

# ASPECTS OF EXPERIMENTAL DESIGN FOR THE PERCEPTUAL EVALUATION OF VIOLIN QUALITIES

Charalampos Saitis<sup>1</sup>, Bruno L. Giordano<sup>2</sup>, Claudia Fritz<sup>3</sup>, Gary P. Scavone<sup>1</sup>

<sup>1</sup>Computational Acoustic Modeling Laboratory, Centre for Interdisciplinary Research in Music Media and Technology, Schulich School of Music, McGill University, Montreal QC, Canada, H3A 1E3

<sup>2</sup>Music Perception and Cognition Laboratory, Centre for Interdisciplinary Research in Music Media and Technology, Schulich School of Music, McGill University, Montreal QC, Canada, H3A 1E3

<sup>3</sup>Lutheries-Acoustique-Musique, Institute Jean le Rond d'Alembert, Université Pierre et Marie Curie, UMR CNRS 7190, 4 place Jussieu, 75005 Paris, France

## 1. INTRODUCTION

In this paper, we present a method for the perceptual evaluation of violin qualities. In particular, we discuss the different experimental, musical and logistical issues concerning the design of reliable playing tests to explore the perceptual processes involved when the violinist interacts with the instrument in a musical context – for example, when comparing different violins in order to purchase a new one. Preliminary results are provided.

## 2. METHOD

### 2.1. Playing vs. listening tests

Psychoacoustic experiments based on listening tests using recordings, synthesized sounds or live performance have several disadvantages. Recorded sounds often lack the “naturalness” of live performance. Similarly, synthesized tones often sound rather unmusical [1]. And when using live players, listeners tend to focus more on the performer than the instrument: A “good” player is likely to make a “bad” violin sound “good” and vice versa. Most importantly, vibro-mechanical and tactile properties, such as “responsiveness,” cannot be completely assessed without direct interaction with the instruments [2]. Playing tests are therefore more ecologically valid than listening tests.

A potential problem with playing tests is how to evaluate sound radiation from the violin and timbre-related qualities such as “richness.” Radiation depends on the distance from the sound-producing source as well as the acoustics of the space in which the experiment is conducted. In the first case, the same individual might possibly assess the sound quality of the same instrument differently when in the role of player versus listener at a different location. In the second case, using a reverberant or acoustically dry room may affect how the sound is perceived. For our experiments, we used a rather dry room with a surface of  $27m^2$  and reverberation time of approximately 0.18s in order to help players focus more on the vibrational response of the violins than on their sound,

### 2.2. Visual information

Anecdotal evidence strongly suggests that some visual information, such as the color of the varnish, the grain of the wood, or identifying marks on the violin, may influence judgement. To help minimize the effects of such visual cues as much as possible in listening tests involving live performance, the listeners or the performers or both are usually blindfolded. Another approach is to have the instruments played behind a physical divider [3]. However, blindfolding does not appear to be a viable solution for playing tests. To circumvent the potential impact of visual information on judgment while ensuring a certain level of comfortability for the musicians, as well as safety for the instruments, we used low light conditions and asked participants to wear a pair of extra tinted sunglasses. In this way, violinists can provide objective assessments while still retaining some visual contact with the instruments.

### 2.3. Choice of bow

A critical issue when conducting violin playing tests is the choice of bow. In this study, two alternatives were considered: asking players to use their own bow (or the one they are most familiar with) or choosing a common bow across all participants (the option of instead using a common set of different bows across all players was discarded due to logistical constraints concerning the duration of the experimental session). Although neither option is ideal, using a common bow is likely to raise the very same quality question as with the violin [4]. Moreover, participants may feel uncomfortable with a bow they are not familiar with. We therefore preferred the first option. This was also felt to be more typical of how violinists assess instruments while in the process of purchasing one.

### 2.4. Violins

Two pilot studies with five and nine violins respectively were run to help optimize the number and type of instruments to evaluate so as to increase the generality of the results. Accordingly, eight violins of different make (Europe, Québec, China) and age (ranging from 1840 to 2010) were selected in coordination with local luthiers. The strings, bridge, and chinrest were optimally setup for each

violin by the luthiers prior to the experiment. All violins had identical shoulder rests (Kun Original model).

To avoid potential biases caused by the "mere exposure" effect by which familiarity with a stimulus object amplifies preference toward it [5], the player's own violin was not included in the evaluation set. Instead, participants were permitted to use their own violin as a reference point during the experiments.

### 2.5. Data quality

To obtain reliable results from quality assessments, it is important to consider the statistical validity of the experimental procedure. The number of players can be maximized to better estimate the extent of inter-subject differences. For the purposes of our experiment, twenty violinists with at least fifteen years of performance experience were selected. To increase accuracy of intra-subject consistency and individual-specific preference data, players were asked to repeat the evaluation task five times. To obtain a stronger testing of intra-subject consistency, violinists were asked to return for a second, identical session within a period of three to seven days later.

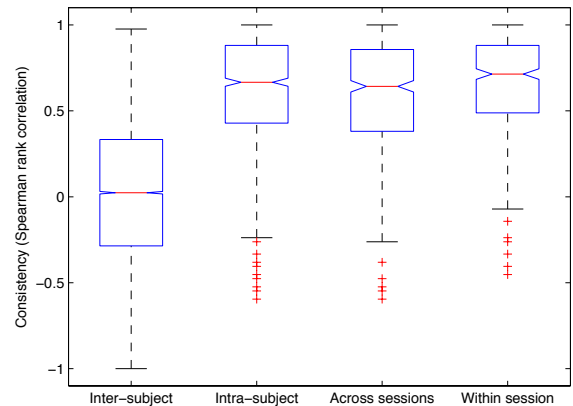
Maximizing task repetitions as much as possible is desirable, but there are logistical constraints such as the total duration of the experimental session or physical limitations such as fatigue that must also be considered. Although having many repetitions can help reduce the experimental noise in the data, fatigue may have the opposite effect.

## 3. RESULTS

Participants were asked to play the different violins and order them by preference. For ecological validity reasons, no constraints were imposed on the evaluation process (e.g. specific repertoire). Preference judgments were collected as a measure of subjective evaluation of the quality of a violin based on choice behavior [6]. The Spearman's rank correlation coefficient across rankings from different trials was used as a measure of intra- and inter-subject consistency. A Wilcoxon signed-rank test was then adopted to test (a) whether inter-individual agreement was significantly different from intra-individual consistency, and (b) whether intra-subject consistency between the two sessions was significantly different from intra-subject consistency within one session.

Figure 1 depicts the results for inter-individual agreement and intra-individual consistency across the two sessions, and intra-individual consistency between the two sessions and within one session in the form of box plots (from left to right respectively). Results indicate that violin players are relatively self-consistent when evaluating different instruments in terms of preference. However, a significant amount of disagreement between violinists is observed ( $p < 0.001$ ). Furthermore, participants appear more self-

consistent in the same session than across the two sessions ( $p = 0.0045$ ).



**Figure 1. Box plots of intra- and inter-subject variation: 1 corresponds to perfect consistency and -1 corresponds to perfect inconsistency (exactly opposite ranking between different trials); The red line is the median; Box brackets show the interquartile range; Red crosses depict outliers.**

## 4. CONCLUSIONS

We have presented a method for evaluating violins that is well-controlled by scientific standards. The musical and logistical difficulties related to the experimental design have been discussed. We applied the method to examine intra- and inter-subject variability in violin players across a certain set of violins using preference rankings. Experimental control is necessary to obtain reliable quantitative data, but a certain amount of musical context is required for such information to be meaningful. No single preference ranking can satisfactorily capture inter-individual differences. Preference rankings are still an important area of study in the perceptual evaluation of string instruments. Results suggest the need for further work to explain the large inter-individual variability in the preference for violins.

## REFERENCES

- [1] Wright, H. (1996). *The Acoustics and Psychoacoustics of the Guitar*, Ph.D. thesis, Department of Physics and Astronomy, University of Wales, Cardiff, UK
- [2] Fritz, C., Muslewski, A. and Dubois, D. (2010). "A situated and cognitive approach of violin quality," in Proc. International Symp. Musical Acoust., Sydney and Katoomba, Australia.
- [3] Petiot, J.-F. and Caussé, R. (2007). "Perceptual differences between cellos: A subjective/objective study," in Proc. International Symp. Musical Acoust., Barcelona, Spain.
- [4] Caussé, R., Maignet, J. P., Dichtel, C. and Bensoam, J. (2001). "Study of violin bow quality," in Proc. International Symp. Musical Acoust., Perugia, Italy.
- [5] Zajonc, R. B. (1968). "Attitudinal effects of mere exposure." *J. Pers. Social Psych.*, **9**, 1-27.
- [6] Giordano, B. L., Susini, P. and Bresin, R. (2011). "Experimental methods for the perceptual evaluation of sound-producing objects and interfaces," in *Sonic Interaction Design*, edited by K. Franinovic and S. Serafin, MIT Press, Boston, MA.