

## PERCEPTUAL DIFFERENCES BETWEEN CELLOS: A SUBJECTIVE/OBJECTIVE STUDY

*Jean-François PETIOT<sup>1)</sup>, René CAUSSE<sup>2)</sup>*

<sup>1)</sup> *Institut de Recherche en Communications et Cybernétique de Nantes (UMR CNRS 6597) - 1 rue de la Noë, BP 92101, 44321 Nantes Cedex 3 France,*

*petiot@ircsyn.ec-nantes.fr*

<sup>2)</sup> *Institut de Recherche et Coordination Acoustique Musique (UMR CNRS 9912) 1 place Igor Stravinsky, 75004 Paris France, Rene.Causse@ircam.fr*

### Abstract

This paper addresses the characterisation of cello sounds. In order to study to which extent a jury perceives differences between two cellos in playing conditions, we carried out two blindfolded hearing tests, involving two instruments and two professional musicians: (1) An evaluation test of the couple musician-instrument on a structured scale, for 5 different attributes defined by adjective-pairs. The assessment of the jury was based on the same musical sequences, played by the musicians; (2) A comparative test, based on the ranking of the couple musician-instrument on 4 different attributes. The same short musical fragment was played successively by the musicians. All the played musical sequences have been recorded, and metrics based on the acoustic signal (playing frequency, spectral centroid, signal/noise ratio) were calculated in order to interpret the perceived differences. The results show that for the evaluation test, the inter-subject differences of the jury are too large and do not allow the definition of a significant instrument effect. For the comparative test, the agreement between the subjects is better and significant differences between the instruments and the musicians can be observed, explained by the signal processing of the played sounds.

### INTRODUCTION

The study of the quality of musical instruments is particularly interesting to help their development and to improve their design. In the literature, we distinguish two kinds of studies which address this goal: (1) subjective studies, where the quality is assessed by listeners or players [Pratt & Bowsher, 1978] during evaluation tests; (2) objective studies, where the quality is evaluated by physical measurements on the instruments [Pratt & Bowsher, 1979]. Most of the time, links are proposed between these two studies, in order to explain *a posteriori* subjective attributes by physical measurements. In order to propose a model for predicting certain qualities of an instrument, the approach consists in discovering correlations, for a set of instruments, between the subjective response (given by the subject) and measurements (made on the signal of the sounds, or directly on the instrument itself) [Plitnik & al., 1999]. This approach has been used for example for the study of loudspeakers [Lavandier, 2005], the study of guitars [Wright, 1996] or trumpets [Poirson et al., 2007] and is classical in room acoustics.

The main difficulty with this approach is to get reliable subjective data. Indeed, the subjective assessments are most of the time dependent on cultural and training aspects of the subject, and subjected to inherent inter-individual differences. This work is in this context. On request of two professional musicians, we wanted to know if what can be called a “Stradivarius effect” is perceptible by listeners. In other words, our aim was to study in which extend a jury of subjects perceive consensual and/or reliable differences between two high-end cellos. More precisely, the objectives of this work are:

- To characterise, by the way of hearing tests with a jury, the perceived differences between two high-end cellos
- To try to explain the differences and to link them to attributes of the acoustical signal

In this paper, we present firstly the experimental protocol we designed for the assessments of the instruments. The different tests and the design of experiments are detailed. Secondly, the results of the subjective tests are presented. Finally, the main results of the signal analysis are given, in relation with the conclusions of the subjective tests.

## **EXPERIMENTAL PROTOCOL**

### **Jury and room**

The jury was made of 6 participants, each of them being involved in the musical acoustics or instrument making sector. The jury was blindfolded, all the instrument have been played behind a curtain (figure 1). The tests were carried out in the “ESPRO” room of the IRCAM, the duration of the session was approximately 2 hours.



*Figure 1 : picture of the assessment session with the jury*

### **Instruments and musicians**

Two professional musicians, denoted “X” and “Y”, played the instruments for the tests. Two instruments (their own instruments) have been considered for the study:

- Instrument “R”, French maker, modern, 2005
- Instrument “C”, Italian maker, ancient, Venice, 1680

### **Assessment tests**

Given that we wished to asses the instrument effect, the musician effect, and eventually the interaction musician\*instrument, we tested all the combinations

musician\*instrument, and we introduced repetition in the presentation. The full factorial design is made of  $2 \times 2 = 4$  cells (all the possible combinations). It has to be noticed that the musicians used their own bow whatever instrument they played. In collaboration with the two musicians, two listening tests were designed:

- Test n°1: evaluation test: assessment on a bipolar structured scale, for a set of 6 attributes concerning the sound and the radiation of the instrument
- Test n°2: comparative test : ranking of the instruments, for a set of 4 attributes

*Test n°1: evaluation test*

Six attributes were proposed for the characterisation of the instruments. Concerning the sound, 5 attributes, defined by a bipolar scale: *neutral-rich*; *not resonant-resonant*; *nasal-round*; *magnitude neutral-rich*; *timber homogeneity*. Concerning the radiation, 1 attribute: *narrow-large*.

In order to test the repeatability of the subjects, 3 repetitions of the configurations were introduced in the presentation. The factorial design proposed to the jury was finally made of  $2 \times 2 \times 3 = 12$  configurations. The stimuli of the test n°1 have been proposed by the musicians. It was a musical sequence of around 2 minutes, mixing various style and dynamics, repeated for all the configurations. For the evaluations, the musicians played the stimuli, and the jury assessed progressively the configuration during the listening. The assessment sheet for test N°1 is given figure 2. It is made of 6 structured scales on which the jury had to indicate his/her assessment.

<b>Epreuve 1 : cotation</b>				
<b>LE SON (timbre de l'instrument)</b>				
<b>neutre – riche</b>				
ni neutre ni riche				
Très neutre	Assez neutre		Assez riche	Très riche
1	2	3	4	5
<b>peu résonant – résonant</b>				
Très peu résonant	Assez peu résonant	ni résonant ni peu résonant	Assez résonant	Très résonant
1	2	3	4	5
<b>nasal – rond</b>				
ni nasal ni rond				
Très nasal	Assez nasal		Assez rond	Très rond
1	2	3	4	5
<b>Amplitude neutre – riche</b>				
faible	Assez faible	moyenne	Assez grande	Très grande
1	2	3	4	5
<b>Homogénéité du timbre</b>				
Très peu homogène	Assez peu homogène	ni / ni	Assez homogène	Très homogène
1	2	3	4	5
<b>LE RAYONNEMENT DU SON</b>				
<b>étroit – large</b>				
ni étroit ni large				
Très étroit	Assez étroit		Assez large	Très large
1	2	3	4	5

Figure 2 : assessment sheet for the test n°1

*Test n°2: comparative test*

Four attributes were proposed for the ranking of the instruments: *powerful-bright-full-directive*. In order to test the repeatability of the subjects, 2 repetitions of the configurations were introduced in the presentation. The factorial design proposed to the jury was finally made of  $2 \times 2 \times 2 = 8$  configurations.

The stimuli of the test n°2 have been proposed by the musicians. It consists of a short musical fragment of around 5 seconds, specific for each attribute. The two musicians were asked to play successively this fragment, and the jury had to rank the instruments just after, according to the attribute. The assessment sheet for the test n°2 is given figure 3.

<b>Epreuve 2 : classement</b>	
<b>Classer les instruments du plus puissant (1) au moins puissant (2)</b>	
Instrument A	Instrument B
<input style="width: 30px; height: 30px;" type="text"/>	<input style="width: 30px; height: 30px;" type="text"/>
<b>Classer les instruments du plus clair (1) au moins clair (2)</b>	
Instrument A	Instrument B
<input style="width: 30px; height: 30px;" type="text"/>	<input style="width: 30px; height: 30px;" type="text"/>
<b>Classer les instruments du plus ample (1) au moins ample (2)</b>	
Instrument A	Instrument B
<input style="width: 30px; height: 30px;" type="text"/>	<input style="width: 30px; height: 30px;" type="text"/>
<b>Classer les instruments du plus directif (1) au moins directif (2)</b>	
Instrument A	Instrument B
<input style="width: 30px; height: 30px;" type="text"/>	<input style="width: 30px; height: 30px;" type="text"/>

Figure 3 : answer sheet for the test n°2

### Recordings and signal processing

All the stimuli were recorded during the session. Specifically, the stimuli for the test n°1 comprised the open strings of the cello, C2, G2, D3, A3 (figure 4), played at a dynamic *mf*.



Figure 4 : notes of the cello recorded for each configuration, basis for the signal processing

For each configuration (musician\*instrument), three repetitions of each note were available. For each note, the playing frequency, the spectral centroid  $S_c$ , and the signal-noise ratio  $R_{sn}$  (ratio of the total intensity of the signal on the intensity of the non-harmonic part of the signal), were computed. The synchronous detection method was used to calculate the level of each harmonic of the note.

## RESULTS OF THE SUBJECTIVE ASSESSMENTS

### Test n°1: assessment tests

For all the attributes, the box-plot of the assessments of the two instruments “R” and “C” (whatever musician is playing) by the 6 subjects are given figure 5. These raw data don’t show a clear instrument effect (the average value of the evaluations, represented by the red cross, is rather similar for all the attributes, except for “timber homogeneity”, for which the instrument R seems to be more homogeneous than C). The inter-subjects difference and/or the error of repeatability of the subjects seem to be rather high.

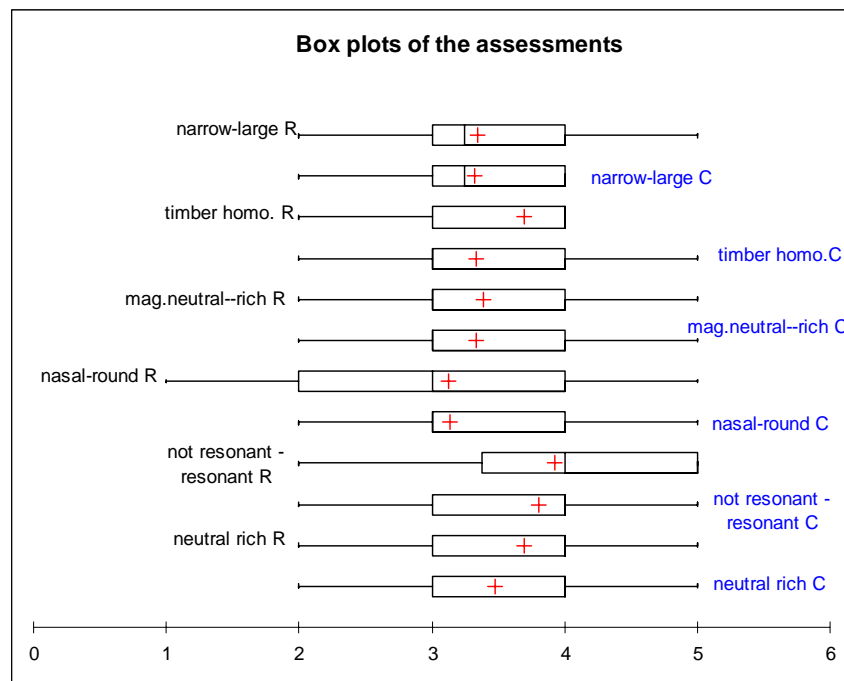


Figure 5 : Box-plots of the assessments for the test n°1.

A two-way ANOVA with interaction (factors *instrument* and *musician*) shows that the only significant effect ( $p$ -value $<5\%$ ) is the instrument effect for the attribute *timber homogeneity* ( $p$ -value = 4%). Table 1 gives the  $p$ -values of the ANOVA for all the factors and all the attributes.

<b>p-value</b>	<i>neutral-rich</i>	<i>not resonant - resonant</i>	<i>nasal-round</i>	<i>magnitude neutral-rich</i>	<i>timber homogeneity.</i>	<i>narrow-large</i>
instrument effect	0,170	0,479	0,846	1,000	0,040	0,481
musician effect	0,894	0,550	0,928	0,700	0,436	0,467
Interaction musician*instrument	0,756	0,081	0,741	0,785	0,541	0,279

Table 1:  $p$ -values of the two-ways ANOVA

This analysis confirms the examination of the raw data: the differences between the instruments are weak and/or the subjects are not in agreement for their evaluations and/or the repeatability of the subjects is weak. In conclusions, after a debriefing with the subjects, the test n°1 seems to be rather difficult for the subjects, they had not a clear idea of what they had to evaluate.

### Comparative tests

For all the attributes, each subject provided a rank of the 2 instruments (they had also the possibility to rank them placed equal). For the instrument R (arbitrarily), we computed the number of times it has been ranked first (1), second (2) or placed equal (1,5) for all the evaluations. Figure 6 shows the results of the ranking by all the subjects for the four attributes. The ranking of instrument L can be directly deducted of this figure 6, by inverting the rank “1” by “2”.

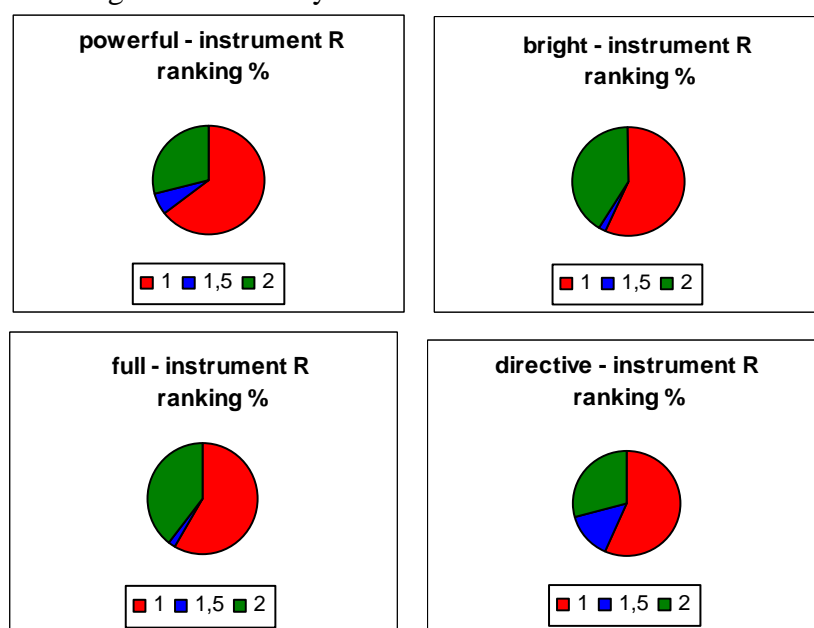


Figure 6 : ranking % for the test n°2.

The raw data (percentage of rank) shows that instrument R seems to be more powerful, more bright, more full and directive than instrument C (the rate of rank 1 is higher than 50%). A Friedman test showed that the following propositions are significant (p-value<0.1%):

- R is more *powerful* and *directive* than C
- The musician X plays more *powerful* and *directive* than Y
- For the musician X, he played more *full* with C than with R (C is his own instrument), and more *directive* with R than with C
- For the musician Y, he played more *powerful* and more *full* with R than with C (R is his own instrument)

A debriefing of the test n°2 with the subjects showed that this test was very intuitive and simple. This test n°2 gives more consistent results than the test n°1.

## RESULTS OF THE SIGNAL ANALYSIS

Globally for the four notes studied, a one-way ANOVA with the factor 'instrument' shows first that the effect of the instrument on the signal-noise ratio  $R_{sn}$  and on the spectral centroid  $S_c$  is not significant. A study note by note is necessary. Table 2 gives, for each note, the average value of the spectral centroid, and the p-value of the instrument effect of the ANOVA.

	C2		G2		D3		A3	
	R	C	R	C	R	C	R	C
$S_c$	8,912	8,666	8,534	7,060	6,435	6,062	6,024	7,018
Instrument effect	p-value = 3.8%		p-value = 0.5%		p-value = 6.8%		p-value = 0.2 %	

Table 2: value of the average Spectral centroid  $S_c$  and p-values of the ANOVA

- For the notes C2, G2, D3, the spectral centroid is higher for instrument R than for instrument C
- For the note A3, the spectral centroid is higher for instrument C

Furthermore, the absolute value of the spectral centroid for the note A3 and instrument C is not really consistent with the  $S_c$  for the notes C2, G2, D3: the string A3 seems to stand out. A debriefing with the musicians and the instrument maker after the session confirmed this fact. The results of this analysis are in accordance with the conclusions of the test n°1: the timber of instrument R is more homogeneous than those of instrument C.

No effect of the musician on the spectral centroid or on the signal- noise ratio was significant, neither globally nor note by note: the effect of the instrument is always more important than those of the musician.

## CONCLUSIONS

In this paper, we presented a study of the perceived differences between to high-end cellos during instrument playing. Two kinds of tests were carried out with a blindfolded jury of six participants: an evaluation test and a comparative test.

For the evaluation test, large differences between subjects and few significant differences between instruments were noticed.

The comparative test provided fairly consensual results: it seems possible to evaluate certain attributes of sound quality in this way. The effect of the instrument was always greater than the effect of the musician. In order to explain the perceived differences between the instruments, metrics based on the signal of the open strings were calculated. The value of the spectral centroid of the sounds was consistent with the conclusions of the listening tests.

## ACKNOWLEDGEMENTS

The authors wish to thank the musicians Etienne CARDOZE and Raphaël PIDOUX, the jury Matthias DEMOUCRON, Nicolas RASAMIMANANA, Pierre CARADOT, Michel ORIANO, Cécile LENOIR and Emilie POIRSON for their substantive input.

**REFERENCES**

- Lavandier M. (2005). Différences entre enceintes acoustiques : une évaluation physique et perceptive. *Thèse de doctorat de l'université d'Aix Marseille II*, 19 décembre 2005.
- Plitnik G. R, Lawson B.A. (1999). An investigation of correlations between geometry, acoustic variables, and psychoacoustic parameters for French horn mouthpieces. *J. Acoust. Soc. Am.* **106**, 1111-1125.
- Poirson E., Depincé P., Petiot J-F. (2007). User-centered design by genetic algorithms: Application to brass musical instrument optimization. *Engineering Applications of Artificial Intelligence*, 20 (2007) 511–518.
- Pratt R.L., Bowsher J.M. (1978) “The subjective assessment of trombone quality”. *Journal of Sound and Vibration* 57, 425-435.
- Pratt R.L., Bowsher J.M. (1979). “The objective assessment of trombone quality”. *Journal of Sound and Vibration* 65(4), 521-547.
- Wright H. “The acoustics and psychoacoustics of the guitar” (1996). PhD Thesis, University of Wales, Cardiff, (1996).