interest in Chinese traditional instruments, researchers face a shortage of datasets to support their studies. While several Western bowed-string databases are available [29–37], fewer Chinese music datasets have been established. Notable Chinese datasets include the Chinese folk music-image dataset by Xing et al., the DCMI dataset by Zijin Li et al., and Zhang's JinYue and CCOM-HuQin datasets [24–28]. These datasets offer valuable resources for cross-media retrieval, regional representativeness, and emotion, scene, and imagery recognition in Chinese music.

C. CROSS-CULTURAL INSTRUMENTS AND EMOTIONAL PERCEPTION

With the surging popularity of Chinese music, researchers have begun exploring cross-cultural instruments and their impact on emotional perception [21,22,38–40]. Studies have collected data from participants from different countries, focusing on the cultural influences on music perception.

3. METHODOLOGY

This section outlines the methodology employed in the study to investigate the emotional characteristics of cross-cultural instruments, specifically the erhu and violin, from Chinese and Western cultures.

A. DATA COLLECTION

To address the limitations of previous research that relied on single pieces [19, 20], single performances [21], or MIDI files [22], our study adopted a comparative approach that included multiple performances of each excerpt. We selected well-known pieces commonly performed on both the erhu and violin, resulting in a total of 14 excerpts from 12 compositions. These excerpts encompassed a diverse range of performers in terms of gender, age, and nationality.

In this study, we adapted the emotional categorization scheme to include Happy, Sad, Agitated, and Calm in a more general way (As Figure 1).

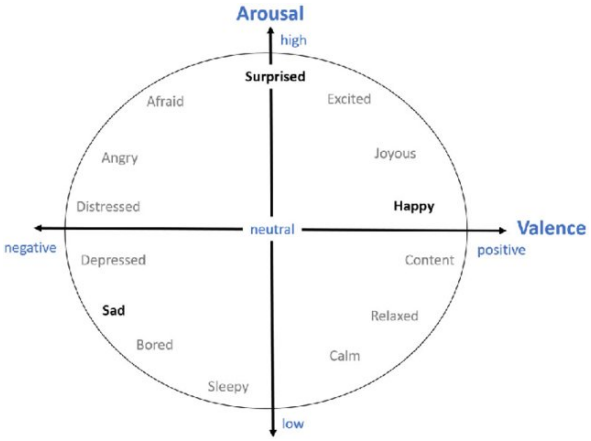

Figure 1: The 2D valence-arousal model of emotion proposed by Russel [41].

B. HYPOTHESES

Before conducting the listening test, two hypotheses were formulated to examine the generalizability of previous findings on the erhu and violin:

Hypothesis 1 *The violin conveys more positive emotions, while the erhu is perceived as sadder.*

Hypothesis 2 *The violin is better at conveying high-arousal excerpts (i.e., the violin is more Happy and more Agitated than the erhu).*

C. LISTENING TEST

The listening test was designed as an absolute comparison to evaluate the emotional characteristics of the erhu and violin. Participants listened to all 146 excerpts, which were presented in a randomized order. After listening to each excerpt, participants selected multi-choice answers from the four emotional categories (Happy, Sad, Agitated, and Calm) to indicate the emotions they perceived in the excerpt.

i. Stimuli

The 14 excerpts with in total 146 performances (In Table 1) were carefully selected from well-known compositions commonly played by both the erhu and violin. The excerpts were sourced from commercial CD recordings and the Internet, and minor adjustments were made to ensure consistency in volume and prevent distracting pops.

Table 1: Selected Excerpts Table

Piece Name	Category	No. of Performances(Erhu/Violin)
Erquan	Chinese	5/5
Zigeunerweisen, Op. 20 - part1	Western	5/5
Zigeunerweisen, Op. 20 - part2	Western	5/5
Zigeunerweisen, Op. 20 - part3	Western	5/5
Introduction and Rondo Capriccioso	Western	5/5
Serenade/Schubert	Western	6/6
Spring in Xinjiang	Chinese	4/4
Hora Staccato	Western	5/5
Czardas	Western	5/5
Singing the Night Among Fishing Boats	Chinese	5/5
Flight Of The Bumblebee	Western	6/6
The Sun shines on Tashikuergan	Chinese	5/5
Wonderful Night	Chinese	5/5
Meditation	Western	7/7

ii. Test Procedure

The listening test was conducted online using the Qualtrics survey system. Participants received brief instructions on how to run the test on their computers and familiarized themselves with the emotional categories (See Figure 2). They were required to use high-quality wired headsets and listen in a quiet environment with minimal background noise. The test took approximately 30 minutes to complete, and participants had the option to listen to each excerpt more than once.

Upon completing all 146 questions, participants were also required to submit video recordings for an honesty check.

Figure 2: Online Test for Example Question

iii. Participants

The subjects in this study were students enrolled in the Computer Music course at HKUST, who volunteered to participate in the listening test as part of their course requirements. A total of 46 participants took part in the test, including 36 males and 10 females, with an average age of 21.5 years. The participants were fluent in English and had no hearing issues. Out of the 46 participants, 41 were from China (including Hong Kong), and 30 had musical training experience ranging from 0.5 to 16 years.

To assess participants' familiarity with the erhu and violin, a single-choice question was included, as shown in Figure 3, to determine whether participants had prior experience listening to or playing either instrument. Out of the 46 participants, 14 indicated that they were familiar with both instruments, 14 were familiar with the violin, 1 was familiar with the erhu, and 1 reported playing both instruments.

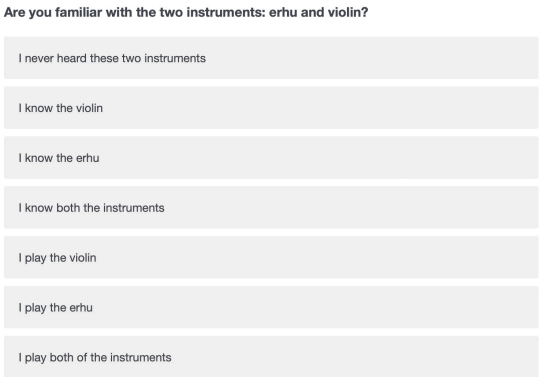

Figure 3: Online Test for Familiarity Question

D. DATA ANALYSIS

The data analysis involved several steps. First, the emotional distribution of each performance was calculated using counting percentages and summed it all up by excerpt. Linear regression was then employed to compare the emotional characteristics of the erhu and violin. The analysis focused on the emotional intensity of the four categories (Happy, Sad, Agitated, and Calm) and their combinations. The emotional intensity was calculated as the percentage of listeners who selected each emotional category for each excerpt. Linear regression models were used, with the instrument (erhu or violin) as a factor and the emotional intensity as the dependent variable.

4. RESULTS

In this study, we conducted an in-depth analysis to investigate the impact of instruments on the emotional response of listeners to Chinese and Western classical music excerpts.

A. FREQUENCY ANALYSIS AND EMOTIONAL CATEGORIES

To begin, we summarized the results of the listening test by creating frequency tables for each performance. We calculated the percentage of each emotion category for every performance as Figure 4 shows (*Spring in Xinjiang* as an example).

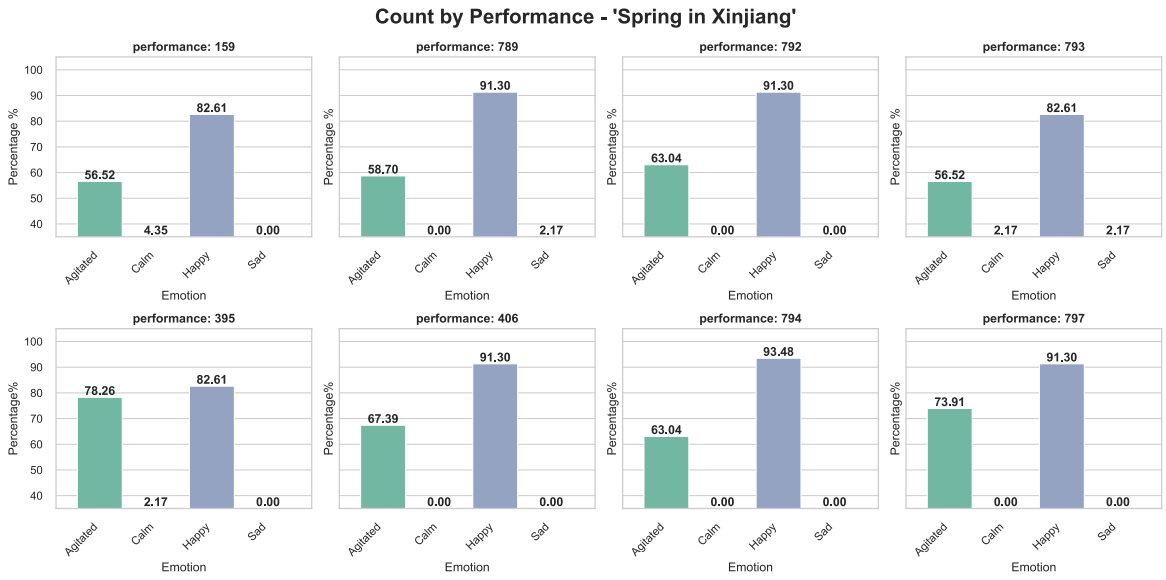

Figure 4: Example Excerpt of Counting Percentage by Performance

Notably, we observed that most excerpts elicited the top two represented emotions across both instruments. Additionally, the high/low arousal pattern appeared to be similar for both the erhu and violin within the same excerpt.

We aggregated the pecrformance data and analyzed the distribution of emotions in each excerpt, as shown in Figure 5. The results revealed a consistent trend: for the majority of excerpts, the erhu had a higher percentage in the Sad emotional category (i.e. *Zigeunerweisen, Op. 20 - 1, Zigeunerweisen, Op. 20 - 2, Zigeunerweisen, Op. 20 - 3, Serenade/Schubert, Meditation, Wonderful Night, Meditation*), while

the violin had a higher percentage in the other three emotions. This observation suggests that the choice of instrument influences the emotional expression in specific categories.

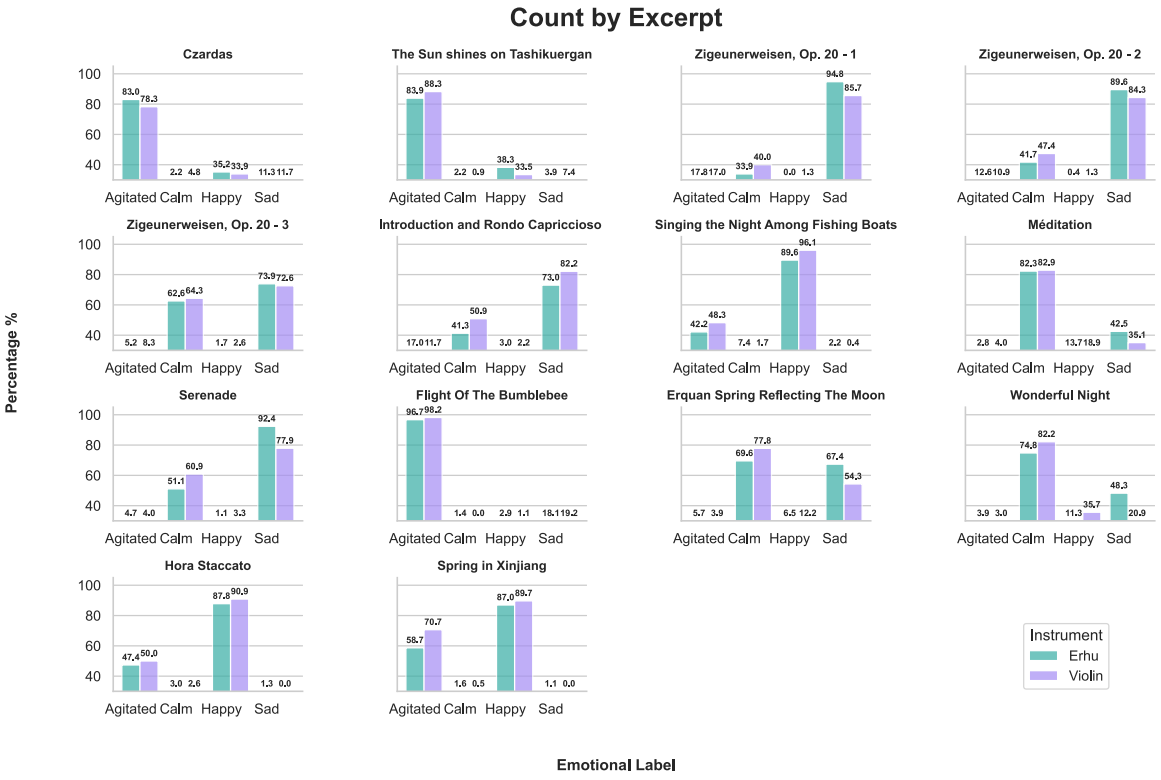

Figure 5: Aggregated Counting Percentage by Excerpt

Furthermore, we noted that the top two emotions consistently emerged as the dominant categories for each excerpt. This finding indicates that these primary emotions play a pivotal role in determining whether an excerpt elicits a high or low arousal response.

B. EXCERPT'S AROUSAL DETERMINATION

We proceeded to determine the arousal level of each excerpt based on the following rule: *if the Count (Happy + Agitated) is greater than the Count (Sad + Calm), the excerpt is classified as high arousal; otherwise, it is low arousal.* By applying this rule, we identified six high-arousal excerpts (See Figure 6): *Zigeunerweisen, Op. 20, Introduction and Rondo Capriccioso, Serenade/Schubert, Erquan, Wonderful Night*, and *Meditation*. Furthermore, we identified eight low-arousal excerpts: *Spring in Xinjiang, Hora Staccato, Czardas, Singing the Night Among Fishing Boats, Flight Of The Bumblebee*, and *The Sun shines on Tashikuergan*.

Interestingly, we found that the instrument factor had a notable impact on the emotional distribution of low-arousal excerpts, with significant differences observed in the percentages of Sad or/and Calm between the erhu and violin. In contrast, the percentages of Happy or/and Agitated did not exhibit significant differences between the two instruments for high-arousal excerpts.

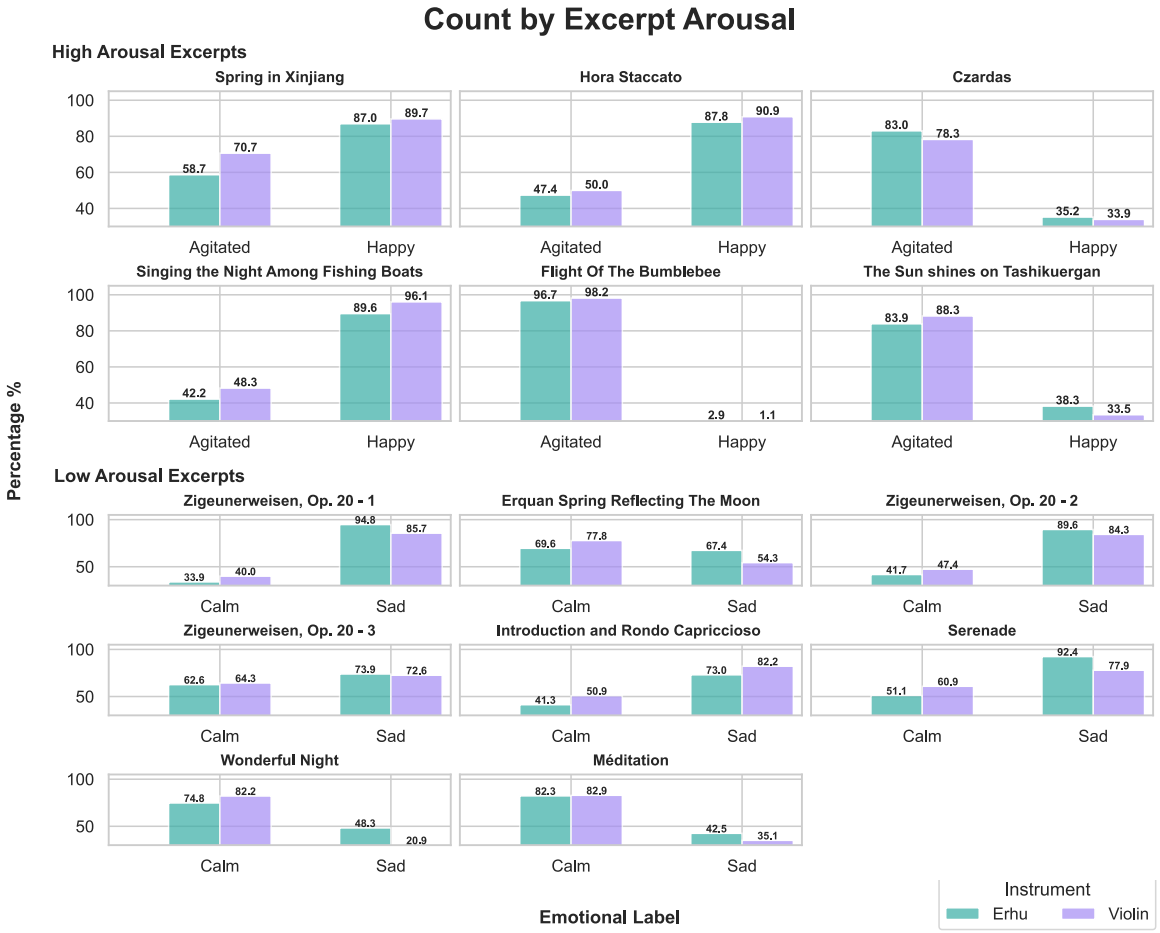

Figure 6: Counting Percentage by Arousal

C. LINEAR REGRESSION ANALYSIS

To further investigate the instrument factor's influence on emotional arousal, we conducted linear regression analysis.

Upon examining the coefficient values (See Figure 7, Table 2 and 3), we observed that the instrument factor showed significant differences across excerpts in terms of specific emotions. For low-arousal excerpts, the instrument factor played a crucial role, particularly in influencing emotions such as Sad and Calm. In contrast, for high-arousal excerpts, only two exceptions exhibited significant differences with respect to Happy and Agitated emotions.

These findings demonstrate that the instrument factor has a more pronounced impact on emotional responses in low-arousal excerpts, while its influence in high-arousal excerpts is relatively limited.

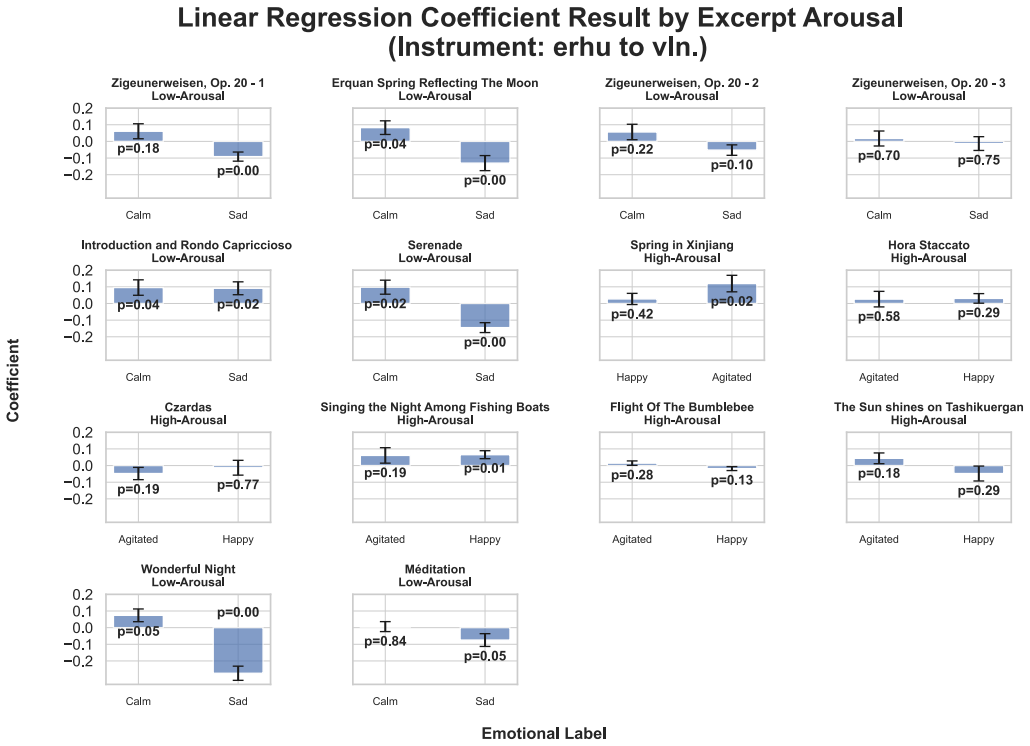

Figure 7: Instrument Factor - Linear Regression Result (by excerpt arousal)

D. SPECIAL CASE: *WONDERFUL NIGHT*

We encountered a special case in *Wonderful Night*, where the top three emotions displayed significant differences concerning the instrument factor. This suggests that factors other than the instrument, such as performance, might contribute to the observed emotional responses. Further experiments are necessary to validate this hypothesis.

Table 2: LR Significant Result (by emotion)

Piece Name	Arousal	Emotion	Coef.	F	FP	Std. Error
Spring in Xinjiang	High	Agitated	0.12	5.82	0.02	0.05
Erquan	Low	Calm	0.08	4.07	0.04	0.04
Introduction and Rondo Capriccioso	Low	Calm	0.1	4.26	0.04	0.05
Serenade/Schubert	Low	Calm	0.1	5.39	0.02	0.04
Singing the Night Among Fishing Boats	High	Happy	0.07	7.43	0.01	0.02
Zigeunerweisen, Op. 20 - 1	Low	Sad	-0.09	11.08	0	0.03
Erquan	Low	Sad	-0.13	8.33	0	0.05
Introduction and Rondo Capriccioso	Low	Sad	0.09	5.56	0.02	0.04
Serenade/Schubert	Low	Sad	-0.14	23.82	0	0.03
Wonderful Night	Low	Sad	-0.27	41.42	0	0.04

Table 3: LR Result by Coefficient (DES)

Piece Name	Arousal	Emotion	Coef.	F	FP	Std. Error
Wonderful Night	Low	Sad	-0.27	41.42	0	0.04
Serenade/Schubert	Low	Sad	-0.14	23.82	0	0.03
Erquan	Low	Sad	-0.13	8.33	0	0.05
Spring in Xinjiang	High	Agitated	0.12	5.82	0.02	0.05
Serenade/Schubert	Low	Calm	0.1	5.39	0.02	0.04
Introduction and Rondo Capriccioso	Low	Calm	0.1	4.26	0.04	0.05
		Sad	0.09	5.56	0.02	0.04
Zigeunerweisen, Op. 20 - 1	Low	Sad	-0.09	11.08	0	0.03
Erquan	Low	Calm	0.08	4.07	0.04	0.04
Singing the Night Among Fishing Boats	High	Happy	0.07	7.43	0.01	0.02

5. CONCLUSIONS AND DISCUSSIONS

In this study, our objective was to investigate the emotional response of listeners to Chinese and Western classical music excerpts performed on two distinct instruments: the erhu and the violin. Through a comprehensive listening test, we examined the impact of both the instrument and performance factors on listeners' emotional experiences.

Our hypotheses regarding the emotional perception of the erhu and violin were supported by the results, revealing statistically significant differences. The erhu consistently evoked a sadder emotional response, while the violin elicited perceptions of happiness, calmness, and agitation.

Notably, we found that the instrument factor played a more significant role in shaping emotional responses for low-arousal excerpts, with noteworthy differences observed in most cases. However, for high-arousal excerpts, the instrument factor did not exhibit the same level of significance, suggesting the influence of other factors in determining emotional experiences.

These findings underscore the complexity of emotional responses to classical music, which are influenced by multiple factors, including the instrument and performance aspects. Further investigations are warranted to explore the impact of additional factors such as cultural background and personal experiences, which may contribute to a more comprehensive understanding of emotional responses to music.

In conclusion, this study provides valuable insights into the emotional responses of listeners to Chinese and Western classical music performed on different instruments. It emphasizes the importance of considering multiple factors when studying emotional experiences in music and opens up new avenues for future research in this field.

REFERENCES

[1] K. R. Scherer and M. R. Zentner, "Emotional effects of music: Production rules." 2001.

[2] W. L. Sin, B. Y. Chang, X. Ma, and A. Horner, "The acoustics features and their relationship to the emotional characteristics of rain sound effects," Ph.D. dissertation, Hong Kong University of Science and Technology, 2019.

[3] 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: An Interdisciplinary Journal*, vol. 30, no. 1, pp. 49–70, 2012.

[4] B. Wu, S. Wun, C. Lee, and A. Horner, "Spectral correlates in emotion labeling of sustained musical instrument tones." in *ISMIR*, 2013, pp. 415–420.

[5] 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," in *ICMC*, 2014.

[6] ——, "Emotional predisposition of musical instrument timbres with static spectra." in *ISMIR*, 2014, pp. 253–258.

[7] ——, "The correspondence of music emotion and timbre in sustained musical instrument sounds," *Journal of the Audio Engineering Society*, vol. 62, no. 10, pp. 663–675, 2014.

[8] C.-j. Chau, B. Wu, and A. Horner, "Timbre features and music emotion in plucked string, mallet percussion, and keyboard tones," in *ICMC*, 2014.

[9] ——, "The emotional characteristics and timbre of nonsustaining instrument sounds," *Journal of the Audio Engineering Society*, vol. 63, no. 4, pp. 228–244, 2015.

[10] ——, "The effects of early-release on emotion characteristics and timbre in non-sustaining musical instrument tones," in *ICMC*, 2015.

[11] C.-j. Chau and A. Horner, "The effects of pitch and dynamics on the emotional characteristics of piano sounds," in *ICMC*, 2015.

[12] 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.

[13] S. J. Gilburt, C.-j. Chau, and A. Horner, "The effects of pitch and dynamics on the emotional characteristics of bowed string instruments," in *International Computer Music Conference (ICMC), Utrecht*, 2016, pp. 405–410.

[14] C.-j. Chau, S. J. Gilburt, R. Mo, and A. Horner, "The emotional characteristics of bowed string instruments with different pitch and dynamics," *Journal of the Audio Engineering Society*, vol. 65, no. 7/8, pp. 573–588, 2017.

[15] C.-j. Chau and A. Horner, "The emotional characteristics of mallet percussion instruments with different pitches and mallet hardness," in *Proceedings of the International Computer Music Conference*, 2016, p. 401.

[16] H. T. Chan, B. Y. Chang, and A. Horner, "The emotional characteristics of woodwind musical instruments with different pitch and dynamics," in *Proc. Int. Comp. Music Conf.(ICMC)*, 2019.

[17] H. T. Chan, R. Mo, C. Keyes, and A. Horner, "The emotional characteristics of brass musical instruments with different pitch and dynamics," in *Proc. Int. Comp. Music Conf.(ICMC)*, 2019.

[18] L. Yang, E. Chew, S.-K. Rajab *et al.*, "Vibrato performance style: A case study comparing erhu and violin," 2013.

[19] D. Lee, "A comparison of the emotional characteristics of the violin and erhu on the butterfly lovers concerto," Ph.D. dissertation, The Hong Kong University of Science and Technology, 2020.

[20] D. Lee, W. Song, and A. B. Horner, "A head-to-head comparison of the emotional characteristics of the violin and erhu on the butterfly lovers concerto," 2020.

[21] X. Wang, Y. Wei, L. Heng, and S. McAdams, "A cross-cultural analysis of the influence of timbre on affect perception in western classical music and chinese music traditions," *Frontiers in Psychology*, vol. 12, p. 732865, 2021.

[22] J. Liu, S. Wang, Y. Xiang, J. Jiang, Y. Jiang, and J. Lan, "Comparison and analysis of timbre fusion for chinese and western musical instruments," *Frontiers in Psychology*, vol. 13, p. 878581, 2022.

[23] W. Song and A. B. Horner, "Uncovering the differences between the violin and erhu musical instruments by statistical analysis of multiple musical pieces," in *Proceedings of Meetings on Acoustics*, vol. 50, no. 1. AIP Publishing, 2022.

[24] B. Xing, K. Zhang, S. Sun, L. Zhang, Z. Gao, J. Wang, and S. Chen, "Emotion-driven chinese folk music-image retrieval based on de-svm," *Neurocomputing*, vol. 148, pp. 619–627, 2015.

[25] X. Liang, Z. Li, J. Liu, W. Li, J. Zhu, and B. Han, "Constructing a multimedia chinese musical instrument database," in *Proceedings of the 6th Conference on Sound and Music Technology (CSMT) Revised Selected Papers*. Springer, 2019, pp. 53–60.

[26] Z. Li, X. Liang, J. Liu, W. Li, J. Zhu, and B. Han, "Dcmi: A database of chinese musical instruments," 2018.

[27] K. Zhang, X. Wu, R. Tang, Q. Huang, C. Yang, and H. Zhang, "The jinyue database for huqin music emotion, scene and imagery recognition," in *2021 IEEE 4th International Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2021, pp. 314–319.

[28] Y. Zhang, Z. Zhou, X. Li, F. Yu, and M. Sun, "Ccom-huqin: an annotated multimodal chinese fiddle performance dataset," *arXiv preprint arXiv:2209.06496*, 2022.

[29] G. B. CarmineCella. (2019-2023) Sol. IRCAM. [Online]. Available: https://forum.ircam.fr/projects/detail/fullsol/

[30] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "Rwc music database: Popular, classical and jazz music databases." in *Ismir*, vol. 2, 2002, pp. 287–288.

[31] ——, "Rwc music database: Music genre database and musical instrument sound database," 2003.

[32] V. Lostanlen, J. Andén, and M. Lagrange, "Extended playing techniques: the next milestone in musical instrument recognition," in *Proceedings of the 5th international conference on digital libraries for musicology*, 2018, pp. 1–10.

[33] J.-F. Ducher and P. Esling, "Folded cqt rcnn for real-time recognition of instrument playing techniques," in *International Society for Music Information Retrieval*, 2019.

[34] H. von Coler, "Tu-note violin sample library–a database of violin sounds with segmentation ground truth," in *Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18), Aveiro, Portugal*, 2018, pp. 4–8.

[35] H. von Coler and A. Lerch, "Cmmsd: A data set for note-level segmentation of monophonic music," in *Audio Engineering Society Conference: 53rd International Conference: Semantic Audio*. Audio Engineering Society, 2014.

[36] K. Subramani and P. Rao, "Hprnet: Incorporating residual noise modeling for violin in a variational parametric synthesizer," *arXiv preprint arXiv:2008.08405*, 2020.

[37] A. Elovsson and O. Lartillot, "A hardanger fiddle dataset with performances spanning emotional expressions and annotations aligned using image registration," 2021.

[38] J. H. Lee and X. Hu, "Cross-cultural similarities and differences in music mood perception," *IConference 2014 Proceedings*, 2014.

[39] X. Hu and Y.-H. Yang, "Cross-dataset and cross-cultural music mood prediction: A case on western and chinese pop songs," *IEEE Transactions on Affective Computing*, vol. 8, no. 2, pp. 228–240, 2017.

[40] L. Heng, *Timbre in the Communication of Emotions Among Performers and Listeners from Western Art Music and Chinese Music Traditions*. McGill University (Canada), 2018.

[41] Y. Tsalamlal, M.-A. Amorim, J.-C. Martin, and M. Ammi, "Modeling emotional valence integration from voice and touch," *Frontiers in Psychology*, vol. 9, p. 1966, 2018.