

Developing methodologies to study perceived sound qualities of violins*

Claudia Fritz^{1,**} , George Stoppani², Unai Igartua³, and Jim Woodhouse⁴ 

¹Institut Jean Le Rond D'Alembert, UMR 7190, CNRS/Sorbonne Université, 4 Place Jussieu, 75005 Paris France

²Violin maker, Manchester, 31A Derbyshire Road, Manchester M40 1QN, UK

³Bele, Conservatorio de Musica – Juan Crisostomo de Arriaga, Bilbao, Ibarrekolanda Plaza 1, 48015 Bilbao, Spain

⁴Cambridge University Engineering Department, Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK

Received 15 April 2024, Accepted 12 March 2025

Abstract – Violin makers and acousticians have long sought correlations between sound qualities, construction parameters and vibroacoustic measurements. This is challenging for three main reasons: it is difficult to build violins reliably enough to ensure that the change in the sound is only a result of the intended change of construction parameters; no clear link has been found so far between measurements and perceived qualities; and when listening to violins being played, differences tend to be smoothed out by the players who adapt very quickly. Therefore, while we have previously preferred using players in our experiments to maximise ecological validity and account for the complexity of the interaction between the player and the instrument, in this study we test whether other methods that reduce the player's influence, though more artificial, may be useful for exploring the impact of certain construction parameters on the sound. In the context of a set of violins built with controlled thickness variations in their plates, we conducted two listening tests, based on real recordings of a player and a bowing machine, along with synthesised sounds created from an excerpt recorded with piezo sensors by convolution with radiation measurements in an anechoic chamber. The hybrid synthesis was found to be the most effective of the three in highlighting instrument differences, capturing properties perceived in the actual instruments, and correlating with radiation measurements.

Keywords Violin, Perception, Construction parameters, Sound quality, Loudness

1 Introduction

Since the pioneering work by Carleen Hutchins and Frederick Saunders in the 1950s, many studies by researchers and violin makers have been dedicated to a search for relationships between construction parameters, acoustical properties and sound qualities of violins [1]. The emphasis was originally on the physics, based on various acoustical measurements on different violins which were selected on their global sound quality based on informal testings or even on assumptions that they depended on the age [2–5]. Though there are a few mentions of methods for the evaluation of violins in the second half of the 20th century [6, 7], how violins were actually perceived by players and listeners was only scientifically

studied in the 21st century. Some of these studies were specifically designed to explore the links between the perception and some construction or acoustical parameters via playing or listening tests [8–15], or address some particular questions like the long-term effect of whether a violin was played regularly or not [16], or the comparison of old Italian violins with their contemporary counterparts [17, 18]. Other studies were more general and intended to better understand how violinists evaluate violins: vocabulary used, intra- and inter-agreement, etc. [19–21]. Finally, a large number of studies were dedicated to violin timbre alone, searching for semantic labels that best characterise timbre and for their respective acoustic correlates [22–28].

Despite these numerous studies, as well as recent advances in analysing vibroacoustic measurements [29, 30], the challenge of linking the qualities of violins to their vibroacoustic properties is still elusive, as acknowledged by [18] too. Linking them to construction parameters has been so far elusive as well [31], mainly

*Dedicated to the memory of George Stoppani, who died before the paper was finalised.

**Corresponding author:

claudia.fritz@sorbonne-universite.fr

because violins are made by hand, and so it is difficult to assess whether a perceived change is due to the intentional variation or to irreducible variability in wood properties and geometric tolerance. However, even in the case of recent projects [32, 33] that have used modern technology, in particular CNC routing, to control with high accuracy the constructional parameters, the results of perceptual tests were less clear than had been hoped. One possibility is that the methodologies used for the tests were not optimal within this context, for various possible reasons: the complexity of the task for the players and listeners, the strong influence of the player and/or the excerpt being played in the listening tests, the ordering of the violins, etc. This study will investigate one aspect of that question.

The qualities of a violin can be described in three broad categories [21]: the playability (response, comfort), the loudness and the timbre quality (which we will consider equivalent to tone quality, similarly to [28]). Vibroacoustic properties are commonly obtained by measurements relying on removing the player and replacing them by a controlled excitation via an impact hammer, and looking at frequency responses (measured with a microphone or a mechanical sensor). Many such tests have been conducted by scientists in the field (see many of the previous references) as well as by makers [34, 35]. Makers have concentrated particularly on measurements performed using a microphone, of the transfer function from force at the bridge to radiated sound pressure, often called “radiativity”. Such measurements would be expected to relate to loudness and tone quality, though they cannot provide any direct information on playability.

With the ultimate goal of correlating perceptual judgements with radiativity measurements, in this study we focus on listener evaluations of loudness and timbre quality. The hope is that listeners will be able to make judgements independent of playability issues, unlike players for whom the different aspects may interact, and that these judgements by listeners may be more easily correlated with acoustical measurements, conducted with an impact hammer and not involving a player.

Of course we do not advocate total disregard of the player’s responses, but we recognise that a player has a more complex relationship with the instrument than a listener does. When a player makes judgements of preference between instruments, they are combining different sources of information. A different methodology is needed to investigate such player judgements, and a wider range of acoustical and structural tests should probably be included in any search for correlations.

Even for the more limited objective of investigating listener judgements, there are many issues affecting the methodology of any test. How to listen to the violins? Played by whom? What excerpt? In which venue? Live or via recordings? Many authors have tried to favour ecological validity [36], but there is a fundamental conflict of requirements. We would like the listening judgements to be made in the most natural conditions, but we want to collect data that is free of bias, in sufficient quantity to allow meaningful statistical analysis.

Our natural listening involves combining all our senses plus many kinds of prior knowledge about what we are observing. When we know the “answer” we can usually hear it, which does not constitute a valid psychoacoustic experiment. Our compromise, so far, has been to allow the player to perform the selected material as they wish, but to deprive the audience of knowing which instrument they are playing by interposing an acoustically transparent screen. Musical samples usually consist of short musical phrases from famous violin concertos, covering the full range of the instrument (e.g. [37]). The player may be asked to wear dark goggles so that they cannot see which instrument is in their hands. Others like Nastac et al. [38] preferred to record one violinist in a concert hall and use the recordings in a listening test. In both cases, the listening test relies on the reproducibility of the player, which appears to be lower than expected due to a strong (presumably unconscious) wish to make their “own sound”. Their goal is to deliver a “musical” performance. Thus, differences between instruments seem to be smoothed out, to the extent that listeners are not even able to tell whether a player plays twice on the same violin or two different violins in a row [39]. In addition, if the player plays slightly differently, only very experienced listeners may be able to tell whether the differences heard are due to differences in playing or differences between instruments. This is why some authors have preferred to rely on recordings made with a bowing machine (e.g. [15]).

The specific agenda for the present study is to take advantage of the availability of a special set of well-characterised violins to perform perceptual comparisons using sound files generated by three very different approaches, all of which have been used in earlier published studies. The goal is to explore and compare, via two separate tests, these different versions of “the sound of a violin”, in order to shed light on which might be best for capturing differences between instruments, for revealing a recognisable “identity” of the violin when heard live, and also for maximising the likelihood that perceived sound differences can be correlated with acoustical properties and construction parameters. To this end, a set of violins (described in Sect. 2.1) whose manufacturing was carefully controlled were recorded with three different techniques (described in Sect. 2.2): live recordings, recordings of single notes generated by a mechanical bowing machine, and synthesised recordings made by combining a pre-recorded extract of bridge force with a measurement of the radiativity of each violin. The recordings were then used in two perceptual tests that are described in Section 2.3: a recognition test and an online listening test. Results are presented and discussed in Section 3.

2 Material and methods

2.1 Violins used in the study

The heritage of high quality old instruments includes many kinds of graduation scheme, but rather little is

known about how the playing quality is affected and how the decisions were made. To this end, a number of violins have been made by the authors and their collaborators to study different graduation patterns [40]. Here, only the so-called Bilbao set [13, 32] was used. It consists of six instruments that were carefully built to investigate the influence of the plate thickness on the sound qualities: three instruments with medium backs, each paired with a pliant/thin (V9), normal (V5), or resistant/thick top (V4); similarly, three with medium tops, each paired with a pliant/thin (V13), normal (V11), or resistant/thick back (V1). The two examples of medium top paired with medium back (V5 and V11) serve as a control. Whenever possible, thicknesses were kept uniform to avoid the effects of localised differences. Wood for tops and backs were chosen to be closely matched in density and sound speeds – all tops (*Picea abies*) and backs (*Acer pseudoplatanus*) were cut from the same trees. Greater control was achieved by having all twelve plates and scrolls cut by a CNC router, based on the Stradivarius Huberman model (1713). The outside surface was not changed during the experiment, as the graduation was performed entirely on the inside surface.

The differences in thickness were chosen to maximise differences in the characteristic drive-point impedance Z (the impedance averaged over all possible drive points) – which is a better indicator of the vibratory behaviour of the plates than their static properties like weight and thickness – while still staying in the usual range for contemporary making. This characteristic impedance is estimated using the formula proposed by Davis [41], derived by replacing the arch and thickness variations of the violin plates by a rectangular plate of an effective orthotropic material and applying orthotropic thin-plate bending theory [42]:

$$Z = 4M\sqrt{\frac{2}{13}(f_2^2 + f_5^2)}$$

where M is the mass of the plate and f_2 and f_5 are respectively the frequencies of mode 2 and mode 5 in free conditions. The plates were thus graduated to reach a certain impedance target, which was chosen in the middle of the usual range (based on empirical knowledge and measurements gathered by the second author) for the medium plates. An increase/decrease of 25% was chosen for the target of the thick/thin plates to induce a significant change while still staying in the usual range. For the medium plates, reaching a specific target also allowed compensation for unavoidable variations in the wood properties (density and Young's modulus) between the wedges used for carving the plates.

Weights, averaged thicknesses and the plate impedances for the 12 plates are listed in Table 1. The tight control over the construction parameters of these violins provides an unprecedented opportunity for exploring correlations with sound qualities.

2.2 The three reproduction techniques

2.2.1 Live recordings: a real player and real violins

We recorded a semi-professional violinist in a large seminar room, playing a short excerpt from the Glazunov concerto (Fig. 1) on the six violins described above. This excerpt was chosen for various reasons. First, the goal of this study was to compare different recording methodologies rather than obtain full evaluations (across all strings) of the violins. Listening tests are very tiring so excerpts should remain short. Players and makers agree that playing the G string reveals best the quality of a violin, perhaps because of the large number of harmonics of the notes played on this string. Finally, this excerpt was available in the database of piezo recordings required for the hybrid synthesis (see Sect. 2.2.3) and so could offer direct comparisons between the techniques.

The position of the player was kept constant relative to the omnidirectional microphone (DPA 2006C) throughout the session, at a distance that had been adjusted at the beginning of the session to obtain pleasant recordings (Audio file Violin5_Live ).

The instructions that were provided to the player for the live recordings were similar to previous studies (e.g. [37]): he was asked to play as well as possible on each violin. What matters for players is the musical intention and a good sound (for them). Playing different violins using exactly the same control parameters is not a realistic aim for a player. These instructions may well lead to many (unknown) adjustments by the player, to some “smoothing” of the differences (as already shown in [39]) and even to some timbral changes. But that does not mean that no differences can be heard, even in loudness, as shown in [37], where a significant loudness difference was perceived between old and new instruments.

2.2.2 Bowing machine: an artificial player and real violins

The goal of this methodology is to replace a not so reproducible player by a reproducible excitation (though artificial). The bowing machine is shown in Figure 2. The six violins were recorded on each of the four open strings with the same omnidirectional microphone as for the live recordings (DPA 2006C) placed vertically above the bridge at a distance of 185 mm when the violin lays flat. That distance was adjusted to have, qualitatively, a strong direct sound with a bit of reverberation, in a quiet laboratory room. The violins were rotated for the different strings in order to keep a bowing direction tangent to the bridge curvature.

2.2.3 Hybrid sound synthesis: a real player and virtual violins

The hybrid synthesis is based on the convolution of the recording of an input signal (the force applied by the bowed strings on the bridge) during a live performance with the inverse Fourier transform of a frequency response function of the chosen violin [8]. This methodology relies

Table 1. Weight, averaged thickness and plate impedance for the six tops (before adding the bass bar) and six backs. The $\sim$ for the medium plates indicate that the values that are reported correspond to the average over the four similar plates. The labels of the violins to which they correspond are indicated in brackets.

	Tops (without bass bar)			Backs		
	Thin (V9)	Medium (V1, V5, V11, V13)	Thick (V4)	Thin (V13)	Medium (V5, V9, V11, V13)	Thick (V1)
Weight M [g]	55	~ 63	69	85	~ 100	115
Thickness [mm]	2.0 in average	~ 2.6 in average	3.4 in average	2 in the upper and lower bouts; 4.3 in the island area	2.6 in the upper and lower bouts; 4.5 in the island area	3 in the upper and lower bouts; 5 in the island area
Plate impedance $Z/1000$ [N · s · m ⁻¹]	26	~ 32	40	48	~ 60	75
Free plate eigenfrequencies (mode 2 and 5) [Hz]	110 280	~ 126 ~ 301	155 337	154 313	~ 182 ~ 338	210 360



Figure 1. Excerpt from the beginning of the Glazunov concerto.

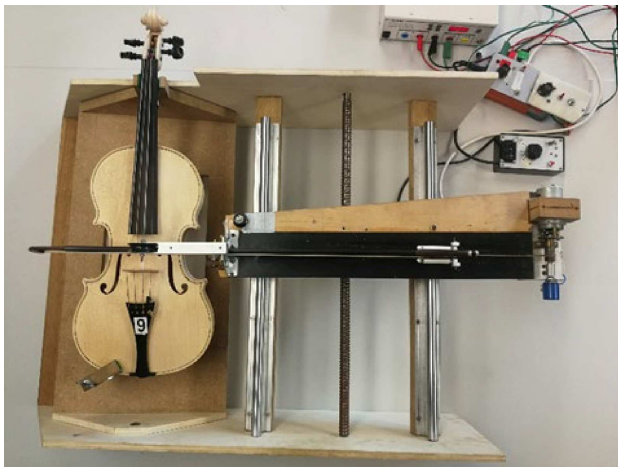


Figure 2. Bowing machine. Audio file Violin5_Bowing Machine.G . Audio file Violin5_BowingMachine.D .

on the fact that to a good first approximation, the body motion of a bowed instrument has little backward influence on the string motion (for more detail, see [8]). Note that this methodology would not transfer to a plucked-string instrument, because in that case the decay time of each string overtone will be influenced by coupling to the instrument body, and so other methodologies are needed (e.g. [9]).

For the input signal, we used recordings of the same Glazunov excerpt made on a violin whose bridge was instrumented with a piezoelectric force sensor under the G string. For the frequency response function, we used a

complex average of 12 radiation measurements, made in an anechoic chamber, with a hammer excitation, and an omnidirectional microphone (see Fig. 3). The violin could rotate around the vertical axis, and the measurements were taken at six different angles (every 60°) between the violin plane and the microphone for two excitation directions (tangential and perpendicular to the bridge). The distance between the microphone and the bridge was set to 20 cm when the microphone was at the front.

This hybrid synthesis makes it possible to drive different virtual violins with the identical realistic forcing waveform (measured during real playing) so that sound differences can be compared with no complications arising from variations in playing (Audio file Violin5_HybridSynth .

2.2.4 The issue of room acoustics

The recordings for the three reproduction techniques were made in different acoustical environments. The room used for each technique was chosen to be most suitable for the recordings/measurements that were required for that technique. Live recordings were done in the kind of space a player would choose for trying instruments, a rather large room with enough reverberation to make the recordings pleasant and natural. The bowing machine was installed in a quiet laboratory room; and finally the radiation measurements were conducted in an anechoic chamber in order to maximise the signal to noise ratio (and therefore the robustness of the method) especially at high frequency. Consequently, the room acoustics is inherent to each reproduction technique.

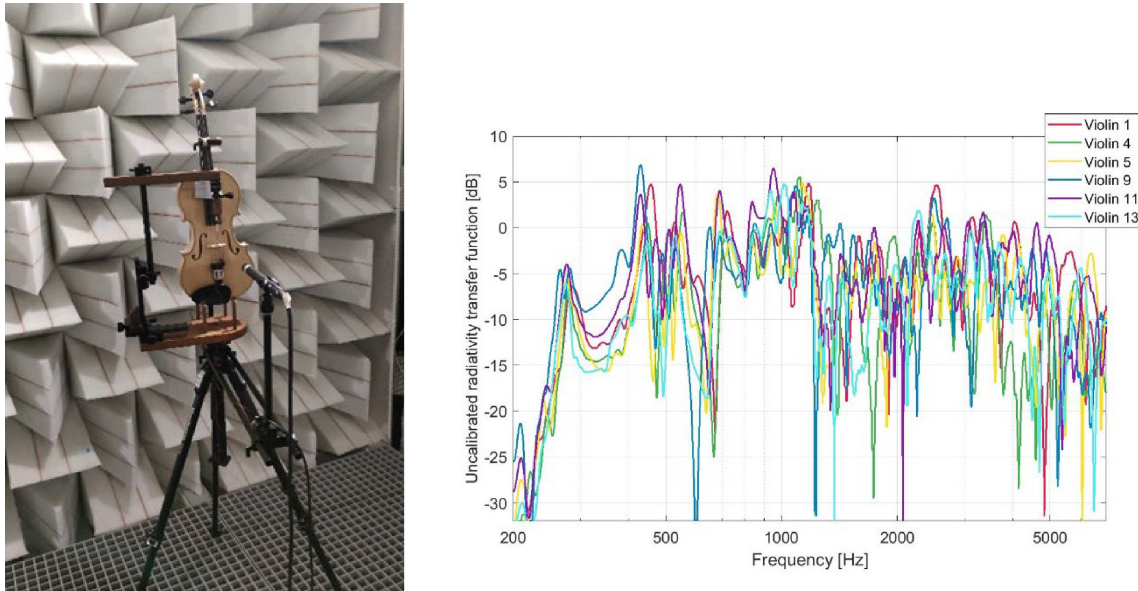


Figure 3. Left: Radiation measurement rig in the anechoic chamber. Right: the uncalibrated radiativity transfer function (magnitude), averaged over the twelve measurements (six positions and two excitation directions) for each of the six violins. For readability purposes, the curves are smoothed over half a semitone.

While room acoustics has been shown to modify the the “absolute” timbre of an instrument [43, 44] and how a player plays [45], there is no evidence that it can affect the relative timbral differences between two instruments [17] nor that it will have a larger effect than changing the player, who is inherent to the reproduction technique too, and whose influence on the evaluation of violins by listeners is small [37].

We decided to avoid any spatial information in the sound files for all three methods. The hybrid synthesis as implemented here is intrinsically monophonic, so we decided to use high-quality monophonic recordings for the other two methods. It is accepted that the spatial nature of the sound field radiated by a live violin may well contribute to perceptions of the kind investigated here, raising issues that could be explored in future work.

2.3 Perceptual tests

Many listening tests can be designed with recordings from this database. The goal of the present study was not to compare the recordings made with different methodologies directly with each other, but rather to compare violins within a set for each reproduction technique, and explore whether listeners’ evaluations are similar in the three cases, qualitatively and quantitatively. This was achieved with two separate tests described hereafter.

2.3.1 Recognition test

Methodology. The goal of this first test was to explore in a very direct way how representative of “reality” each reproduction technique is. For each technique, the

task was simply to match the physical violins with their recordings. A diagram of the experimental set-up is provided in Figure 4.

The experiment took place in an auditorium, during a European Science and Violin Making workshop organised yearly by the authors. The participants could listen to four recordings of the same type via headphones and had to match them with the four physical violins that were laid on the table. They were free to play them as they wished. As not all participants were good violinists, these violins had been played, in the same room, 1 or 2 days before, by a professional violinist, during a session in which the participants could take notes about each violin. During the recognition test, they were allowed (and even recommended) to use their notes as well as previous notes they could have, as these participants were all acquainted with the Bilbao violins and had listened to and played them on different occasions. Only four out of the six violins were used, as a pilot test showed that using the six violins was too difficult and too tiring. As a consequence, one of the twin violins (V11) was removed as well as the one with the thick back (V1). In addition, for the recordings with the bowing machine, only those made on the G string were used.

The test consisted in three separate recognition tasks, one for each of the reproduction techniques. The four recordings presented on the screen as A, B, C, D (see Fig. 4) had been attributed randomly to the four violins for each technique, but were the same for all participants.

Participants. The test was taken by 13 participants, among whom were 9 violin makers, 1 violinist and 3 acousticians. They were regular members of the



Figure 4. Experimental set-up for the recognition test.

Science and Violin Making workshop, and were thus well acquainted with the Bilbao violins.

2.3.2 Online listening test

Methodology. The second test aimed at quantifying the difference in evaluation of all six violins across recording methodologies. We decided to focus on the two main criteria relevant to players (see Sect. 1), which can also be evaluated by listeners: loudness (perceived sound intensity) and timbre quality. For the latter, we were interested, similarly to [10], in a holistic evaluation linked to listeners’ preferences (with high timbre quality being a timbre that is liked and low timbre quality being one that is disliked), regardless of which specific aspects of timbre – being a multi-faceted auditory attribute [46] – were responsible for this overall quality. This approach differs from other studies like [18], which focused on different facets of timbre, such as brightness, openness, or nasality. Obviously, it would be possible to perform similar tests in future work with a more fine-grained methodology.

We opted for an online test which allows a larger number of participants (in particular the violin makers who took part in the Bilbao project and who live across Europe), without necessarily compromising the results due to a lesser control [47]. As the differences between the instruments can be subtle, the participants were recommended, at the beginning of the test, to use the best audio equipment possible (good quality headset). We used the option Audio Perceptual Evaluation [48] available in the Web Audio Evaluation Tool (WAET) [49, 50]. This consists of comparing and ranking stimuli on a criterion, using 0–10 scale.

Each test was divided in two sub-tests, one for each criterion. Each sub-test consists of a series of four pages (in random order, presented once): on each one, recordings made with a given reproduction technique (real player, bowing machine on each the G and the D strings, hybrid synthesis) for a given violin set are compared. The

recordings made with the bowing machine on the A and E strings were excluded from the test as it would have led to a test that was too tiring for the participants. As timbre quality differences are hard to evaluate when there are loudness differences, the stimuli were normalised in loudness for the timbre sub-test. This was done via the loudness normalisation option within WAET [47], which is based on the computation of the LUTS (loudness unit related to full scale) integrated loudness [51]. Figure 5 illustrates one page of the timbre quality test and one page of the loudness test. The green bars correspond to the stimuli. Listeners had first to click on them to listen to all stimuli (in any order and as many times as they wanted) and could then move them freely along the scale to order them according to their judgements). 0 on the left corresponds to the poorest timbre quality/softer stimulus while 10 on the right corresponds to the highest timbre quality/loudest stimulus.

Participants. The link to the test was sent to our network, in order to reach what we could call expert listeners. Twenty seven anonymous participants across the world completed the test: 15 violin makers, 6 violinists, 4 other musicians and 2 others, with an average age of 45 years.

3 Results

3.1 Recognition test

The recognition test does not give any information about the different violins, but it gives a very clean comparison of the three reproduction techniques. In each of the three sub-tests, the participants are comparing a set of monophonic recordings directly with the actual violins, either by playing them on the spot or based on their notes taken during earlier expert performances.



Figure 5. Listening test interface: Ranking by timbre (top) and ranking by loudness (bottom).

For any individual participant, the probability of correctly identifying all four violins purely by chance is 0.039, just below what is commonly used to reject the null hypothesis (here that the participant did the identification by chance). Therefore, only the participants who did a perfect identification can be considered as not having done it by chance. These results are reported in the first column of Table 2 for each reproduction technique.

In a similar way, the cumulative probability of having more than X participants correctly identifying a given violin is below the 0.05 threshold for $X = 7$ ($p = 0.019$) and just above for $X = 6$ ($p = 0.056$). The number of violins which can be considered as being recognised above chance (i.e. by at least 7 participants) or just below chance (by at least 6 people) is reported, for each technique, in the second column of Table 2.

Although this test was quite small in scale, both in terms of number of violins and number of participants, the results give a very clear answer. The two columns lead to the same conclusion: the hybrid synthesis method is by far the best for this particular task. By that method, all violins were recognised by at least seven people, and slightly under half the participants were able to achieve a perfect identification. The live recordings proved to be by far the worst: no participant guessed better than chance and none of the violins was identified better than chance. The recorded notes from the bowing machine were intermediate, with a small number of participants achieving a perfect identification and one individual violin being recognised by at least seven people.

There are perhaps three main factors that have influenced this pattern: room acoustics, realism of playing, and the inevitable variability of a human player repeating the same passage on different instruments. A fourth factor, spatial information about the sound field, is absent in all three reproduction techniques and so should not have influenced their relative performance. The hybrid synthesis and the bowing machine both produced consistent “performances”, while the live recordings did not. The hybrid synthesis and the live recordings both featured realistically complex human performance, while the bowing machine did not. Both these factors are thus qualitatively consistent with the results.

The role played by room acoustics is less clear. The room in which the real violins were tested had broadly similar acoustics to the one used for the live recordings, both being fairly large and moderately reverberant spaces, regarded by musicians as suitable spaces for normal evaluation of different instruments. The laboratory in which the bowing machine recordings were made was smaller and somewhat dead, but by no means anechoic. The hybrid synthesis was based on anechoic measurements. On the face of it, all this might lead one to expect the exact opposite pattern of recognition. The real violin evaluations and the live recordings were made in spaces with rather similar acoustics, while the hybrid synthesis involved the biggest contrast with the real evaluations. This mismatch suggests, tentatively, that room acoustics as such played only a small role, with the other two factors being more important.

Table 2. For each recording type, number of participants who succeeded above chance and number of violins that were recognised above chance (or just below).

	Number of participants who succeeded above chance (perfect identification)	Number of violins recognised above chance (at least 7 people) [or almost above chance (at least 6 people)]
Hybrid synthesis	6	all
Bowing machine	3	V4 [and V13]
Live recording	0	none

3.2 Online evaluation test

The online evaluation test gives further information about the relative performance of the different reproduction techniques, and it also gives information about distinctions between the six tested violins.

3.2.1 Timbre and Loudness evaluations

Figure 6 shows the average ratings given to each of the six violins, for the four types of recordings (live recording with a real player, hybrid synthesis and recordings with a bowing machine on the G and D strings), for the two criteria, timbre and loudness. Within each recording type the ratings were shifted by the average over the six violins (so the mean is zero), as the goal was to explore how the violins were compared relatively to each other for each reproduction technique. It might have been anticipated that the results for timbre would show more spread and less consistency than those for loudness, based on the fact that loudness is a clear-cut attribute whereas timbre is in reality multi-dimensional. There is some evidence for this in Figure 6, but the results show some consistency and suggest some underlying differences in the perceived timbre of the instruments – for example, V11 was clearly preferred to V4.

3.2.2 Correlations between the evaluations for different reproduction techniques

Figure 6 shows that ratings can vary quite a lot depending on the methodology and this was confirmed by a repeated measures ANOVA conducted on the timbre and loudness evaluations. The two factors (the violin and the reproduction technique) as well as their interaction were found to be significant (with $p < 0.001$ in all cases) except for the reproduction technique in the case of the loudness evaluations ($p = 0.62$). In order to explore further the influence of the technique on the evaluations, correlations between the ratings for the six violins obtained for every pair of reproduction techniques were calculated (Tab. 3).

The correlations are weak in general, showing that the relative ordering of the violins (in terms of loudness or timbre) can be quite different depending on the reproduction technique. The correlation between BM – D string and the other methods is even weaker and this can be explained by the fact that it is the only method which

involves the D string. The excerpt used for the live recording and the hybrid synthesis was indeed only on the G string. This is in agreement with the idea that violins can have a different tonal character on each of the four strings and shows the difficulty when evaluating globally a violin (across all strings).

The correlations are even weaker for timbre than for loudness, in agreement with the fact that the reproduction technique has a smaller influence on the loudness evaluations than on the timbre ones.

The best correlation is obtained, for both criteria, between the hybrid synthesis and the artificially bowed G-string. This correlation can even be considered as good for loudness. This is broadly consistent with the findings from the recognition test, where the hybrid synthesis and the G-string notes from the bowing machine gave the best performance in terms of recognition, whereas the live recordings gave a very poor performance. The good correlation can perhaps be explained by the fact that both reproduction methodologies remove the variability in the player/instrument interaction, together with the fact that the Glazunov excerpt was played on the G string.

3.2.3 Perceived differences between the violins for the different reproduction techniques

Figure 6 also shows that the difference between the violins seem to be smaller for some reproduction methods. To better illustrate this, the RMS average, over the 15 possible pairs, of the differences between the zero-measured ratings of any two violins was computed for each reproduction type (Tab. 4).

In general, differences are not large but still, there is almost a factor two (slightly over for timbre and slightly under for loudness) between the artificially bowed G string and the live recording. Based on previous studies, the room acoustics [17] has little influence on the relative evaluation of violins, and perceptual comparisons conducted in live listening tests with different players tend to give similar relative judgements [37]. This would tend to confirm previous findings that players tend to smooth out differences [39].

3.2.4 Correlations with the acoustical measurements made on the violin

One of the goals of this study was to explore if a reproduction technique can enhance possible links with

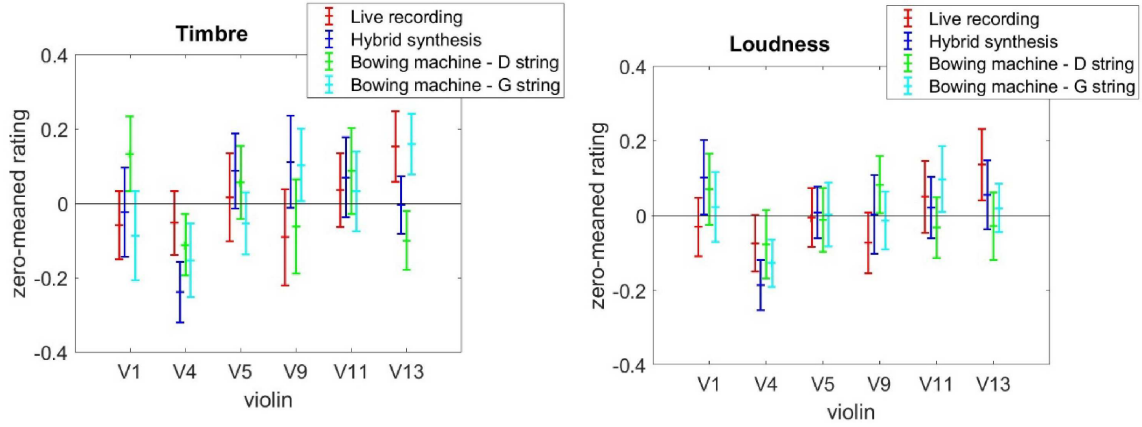


Figure 6. Zero-mean ratings obtained for each of the six violins, using the four reproduction techniques. Results for the two criteria are shown separately, timbre on the left and loudness on the right. These ratings were obtained by subtracting the average over the six violins from the raw ratings, separately for each criterion. Error bars correspond to the 95% confidence intervals.

Table 3. Correlations between the ratings for the six violins obtained for two different reproduction techniques, for both criteria (timbre and loudness).

Correlation	Live/ Hybrid	Live/ BM – D string	Live/ BM – G string	Hybrid/ BM – D string	Hybrid/ BM – G string	BM – D string/ BM – G string
Timbre	0.04	−0.01	0.35	0.41	0.59	−0.23
Loudness	0.47	−0.31	0.58	0.63	0.81	0.29

Table 4. RMS average of the pairwise rating differences computed over all the 15 possible pairs of violins, for the four reproduction techniques and the two criteria.

RMS average of the 15 pairwise rating differences	Live	Hybrid	BM – D string	BM – G string
Timbre	0.12	0.18	0.22	0.27
Loudness	0.11	0.14	0.17	0.20

measured acoustical properties, so a “pseudo-loudness” was calculated from the radiation measurements, to be correlated with the perceived evaluations. The magnitude of the averaged radiated transfer functions plotted in Figure 4 (and used for the hybrid synthesis) were used as the input for the loudness calculation, using the ANSI S3.4-2007 norm [52, 53]. As these transfer functions were uncalibrated, they were artificially divided by a constant value so the average level over the six violins was about 80 dB. The correlation was then calculated between the vector containing the obtained “pseudo-loudness” levels for the six violins and each of the four vectors containing the loudness ratings per recording type. These correlations are provided in Table 5.

The lowest correlation by far is obtained for the live recordings, which confirms that the interaction between the instrument and the player can overrule the intrinsic identity of a violin (as measured with an artificial and repeatable excitation) to the point that listeners cannot

Table 5. Correlation between the perceived loudness evaluations of the six violins for each reproduction type and the computed “pseudo-loudness” levels calculated on the averaged radiated transfer functions.

Live	Hybrid	BM – G string	BM – D string
0.06	0.58	0.77	0.46

guess whether a player plays two different violins or the same violin [39].

3.3 Discussion and relationships with the construction parameters

The various results and analyses presented above lead to the same conclusion: the hybrid synthesis, due to the fact that the playing is controlled but is still natural,

appears as the best methodology to study the qualities of violins, as perceived by listeners. By keeping the playing excitation constant across the studied violins (which is impossible for a real player), this reproduction technique gives by far the best result for recognition of instruments, it enhances perceived differences, and maximises the correlation with acoustical measurements made without a player.

The last goal of this study was to link the construction parameters of these six violins with their perceived qualities. This was secondary to the main goal, the comparison between reproduction techniques, but the test results do give some interesting information despite their limited scope: only the low strings were tested and only for two criteria, which obviously cannot capture all the qualities of a violin. The first observation is that the differences in the perceived qualities are smaller than might have been expected based on the large differences in thickness of the plates (leading to a difference between a medium and a thin/thick plate of around 15% in weight and of 10–15% for the first eigenfrequencies of the free plates, see Tab. 1). This seems to show that the plate thickness has less influence on the finished instruments than it has on the free plates. Gough’s simulations run with a simplified numerical model [54] hint at an explanation in the fact that the vibration patterns of the plates are, in a finished instrument, constrained by the ribs.

Second, if we consider the evaluations obtained with the hybrid synthesis, the only violin that appears significantly different from the others, both in terms of timbre and loudness, is the violin with the thick top (V4). This is in agreement with judgements made by listeners during a live listening test in a concert hall and by players in playing tests [13, 52]. This is the only strong consensus among players and listeners, and that remains valid across most tests: the violin with the thick top appears to be strongly disliked because of its poor sound quality and its quietness. Interestingly the only condition where this is not observed is for the live recordings.

Beyond that, Figure 6 shows that V5, V9 and V11 were all judged to have timbre that was preferred above the mean level almost enough to reach the 95% confidence limits. In terms of loudness, only V1 was judged to have loudness (based on the hybrid synthesis) that was above the mean level by more than the 95% confidence limit. V5 and V11 were the pair with medium tops and backs, so the results suggest a small preference for “normal” instruments in terms of timbre. V9 has a medium back and a thin top, while V1 has a thick back and medium top.

However, something else is observed with the live recordings, but not with the synthesised sounds. The violin with the thin back (V13) is favoured for both timbre and loudness over all the other violins, and significantly over the violins with one thick plate (V1 and V4). Again, this is in agreement with the evaluations of the listeners during the live listening test mentioned above, where V13 was rated the highest among the six violins for timbre and power [13, 52]. Why this listener preference for

V13 appears when the violins are played by a real player but not when they are synthesised is a very interesting question which needs further study. It may lead to a better understanding of playability aspects, related to the interaction of the player with the instrument, which is not taken into account with the hybrid synthesis. The superiority of V13 also appeared when played with the bowing machine on the G string, but only in terms of timbre, not for loudness.

4 Conclusion

Violin makers are seeking tools to quantify perception and relate it to construction differences. This study aimed to explore, through a recognition test and an online listening test, the influence of three reproduction techniques on perceived differences between violins. The goal was to determine which methodology provides the most accurate representation of the actual violins and highlights the largest and most robust differences, which could then be linked to acoustic properties and construction parameters. The study was conducted using a set of six violins, built in a highly controlled manner to minimise unintentional variations in wood properties and craftsmanship, while deliberately varying the thickness of the plates.

Hybrid synthesis, based on recordings of a player using an instrumented violin that are subsequently digitally filtered through a violin’s acoustical response (in this case, the averaged radiated sound measured in an anechoic chamber), clearly emerged as the best methodology in these particular tests. It offers unlimited possibilities for testing various aspects of sound quality and how it is influenced by construction parameters and wood material properties, as well as by measurable vibroacoustical properties as explored earlier in [8] and [9]. The results confirm that recordings made with a player influence perceived sound quality compared to more artificial recordings, probably due to the interaction between the player and the instrument. This interaction can act to smooth out differences, as noted in [39], but this study shows that it can also highlight particular features that still need to be understood. Specifically, why the violin with the pliant back is consistently preferred by listeners across various tests, whether played by a real or artificial player, remains unclear and requires further investigation.

This study also confirms previous findings that a resistant top (with a characteristic impedance of $40,000 \text{ N} \cdot \text{m} \cdot \text{s}^{-1}$) produces an unpleasant and muted sound [55], at least with the Huberman arching design. This is an important step toward better understanding of the various graduation schemes inherited from old Italian instruments: we still lack knowledge about how these affect playability and sound characteristics, and indeed about the rationale behind the original makers’ decisions [56].

Acknowledgments

The authors would like to thank Marion Caumartin and Thibault Ernoult for their help with the live recordings, the design of the listening test interface, and the radiation measurements; Henri Boutin, Benoît Dupeux, Paul Noulet and Silvian Rusu for their help with the recordings made with the bowing machine; Victor Salvador-Castrillo for playing the six violins for the live recordings; Paul Galluzzo for the piezo recording; and finally all the participants to the perceptual tests.

Funding

The construction of the six violins was funded by Erasmus+.

Conflicts of interest

The authors declare no conflict of interest.

Data availability statement

The data are available from the corresponding author on request. The sound files used for the listening tests are available in Recherche Data Gouv, under the reference <https://doi.org/10.57745/VWFP0H>.

References

1. C.M. Hutchins: A history of violin research. *Journal of the Acoustical Society of America* 73 (1983) 1421–1432.
2. C.M. Hutchins: A measurable controlling factor in the tone and playing quality of violins. *Journal of the Catgut Acoustical Society* 1 (1989) 10–5.
3. H. Dunnwald: An extended method of objectively determining the sound quality of violins. *Acustica* 71 (1990) 269–276.
4. E.V. Jansson: Admittance measurements of 25 high quality violins. *Acustica* 83 (1997) 337–341.
5. G. Bissinger: Structural acoustics of good and bad violins. *Journal of the Acoustical Society of America* 124 (2008) 1764–1773.
6. B.A. Yankovskii: Method for the objective appraisal of violin tone quality. *Soviet Physics-Acoustics* 11 (1965) 269.
7. G. Bissinger, F. Gearhart: A standardized qualitative violin evaluation procedure? *Journal of the Catgut Acoustical Society (Series II)* 3 (1998) 44–45.
8. C. Fritz, I. Cross, B.C.J. Moore, J. Woodhouse: Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin. *Journal of the Acoustical Society of America* 122 (2007) 3640–3650.
9. M. Rau, J.S. Abel, D. James, J.O. Smith III: Electric-to-acoustic pickup processing for string instruments: An experimental study of the guitar with a hexaphonic pickup. *Journal of the Acoustical Society of America* 150 (2021) 385–397.
10. C. Fritz, J. Woodhouse, F.P.-H. Cheng, I. Cross, A.F. Blackwell, B.C.J. Moore: Perceptual studies of violin body damping and vibrato. *Journal of the Acoustical Society of America* 127 (2010) 513–524.
11. P. Wrzeciono, K. Marasek: Violin Sound Quality: Expert Judgements and Objective Measurements, in: Z.W. Ras, A.A. Wiczorkowska (eds), *Advances in music information retrieval*, Springer, Berlin, 2010, pp. 237–260.
12. C. Saitis, C. Fritz, B. Giordano, G. Scavone: Bridge admittance measurements of 10 preference-rated violins, in: *Proceedings of Acoustics 2012*, Nantes, France, 23–27 April, 2012.
13. L. Fu, C. Fritz, G. Scavone: Perception of violin sound-post tightness through playing and listening tests. *Journal of the Acoustical Society of America* 150 (2021) 540–550.
14. C. Fritz, V. Salvador-Castrillo, G. Stoppani: The Bilbao project: searching for relationships between sound and playing properties of violins with their construction parameters, in: *Conference on Sound Perception*, Poznan, Poland, September, 2021.
15. C.S. Pannila, A.S. Pannila, D.C. Jayarathna, J.R.P. Jayakody: Evaluation of the sound quality of a classical violin by physically measurable acoustical properties. *Proceedings of the Technical Sessions* 32 (2016) 23–30.
16. Y. Kakegawa, T. Asakura: Effect of physical characteristics of strings on the subjective impression of violin sound. *Acoustical Science and Technology* 44 (2023) 5.
17. R. Inta, J. Smith, J. Wolfe: Measurement of the effect on violins of ageing and playing. *Acoustics Australia* 33 (2010) 1–15.
18. C. Fritz, J. Curtin, J. Poitevineau, H. Borsarello, I. Wollman, F. Tao, T. Ghasarossian: Soloist evaluations of six Old Italian and six new violins. *Proceedings of the National Academy of Sciences of the United States of America* 111 (2014) 7224–7229.
19. C.A. Rozzi, A. Voltini, F. Antonacci, M. Nucci, M. Grass: A listening experiment comparing the timbre of two Stradivari with other violins. *Journal of the Acoustical Society of America* 151 (2022) 443–450.
20. C. Saitis, B.L. Giordano, C. Fritz, G.P. Scavone: Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players. *Journal of the Acoustical Society of America* 132 (2012) 4002–4012.
21. C. Saitis, G. P. Scavone, C. Fritz, B.L. Giordano: Effect of task constraints on the perceptual evaluation of violins. *Acta Acustica united with Acustica* 101 (2015) 382–393.
22. C. Saitis, C. Fritz, G.P. Scavone, C. Guastavino, D. Dubois: Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians. *Journal of the Acoustical Society of America* 141 (2017) 2746–2757.
23. J. Štěpánek and Z. Otčenášek: Rustle as an attribute of timbre of stationary violin tone. *Journal of the Catgut Acoustical Society (Series II)* 3 (1999) 32–38.
24. J. Štěpánek: Musical Sound timbre: verbal descriptions and dimensions. *Proc. of the 9th Int. Conference on Digital Audio Effects (DAFx-06)*, Montreal, Canada, September 18–20, 2006.
25. E. Łukasik: Towards timbre-driven semantic retrieval of violins, in: *5th International Conference on Intelligent Systems Design and Applications (ISDA'05)*, Warsaw, Poland, 8–10 September, 2005.
26. E. Łukasik: Matching violins in terms of timbral features. *Archives of Acoustics* 31 (2006) 227–233.
27. C. Fritz, A.F. Blackwell, I. Cross, J. Woodhouse, B.C.J. Moore: Exploring violin sound quality: Investigating English timbre descriptors and correlating resynthesized acoustical modifications with perceptual properties. *Journal of the Acoustical Society of America* 131 (2012) 783–794.
28. F. Satragno, M. Zanoni, A. Sarti, F. Antonacci: Feature based characterization of violin timbre, in: *Proceedings of the 25th European Signal Processing Conference*

- (EUSIPCO), 28 August–2 September, IEEE, 2017 pp. 1853–1857.
29. S. Giraldo, G. Waddell, I. Nou, A. Ortega, O. Mayor, A. Perez, A. Williamon, R. Ramirez: Automatic assessment of tone quality in violin music performance. *Frontiers in Psychology* 10 (2019) 334.
 30. R. Malvermi, S. Gonzalez, F. Antonacci, A. Sarti, R. Corradi: A statistical approach to violin evaluation. *Applied Science* 12 (2022) 7313.
 31. M. Pezzoli, A. Canclini, F. Antonacci, A. Sarti: A comparative analysis of the directional sound radiation of historical violins. *Journal of the Acoustical Society of America* 152 (2022) 354–367.
 32. C. Fritz: On the difficulty to relate the timbral qualities of a bowed-string instrument with its acoustic properties and construction parameters, in: 2nd International Conference on Timbre (Timbre 2020), Thessaloniki, Greece, 2–4 September, 2020.
 33. C. Fritz, G. Stoppani, U. Igartua, R. Rico, A. Arroita-jauregi, L. Artola: The Bilbao project: How violin makers match backs and tops to produce particular sorts of violins, in: Proceedings of the International Symposium on Musical Acoustics, Detmold, Germany, 13–17 September, 2019.
 34. L. Jost: Work in progress for a PhD. <https://www.fh-zwickau.de/aks/musikinstrumentenbau/forschung/promotionsprojekte/#c26483>
 35. M. Schleske: Empirical tools in contemporary violin making: Part II: Psychoacoustic analysis and use of acoustical tools. *Journal of the Catgut Acoustical Society* 4 (2002) 43–61.
 36. J. Curtin: Measuring violin sound radiation using an impact hammer. *Journal of Violin Society of America: VSA Papers* 22 (2009) 186–209.
 37. J.J. Gibson: The ecological approach to visual perception. Houghton Mifflin, Boston, 1979.
 38. C. Fritz, J. Curtin, J. Poitevineau, F.-C. Tao: Listener evaluations of new and Old Italian violins. *Proceedings of the National Academy of Sciences of the United States of America* 114 (2017) 5395–5400.
 39. S.M. Nastac, V.G. Gliga, M. Mihalcica, A.M. Nauncef, F. Dinulica, M. Campean: Correlation between acoustic analysis and psycho-acoustic evaluation of violins. *Applied Sciences* 12 (2022) 8620.
 40. C. Fritz, R. Viala, V. Salvador Castrillo: Violinist, violin, bow: what can we hear and recognize? in: Proceedings of the International Congress on Acoustics, Gyeongju, South Korea, 24–28 October, 2022.
 41. C. Fritz, G. Stoppani: Developing methodologies to correlate perceived sound qualities of violins with controlled construction parameters, in: Proceedings of Forum Acusticum, Torino, Italy, 11–15 September, 2023.
 42. E.B. Davis: On the effective material properties of violin plates, in: Proceedings of the Stockholm Music Acoustics Conference, Stockholm, Sweden, 30 July–3 August, 2013.
 43. M.E. McIntyre, J. Woodhouse: On measuring the elastic and damping constants of orthotropic sheet materials. *Acta Metallurgica* 36 (1988) 1397–1416.
 44. F.E. Toole, S.E. Olive: The modification of timbre by resonances: Perception and measurement. *Journal of the Audio Engineering Society* 36 (1988) 122–142.
 45. H. Tahvanainen, J. Pätynen, T. Lokki: Perception of bass with some musical instruments in concert halls, in: International Symposium on Musical Acoustics, Le Mans, France, 7–12 July, French Acoustical Society, 2014, pp. 563–568.
 46. M. Kob, S.V. Amengual Garí, Z. Schärer Kalkandjiev: Room effect on musicians’ performance, in: J. Blauert, J. Braasch (eds), *The technology of binaural understanding*, Springer, Cham, 2020, pp. 223–249.
 47. D.J. Hermes: *The perceptual structure of sound*. Springer, 2023.
 48. E. Parizet, D. Grappe, C. Benzecri, C. Coppel: Sound quality assessment: comparison of in-situ and on-line experiments. *Proceedings of the Euronoise, Madeira, Portugal*, 25–27 October, 2021.
 49. B. De Man, J.D. Reiss, APE, Audio perceptual evaluation toolbox for MATLAB, in: Proceedings of the AES 136th Convention, Berlin, Germany, 26–29 April, 2014.
 50. N. Jillings, B. De Man, D. Moffat, J.D. Reiss: Web audio evaluation tool: a browser-based listening test environment, in: 12th Sound and Music Computing Conference, Maynooth, Ireland, 25 July–1 August, 2015.
 51. N. Jillings, B. De Man, D. Moffat, J.D. Reiss, R. Stables: Web audio evaluation tool: a framework for subjective assessment of audio, in: 2nd Web Audio Conference, Atlanta, USA, 4–6 April, 2016.
 52. Recommendation ITU-R BS.1770-4: Algorithms to measure audio programme loudness and true-peak audio level. International Telecommunication Union, 2015.
 53. ANSI S3.4-2007: Procedure for the computation of loudness of steady sounds. American National Standards Institute, New York.
 54. B.C.J. Moore, B.R. Glasberg, T. Baer: A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society* 45 (1997) 224–240.
 55. C. Gough: Violin plate modes. *Journal of the Acoustical Society of America* 137 (2015) 139–153.
 56. V. Salvador-Castrillo: Correlations between construction parameters, vibro-acoustic characteristics and perceptual qualities of a pool of 13 violins, Master dissertation, Sorbonne Université, 2021.
 57. B.C. Stoel, T.M. Borman: A comparison of wood density between classical Cremonese and modern violins. *PLoS One* 3 (2008) e2554.