ACTA ACUSTICA UNITED WITH ACUSTICA
Vol. 101 (2015) 369 – 381
DOI 10.3813/AAA.918833

Perceptual Evaluation of Musical Instruments: State of the Art and Methodology

C. Fritz, D. Dubois
LAM, Institut Jean Le Rond d'Alembert, UMR 7190, CNRS / UPMC Univ. Paris 06, France.
claudia.fritz@upmc.fr

Summary
An object of study in mechanics for more than three hundred years, the violin has only recently been scientifically studied from a perceptual point of view. A range of investigations which have been conducted since 2005, complemented by a few studies of other instruments, offers an illustration of possible methodologies, and serves as a basis for discussing their respective advantages, as well as their limitations and issues. Since methodological choices depend on the goals of the study and on the theoretical (conceptual) choices, this review focuses on new methods borrowed from recent research lines developed in contemporary psychology.

PACS no. 43.66.Jh, 43.75.Cd, 43.75.De

1. Introduction

The violin is one of the most culturally important instruments in Western society, and has therefore been extensively studied by scientists in physical sciences for more than three centuries [1]. Besides the physical modelling of the instrument and its control by the player (e.g. recently [2, 3, 4, 5, 6, 7, 8, 9]), a long-standing goal of violin acoustics research has been to correlate perceptual evaluations of violin qualities with specific properties of physical structure and dynamic behaviour, which could be, in particular, extracted from admittance and/or radiation measurements. Most of this research was grounded in acoustics and based only on the subjectivity of the authors and not on any psychological approach of sensory judgements. For instance, Alonso Moral and Jansson [10] suggested the importance of the signature modes below 600 Hz and the bridge hill in the 2–3 kHz range for violin sound quality, based on bridge admittance measurements on 24 violins, which had previously been played and rated on tonal quality by two professional violinists. Hutchins [11] suggested a correlation between the spacing between the A1 and B1 modes measured on 37 violins and comments about the quality of these violins. These comments were however only made by the respective player or owner of each violin. Dünnwald [12] measured the sound output of 700 violins from a single microphone position and derived a combined parameter based on spectral considerations which, along with the level of the first signature mode (the Helmholtz-like cavity mode called A0), allowed him to categorize the 700 violins into classes, and to separate the "good" violins from the "bad" violins. However, what is meant by good and bad, and who decided which violins were good or bad, is not specified.

It seems this was solely based on the reputation of the makers and/or the owners (famous soloists for instance) and therefore cannot be considered as scientific statements in the domain of psychology of perception to be related with acoustic properties but as opinions to be studied as social sciences issues. Langhoff *et al.* [13] conducted experiments in which violin performances were filtered digitally. They used one violin as a baseline and then modified its frequency response curve (and therefore its impulse response) in several ways, to give enhancement of the A0 mode and of mid-range frequencies (around 1.7 kHz), and creation of a smoother decay towards higher frequencies. This experiment did show that it was possible to compare violin spectra by listening to digitally filtered signals but it did not aim to directly address the question of how people perceived the different sounds created. Their paper only reports the "subjective impressions" of one of the authors and no other participant was involved. Finally, Bissinger conducted a wide range of vibration and radiation measurements [14] of 17 violins and correlated them with quality ratings from "bad" to "excellent". However, the quality ratings were provided by just a single professional player for 12 of the violins, and by Bissinger himself for five. Few details on the rating procedure itself were provided. The correlations showed no significant quality differentiators except for the A0 mode, the radiation of which was significantly stronger for violins evaluated as excellent than for the ones evaluated as bad.

This search for correlations between "subjective" (psychological / perceptual) and "objective" (physical / acoustical) properties of musical instruments is nowadays a general issue in musical acoustics, and the topic of this review. The interest for this issue became larger in the last

Received 25 March 2014,
accepted 12 August 2014.

five decades, and has been growing continuously for the last fifteen years. Examples can be found for all kinds of instruments: starting (to our knowledge) with the piano in 1962 [15], then followed by the trombone [16], French horn [17], trumpet [18, 19, 20, 21], flute [22], bassoon [23], clarinet and saxophone reeds [24, 25, 26, 27], didgeridoos [28, 29], cellos [30], oboe [31], electric guitar [32],… However, what distinguishes these examples from the violin studies mentioned previously is that they contain controlled (at least to some extent) perceptual studies related to experimental knowledge in psychology along a well established paradigm of psychophysics. Even though most of the effort was generally put into the mechanical measurements and the physical modelling, and thus some of the perceptual studies were rather succinct, they contributed to identifying some perceptual properties of the instruments being studied which were not just the views of the authors, and which could therefore serve as an objective/scientific basis for correlations between psychological evaluations and mechanical measurements. However, in the case of the violin, we could not find such controlled study on how violins are evaluated until recently (before 2005), and thus the main problem with the attempts at correlating dynamic measurements with perceptual data in search of quality parameters has always been the lack of convincing and objective perceptual data. Moreover, while many studies actually tried to explain the presumed and largely accepted tonal superiority of the Old Italian violins, and many factors had been proposed and/or investigated to account for it (including properties of the varnish [33, 34]; effects of the Little Ice Age on violin wood [35]; differences in the relative densities of early and late growth layers in wood [36]; chemical treatments of the wood [37, 38]; plate tuning methods [39]), no one actually investigated the fundamental premise of tonal superiority of old violins as a sensory experience, in psychology of perception.

This paper is an attempt to review studies in musical acoustics that address the musical quality of instruments and to reposition the different investigations within their disciplinary domains of concern – mechanics, acoustics, signal processing as well as psychology. Focusing on perceptual studies, our contribution aims at making explicit the premises (within the field of psychology) when designing experimental settings for instrument evaluations, mainly through questions that we had to face when contrasting perceptual tests when listening only and when listening while playing. This issue is not just a technical one leading to suggesting "recipes" but one which addresses what is actually questioned in the different experiments: the sound of an instrument as a mechanical device? As a musical instrument ? And in the latter case, is it possible to forget that the sound is produced (as music) by a musician? The paper is therefore divided in 3 sections. The first two contrast procedures involving listening (only) and playing, which lead us, in a third section, to discuss some methodological issues and to suggest recommendations for improving the investigation of musical instruments quality.

2. From listening tests …

2.1. For comparing modern violins to Old Italians

The question of the presumed tonal superiority of the Old Italian violins can be addressed in a diversity of scientific fields. The excellence of Stradivari violins established as a myth, as a collective representation to be studied within the fields of musicology and sociology, remains to be questioned as individual representations related to sensory experiences in psychology, before being related to the physical characteristics of these instruments, within the field of mechanics. Since 1817 [40], many informal blind listening tests have been conducted and the results all showed that new instruments stand up very well, and often outscore their older, more expensive counterparts. Some tests were particularly famous: a BBC programme in 1977 with violinists Isaac Stern and Pinchas Zukerman and the violin expert and dealer Charles Beare [41]; more recently in Sweden in 2006 with an panel of judges mostly comprising members of the European String Teachers Association [42]; and in 2009 with the British violinist Matth ew Trusler, who played his 1711 Stradivarius and four modern violins made by the Swiss violin-maker Michael Rhonheimer. One of the new violins was made with wood that had been treated with fungi [43]. However, each test has been discredited or dismissed as meaningless by the violin community as they were unscientific and relied on flawed methodology. In particular, they were rarely conducted in the double-blind format, where neither the panel of judges nor the player knows the identity of the violins being evaluated.

A first (to our knowledge) scientific test to address not exactly this specific issue of 'Old Italians versus new violins' but the more general issue of the effect on violins of ageing and playing was conducted by Inta *et al.* in 2005 [44]. A pair of violins that were as similar as possible was commissioned. One instrument was then kept for three years under environmentally controlled conditions in a museum, whilst the other was played regularly by a professional musician. Listening tests were conducted 'live' in a concert hall when the violins were new, and then replicated three years later. Listeners were good amateur violinists and each member of the listening panel became in turn a player. The violins were presented a certain number of times in random order and listeners rated them on a 1 (poor) to 10 (excellent) scale for five criteria worded as: *evenness*, *clarity*, *projection*, *distinctive character* and *warmth*. Instruments were rated for their sound alone (as opposed to performance quality). Results showed no significant differences at the 98% confidence level, for any criteria, between the two violins.

2.2. For searching for correlates between sound evaluations and acoustical characteristics

Almost simultaneously, Fritz *et al.* [45, 46, 47, 48] started to establish quantitative links between acoustical parameters of the instrument body and the perceptual evaluations of a listener, using the methodology of "virtual violins". Representative force waveforms are recorded using normal playing on a violin whose bridge is equipped with a piezoelectric force sensor under each string. These pre-recorded force functions can then be applied to different violins, so that sound differences can be compared with no complications arising from variations in playing. The mechanical frequency response function of these different violins was mimicked using a digital filter, and the output signal for listening tests was generated by applying this filter to the recorded bridge force signal. Once the violin response is represented in digital filter form, it becomes very easy to make controlled variations of a kind which would be virtually impossible to achieve by physical changes to a violin. This methodology is similar to the one used by Langhoff et al. [13] except that the violin bridge was mounted with a force sensor (and not a velocity sensor, from which it was difficult to derive the force). The goals were also different. For instance the aim of Fritz et al. [46] was to report the results of psychoacoustic measures of the ability of musically and non-musically trained subjects to discriminate changes in the frequency and amplitude of single and multiple resonances. This initial study explored two aspects of violin acoustics which received great prominence in the earlier literature as possible indicators of violin sound 'quality': (1) the A0 mode as well as two other individual low-frequency modes of vibration (below 700 Hz), which dominate the low-frequency output of a violin and are usually labelled B1- and B1+ ; and (2) a set of four frequency bands proposed by Dünnwald [12] (190-650 Hz, 650-1300 Hz, 1300-4200 Hz and 4200-6400 Hz). The bridge force signals used were two short bowed single notes, at 196 and 330Hz. For modifications of amplitude, the lowest thresholds were in the range 3-5 dB for individual modes and 1-3 dB for the Dünnwald bands. For modifications in frequency, the lowest thresholds were around 3-5% for individual modes, 1-3% for the first three Dünnwald bands, and around 1% when all frequencies were varied simultaneously. Frequency changes in the 4th Dünnwald band were not detectable.

In [47], Fritz *et al.* explored how the perception of violin notes is influenced by the magnitude of the applied vibrato and by the level of damping of the violin resonance modes. Damping influences the "peakiness" of the frequency response, and vibrato interacts with this peakiness to produce fluctuations in spectral content as well as in frequency and amplitude. Initially, it was shown that thresholds for detecting a change in vibrato amplitude were independent of body damping, and thresholds for detecting a change in body damping were independent of vibrato amplitude. A study of perceptual similarity using triadic comparison of synthesized stimuli (varying in body damping and vibrato amplitude) showed that vibrato amplitude and damping were in this case largely perceived as independent dimensions. A series of listening tests was conducted employing synthesized and recorded performance to probe perceptual judgements in terms of *liveliness* (i. e. how lively and responsive the violin is) and *preference*. The results do not support the conclusion that *liveliness* results from the combination of the use of vibrato and a "peaky" violin response. Judgements based on listening to single notes showed inconsistent patterns for *liveliness*, while preferences were highest for damping that was slightly less than for a reference (real) violin.

The same rationale can be found for other instruments. For instance, Poirson *et al.* [19] investigated the concept of *brightness* for trumpet tones which were generated on a trumpet mounted with a mouthpiece of variable depth in three ways: by a trumpet player, by an artificial player and by physical modelling simulations. This study allowed the authors to find that the magnitude of the impedance peak corresponding to the second harmonic of the tone was highly correlated with *brightness*, and seemed to be the cause. It also allowed them to compare the three ways by which sounds were generated and thus to check the perceptual realism of their artificial mouth as well as their numerical model. Sounds generated by the artificial player or simulated by the harmonic balance technique were found to be perceived in a similar way to the natural sounds when judged on brightness. This augurs well for the use of the artificial player for studying the quality of wind instruments or virtual acoustics techniques in the conception of new instruments.

The results of these studies do not only question the relevance of the acoustic parameters that could influence perceptual judgements. They also lead to questioning the psychological criteria along which the perceptual judgements are made through the use of terms such as *liveliness* and *brightness* and to explicitly integrating the classical psychoacoustic methodologies borrowed from sensory analysis (norms ISO 2003) which consist of collecting lists of words produced by listeners to describe violin sound and then processing them in order to find one word that could be considered, through the negotiation of a consensus, as an adequate term or descriptor (see [49]).

In [48], Fritz *et al.* collected sixty-one common English adjectives used to describe violin timbre and asked violinists to arrange them on a map, so that words with similar meanings lay close together, and those with different meanings lay far apart. The results of multidimensional scaling (MDS) demonstrated consistent use among violinists of many of these adjectives, and highlighted which were used for similar purposes. The authors then investigated the perceptual effect of acoustical modifications of violin sounds produced by a roving of the levels in five one-octave wide bands, 190-380, 380-760, 760-1520, 1520-3040, and 3040-6080 Hz. Pairs of synthesized sounds were presented through headphones and each participant (12 expert violinists, one violin maker and one acoustician/musician) was asked to indicate which of the sounds was more *bright*, *clear*, *harsh*, *nasal*, or *good* (in separate runs for each criterion), These criteria were selected for being widely spread on the three-dimensional MDS map and for being likely related to differences in spectral shape. Increased *brightness* and *clarity* were associated with moderately increased levels in bands 4 and 5, whereas increased *harshness* was associated with a strongly increased level in band 4. Judgements differed among participants for the qualities *nasal* and *good*.

Borrowing new developments of cognitive psychology concerned with "natural" categorisation [50] (see [79] for sensory categories), Bensa *et al.* [51] used a free categorisation task (associated to verbal description of each resulting category) to determine the perceptual influence of two control parameters of a piano sound synthesis model: inharmonicity and "phantom" partials. 17 piano sounds were synthesised so that they varied in terms of these two control parameters, in order to cover a wide range - from a sound with very weak inharmonicity and no "phantom" partials to a sound with exaggerated inharmonicity and a high level of "phantom" partials. The structure and characteristics of the categories obtained by free sorting of these 17 stimuli allowed deriving general conclusions about timbre cognitive processing. In particular, the study showed major differences between the physical and the cognitive descriptions, the first one having a dimensional character, and the second being categorical along family resemblance based on correlates of attributes [50, 80]. Different categorical structures can correspond to the unique description of the stimuli in the physical space and they depend on the strategies of the subjects, based on their expertise and their experience.

2.3. Discussion

The studies described in this section show large variations between listeners when evaluating musical instruments and therefore resonate well with Coggins's remark (about the difference between old and new violins) [42]: "Perhaps the real answer, though, lies not so much in the actual sound that is produced, but more in some intangible interaction between the player and the instrument".

Similarly, regarding the experimental investigation of the perceptual correlates of violin acoustics, Fritz *et al.* [46] acknowledge that their results are only part of the story of violin discrimination, as higher-level perceptual processes are brought into play when a trained violinist compares instruments in a musical setting - for example during the process of choosing a new instrument (in order to buy it). This can explain why differences were obtained in [47] between the judgements by violinists made with synthetic or live performances. Live performance was achieved by playing on an electric violin whose output was filtered in real time by the same filters used to synthesize the sounds for the listening tests. Even if the evaluation is restricted to sound quality, four causes can explain the differences between a playing evaluation and a listening one. First, most listening tests are constructed on short excerpts (even single notes) because of time constraints (to reduce listeners' fatigue and boredom) and experimental requirements for regular control of parameters, but without any theoretical consideration about the equivalence of what is processed (for instance, an excerpt of two notes is processed as two notes while an excerpt of six notes is processed as music [52]). Second, the same instrument can sound very differently when played by different musicians. Third, the sound of an instrument is evaluated differently when listening while being "passive" with respect to the sound production - i.e. listening to the sound produced by someone else - compared to listening while being "active", i.e. while generating the sound. The evaluation during a listening test is indeed made by relying on the sound only and is thus mainly based on the resultant sound without any possible comparison nor control on the nature of the sound and the manner by which it was produced. And finally, in playing tests, sound quality is intrinsically entangled with playing properties (such as playability, response, ...) during the evaluation (by a player) of an instrument, for which the control of the instrument when producing the sound is essential as proved, in the particular case of the violin, by the agency given to the violin in the assessments made during the playing task and the players' statements regarding what is a "good" or a "bad" violin [53]. In short, the sound to be perceived and judged is not the "sound" *per se*. The "same" (acoustically speaking) sound is, psychologically speaking, processed as a musical sound of a violin, differing from noise or speech for example, therefore inducing a different way of listening. In addition, this musical sound is a complex sonic object resulting from an interaction between a player and an instrument; an interaction which intends to produce some specific type of sound/music. Therefore studies on sound quality of musical instruments have to take into account not only acoustical descriptions but also the complexity and diversity of the relations (with the instrument and the player) in which the sound is entangled. As a consequence, the sound quality of an instrument has to be evaluated by taking into account both the instrument and the player, one way being the use of playing tests. This is obviously/a fortiori even more important if the instrument quality (and not just its sound quality) is the object of study. This was already discussed by Pratt and Bowsher in 1978 [16]. Having conducted a preliminary listening test followed by a large scale one, they concluded: "In view of the difficulties experienced by listeners in discriminating between instruments and players, and also since listeners can rate only the timbre of the instruments, it was decided to concentrate on the use of players as subjects for the remaining experiment." However knowledge can still be gained by contrasting results from listening tests (only) and playing tests.

3. ...to playing tests

3.1. For searching for relationships between perceptual evaluations and mechanical measurements of instruments

With the exception of [16] (though it was followed by [54]) and very recent papers on the violin (see section 3.2), papers containing playing tests aimed at establishing correlations and explicit and systematic relationships between perceptual properties and mechanical/acoustical characteristics in order to: 1. search for quality parameters, i.e. the determinants of the quality of an instrument while playing it in the case of the didgeridoo [29], saxophone or clarinet reeds [25, 26, 24], violin bow [55]; 2. check the influence of a single construction parameter, like the geometry of the mouthpiece of the French horn [17], the crook profile of the bassoon [23] or the neck-to-body junction of the electric guitar [32]; 3. check manufacturing consistency, in particular defects like a leak in the bore of a trumpet [21] or differences in bore profiles of oboes [31].

However, there has been little interest until quite recently in better understanding how players evaluate instruments per se, with the search for quality parameters only as a long term goal, after the musicians' evaluation had been directly addressed.

3.2. For studying how violinists evaluate violins

Weinreich wrote in 1993 [56] "*no* [objectively measurable] *specification which successfully defines even coarse divisions in instrument quality is known* (author's italics)" and this still remains a challenge. Finding such a specification would be easier if there were a general consensus among violin players, makers, and dealers on how to rate violin quality. However, it is illusionary to find any consensus on which particular violins are considered better than others and in what particular ways, because players are not relying only on physical parameters but also on psychological properties memorized during previous experience, similar to any sensory evaluation (see for example sensory analyses conducted in food industry, in which global judgements of consumers are clustered and put into relation with analytical descriptions of expert panels [57]). It is therefore important to study how players evaluate instruments, as pointed out by Bissinger at the end of his large-scale study [14] which did not allow him to find conclusive correlations: "What truly defines violin excellence? If the answer is truly excellent violinists, then the reliability/reproducibility of their psychoacoustic judgements must draw more attention." This has been the starting point of recent work that investigates players' experience in terms of self-consistency and inter-individual variability, as well as the contribution of different sensory modalities in their evaluations.

In a series of experiments [58, 59], Saitis *et al.* addressed the question of self-consistency of experienced violinists as well as between-individual agreement. Only what is considered as methodologically important will be reported here, and the reader is invited to refer to the corresponding papers for further details. A first playing test [58] involved 20 skilled violinists who had to rank in order of preference 8 violins selected from different makes, ages and prices. In order to maximise ecological validity and emulate as closely as possible a real situation of evaluating instruments that could happen in the context of purchasing a new instrument, players were asked to use their own bow. The experiment took place in a relatively dry (acoustically) environment as violinists consider such venue best for initial try-outs, as the direct sound from the instrument is not so much coloured by room reflections as in a concert hall [60]. Finally, no playing constraint was imposed on the evaluation process (e.g., specific repertoire). However, though in a real-life search, players would probably start by looking at the instruments, the point was here to circumvent the potential impact of visual information on judgement, while ensuring a certain level of comfort for the musicians and therefore low light conditions were used and participants were asked to wear dark sunglasses. The experiment was divided into two identical sessions, at least three days apart. In each session, after a familiarisation phase, participants had to do the ranking five times (the violins being placed on a table in random order each time).

In a second playing test [58], 13 skilled players had to assess 10 violins (of different make, age and price), by five criteria. To reduce variability of interpretation of the scales across all participants as much as possible, each criterion was presented, in the form of a descriptive sentence alongside a short explanatory text (not provided here, all in English): *the violin is easy to play*, *the violin responds well*, *the violin has a rich sound*, *the violin is well balanced across all strings*, *the violin has a broad dynamic range* and overall preference. The right end of each scale was labelled as "strongly agree" while the left end was labelled as " strongly disagree" . Violins were presented one at a time and the experiment was conducted under the same practical conditions as the previous one.

The first experiment showed that players are self-consistent when assessing different violins and this self-consistency does not appear to correlate with known recorded characteristics of the participants (e.g., years of violin training) . The second experiment showed that violinists agreed on their preference for violins with a rich sound (as produced when playing) and, to a lesser extent, a large dynamic range. However, both experiment showed a large amount of inter-individual variability and that the evaluation of violin attributes widely varied between individual players.

These results invited the authors to go further in the psychological investigation of interindivdual differences. One of many hypotheses about the origin of the large inter-individual differences in violin preference is that players may take varying playing approaches to assess different attributes of the instrument. A third experiment [59] was thus designed to investigate the perceptual evaluation of the concepts of *richness* and *dynamic range* in playing tasks based on prescribed musical material and techniques in order to compare intra-individual consistency and inter-individual agreement in constrained (i.e., playing only certain notes on certain registers) versus unconstrained (i.e., playing a certain excerpt from the violin repertoire) tasks for the cases of these two criteria. 16 skilled players were asked to rank/rate five violins (of different makes, ages and prices), presented simultaneously, on scales named *richness* and then *dynamic range*, first in a constrained task, and then in an unconstrained task, for which they had to indicate their overall preference as well. For each trial, the assessment was done on five scales (one for each violin) presented simultaneously on a computer screen (using on-screen sliders) and participants were instructed to always rate their top choice as 1 and their lowest as 0.

The results of this third study show a higher inter-individual agreement in the playing tasks relative to the previous studies and suggest that the constraints of the tasks should be carefully taken into account when designing the experiment. On the one hand, this observation supports the hypothesis that different violin players may take varying approaches to assess different attributes of the instrument and hence designing focused evaluative tasks may trigger more agreement between individuals. On the other hand, it is possible that participants were able to agree more with each other because they had to evaluate only five violins, a smaller number than in the previous studies. In terms of self-consistency, participants appeared slightly more self-consistent in this study than in the second experiment. To a certain extent, these observations may suggest that when evaluating a set of violins in such experiments, comparing all instruments at one time may be more adequate as a more "ecological" processing of violin evaluation than assessing each violin individually.

In parallel, Wollman *et al.* have been investigating the role of auditory and tactile modalities when evaluating the quality of a violin [61, 62, 63]. In particular, they designed an experiment employing a blind violin evaluation task under different conditions [63]: i) by holding the instruments without producing sound ii) under normal playing conditions, iii) with auditory masking, and iv) with vibrotactile masking. Under each playing condition, 20 violinists evaluated five violins according to criteria related to violin playing and sound characteristics, rated their overall quality and relative preference. Both auditory and tactile modalities appeared important in the violinists' evaluations, but their relative importance was found to depend on the violinist, the violin and the type of evaluation (criteria or preference). In particular, the importance of the sound of a violin to its preference depends on the violin in this task.

The investigation intended to establish as well a correspondence between the different attributes of a violin and the sensory modality they appear to be associated with. Three separate groups of criteria were suggested. One group consists of criteria mainly related to violin sound, namely *sound richness* and *sound palette*, though about a third of the players could still judge these criteria with auditory masking! A second group consists of four criteria that relate to both auditory and tactile modalities, namely *liveliness*, *dynamics*, *loudness/power* and *evenness*. The third group consisting of *responsiveness* and *ease of playing* includes criteria that depend to a large extent on tactile stimulations as cues related to the musician actions while playing. Finally, the overall quality ratings were accurately predicted by the rating criteria, which also proved to be perceptually relevant to violinists, but were poorly correlated with the preference ratings, suggesting that the two types of ratings may stem from different psychological processes to be further identified.

To this end, the design of experimental settings should, instead of being solely based on analytical knowledge in acoustics, result from the musicians conceptualisations (as holistic and multimodal cognitive representations), inferred from psychological and linguistic methods. Such procedures have already been developed in others sensory domains [64] and used in auditory research on soundscapes [65, 66] or musical instruments (piano [51, 67], guitar [68]). For the violin, a first attempt at using verbal data to better understand how violinists evaluate and conceptualise violin quality focused on the differences between playing and listening [53]. The linguistic analysis showed that there are clearly two different objects under consideration for the musician: the *violin* and the *sound*. As far as the psychological evaluation is concerned, musicians mainly focus on their relationship with the instrument while playing (in all the polysemy of the word) with it, the produced sound leading therefore to a different evaluation while listening. The conceptualisation of violin quality evaluated when playing has then been more thoroughly investigated by Saitis et al. [69, 70] using spontaneous preference descriptions by experienced performers collected in a playing-based perceptual evaluation experiment. Upon ordering a set of different violins in terms of preference, players were asked to explain their choices via an open questionnaire. A linguistic analysis confirmed that there are two objects of consideration. It further revealed that the lexicon used by players to refer to *sound* contains mostly descriptive and evaluative adjectives in simple, denominal or deverbal constructions and can be divided in four underlying semantic fields: *texture-temperature*, *action-presence*, *size-volume* and *light*. Furthermore, the constant comparison technique from grounded theory was employed to develop a classification scheme of concepts and the attributes that embody them. At a first level of analysis, three underlying themes of evaluation emerged from the data. The HANDLING refers to the ergonomic aspects of the violin-musician system and relates to concepts such as responsiveness, comfort and flexibility of playing. SOUND comprises descriptions about the quality, quantity and spatiality of the produced sound. And finally RELEVANCE (to the player) refers to quality judgements based on the musical, cultural and emotional involvement of the player.

3.3. Back to comparing modern and Old Italian violins

The detour made in the previous subsections on the analysis of violin evaluation (in general) while playing allows us to go back to one of our initial questions related to the comparison between old and modern violins.

The effect on violins of ageing and playing was investigated as well in [44] with playing tests in addition to the listening tests described above (section 2.1). Each member of the listening panel in turn became a player and had to rate the instruments along ten criteria (*warmth*, *eveness*, *brightness*, *speaking ability*, *playability*, *responsiveness*, *character*, *dynamic range*, *sound preference*, *playing pref-*

erence). The results are similar to those obtained with the listening test: three years of regular playing has not made any statistically significant difference to the performance of one of the pair of violins. However, no statistically significant difference does not mean no difference: the difference may have been too small to be detected at a statistical level with a limited panel. As the number of players/listeners is actually not provided, no firm conclusions can be stated.

These preliminary results along Coggin's comment (see section 2.3) led to an investigation whether there is actually a "true" superiority of the Old Italians among very experienced players. In a first experiment conducted during the 2010 International Violin Competition of Indianapolis 21 experienced violinists were asked to compare three violins by Stradivari and Guarneri 'del Gesu' with three high quality new instruments [71]. The methodology was, for identical reasons, similar to the one presented in section 3.2: room with relatively dry acoustics, reduced lighting, dark goggles, participants' own bow when possible. In the second part of the study, player preferences were explored under conditions designed to maximize ecological validity (within the time constraints): it emulated the way players choose instruments at a violin shop, where they typically try a selection of instruments before selecting one to take home for further tes ting. Participants were thus asked to choose, among the six instruments, the single instrument *they would most like to take home with them* as well as the instruments they considered *best* and *worst* in terms of *range of tone colours*, *estimated projection*, *playability*, and *response*. They were invited as well, at the end of the session, to guess the "making-school" of their take-home instruments. Just 8 of 21 subjects (38%) chose an old violin to take home and subjects seemed not to distinguish between new violins and old.

It is worth noting that these preferences were based solely on the experience of playing the instruments in a rather dry room. Though this raised numerous criticisms after the publication of the study, it was a deliberate choice as violinists consider an acoustically dry room best for initial try-outs, so the direct sound from the instrument is not so much coloured by room reflections. However, the question of how well judgements made in one room carry over into another (in particular a concert hall) is an interesting one. In addition, though players do routinely estimate projection, they typically acknowledge the need to re-test in a hall with trusted colleagues listening. Therefore, a new experiment, which took place in Paris in 2012, was conducted to address both of these issues, among others. In the first part of this study dealing with the player's point of view [72], 10 renowned soloists each blind-tested six Old Italian violins (including five by Stradivari) and six new duri ng two 1h15 sessions – the first in a rehearsal room, the second in a 300-seat concert hall. Like in previous experiments, participants wore modified welders goggles (which together with very low ambient lighting made it impossible to identify instruments by eye), used their own bows, and were encouraged to compare test violins with their own instruments whenever they wished. During sessions in the hall (Auditorium Coeur de Ville in Vincennes, renowned for its excellent acousti cs), soloists had the option of playing with piano accompaniment, getting feedback from a chosen listener, and hearing the violins played by another soloist. When asked to choose a violin to replace their own for a hypothetical concert tour, six of the ten soloists chose a new instrument. A single new violin was easily the most-preferred of the 12. While the soloists found the overall quality of the test instruments, both old and new, to be as high or higher than that of their own violins, they rated on average their favourite new violins more highly than their favourite old for *playability*, *articulation*, and *projection*, and at least equal to old in terms of *timbre*. They readily separated violins they liked from those they did not but were unable to tell new from old at better than chance levels. This emphatically confirms the findings of the Indianapolis experiment – and indeed many informal listening tests conducted over the years. Regarding the second part of the study dealing with the listener's point of view, data are still under analysis.

4. Discussion and recommendations for designing perceptual experiments

This section discusses some practical /empirical issues we encountered when contrasting listening and playing experiments in the context of acoustics research, and that are mainly related to psychological issues. While it is natural for engineers and acousticians to describe with great care the physical measurements and the experimental devices used to carry them out, their perceptual (psychological) studies are often described much more succinctly and and their procedures for questioning people are often just transferred from the procedures they use for physical measurements. However, questioning people (humans) differs largely from measuring the physical world. We will point here to some of the specificities of human functioning that have to be taken into consideration when running psychological evaluation/test/experiments. The data collected from a perceptual experiment result from the processes triggered by the psychological experimental device (stimuli, instructions, procedure). There is thus need for details about the experimental settings and procedures including the selection of the stimulations, the instructions given to subjects, the exact wording of the questions or scales, the language used in verbal data, the type of records, the number of subjects, in order to explicitly identify what the subjects had to process. This requirement for details applies as well to the characteristics of the subjects: age, gender, abilities and practices related to the object of concern (e.g. expertise, familiarity to the objects of concern and/or the tasks), inter-individual variability, ...

In addition, one main concern for researchers is that investigating the musician interaction with the instrument as an agent producing the sound introduces more variability in the loop between mechanical and perceptual processes. We will thus discuss how to control some of these sources of variations in designing experimental settings for psychological investigations.

4.1. Number of participants

With the exception of [32] and the studies presented in sections 3.2 and 3.3, the playing tests were generally conducted with only a few players (less than 10) if not only one or two. The reasons are quite obvious: the tests are thus much easier to run and the analysis of the data is facilitated as the inter-individual variabilities are smoothed (if not completely removed). However, the representativeness of such a small number of players can be questioned, especially if the player is one of the authors who knows the goals of the study, and one may wonder about the generalisation of the results to a larger sample of players. However, there is not necessarily a need for such generalisation. For instance, to know that a few players can distinguish between two factory instruments is enough to investigate the mechanical reason(s) for such a perceptual difference.

Besides, one may want to check first on a very few players whether some relationships (between psychological evaluations and physical measurements) can be found before conducting a large scale experiment.

And finally, one may not be interested in studying the perceptual assessments at an inter-individual level but rather at a intra-individual level: knowing the expertise and the experience of the player, his/her judgement made about different instruments in different conditions do not need to be potentially generalisable as a universal finding to be informative. Even, when the selection of participants relies on knowledge of their training/ skills/ experience/ expertise, i.e. on systematic inter-individual variables, the differential analysis of the different groups of subjects is informative for the understanding of their relation to the instrument (see the classical contrast in psychology of learning between so called “experts” and “novices”.)

4.2. Repetitions of the task

This second aspect is actually critical for psychological measurements. When designing a perceptual experiment aiming at evaluating instruments, acousticians often assume that a perceptual study is not reliable and valid without repetitions of the task (trials). However, if for physical measurements, precision is increased, and signal to noise ratio decreased, by repeating the measurement on the “same object” (at least in classical physics), it is because in most cases, each measurement is considered/supposed to be independent from the previous one. However, in psychology, it cannot be the case: the repetition of a stimulus makes it different just because it has been previously presented and therefore present in subject’s memory. As it is no longer “new”, it therefore induces different processing such as recognition (instead of identification), learning effects, etc., relying on top-down processing from (even short-term) memory. That requires considering each stimulus as dependent on its position in the sequence of presentation and to account for memory processes. Therefore, in contrast with repetitions of measurement in physics that are done to improve precision, the variations within repeated measures on the same participants are informative of the psychological processes under study, in terms of reliability/consistency of the participants or learning effects.

Generally speaking in psychology, and to summarise sections 4.1 and 4.2, inter- as well as intra-individual variations obtained for repeated measurements are not necessarily to be taken as “noise” – or “standard deviation” from systematic variations attributed to controlled variables – that can be reduced by increasing the number of participants or measurements (trials or stimuli) but as a psychological reality that can be interesting to investigate.

4.3. Number of instruments

When considering that the relevant variable is the interaction between an instrument and a musician, the same rationale about repeated measures can be applied to the number of instruments used in the study. In a playing test, one being trained in measuring physical phenomena may be tempted to use a large number of instruments in order to allow a large variety and therefore increase the generality of the study. However, there are two issues.

First, it is difficult to consider different instruments as just exemplars of the “same” instrument: musical instruments are considered, by players, as individuals that are different in their interaction with the players, even if one pretends to have equalised them with respect to the physical variables. This thus questions the extent of the generalisations that can be inferred from the set of instruments used in the experiment. Second, increasing the number of instruments changes the cognitive load, attention, tiredness, etc., of the participants, which will decrease their level of attention and the reliability of their evaluations, in contrast with measurements on physical devices which supposedly do not get tired nor change in their reliability across repeated (independent) measures. Here too, one way to escape this problematic issue is to adjust the experimental situation to approximate a real-life situation within which the processes under study are involved. In the case of buying an instrument for instance, players would rarely be presented with many instruments. So for an evaluation task (like rating instruments on different criteria or ranking them by order of preference), a limited number (less than five or six) may be preferable [59]. However players tend to quickly eliminate instruments they find unsuitable [72], so if the task it to choose their favorite(s), the number of instruments can be increased to ten or twelve as the actual number of instruments with which they deal to perform the task is in effect much lower.

4.4. Separate or joint evaluations?

A playing experiment is time-consuming and difficult to set up (in terms of organisation, loan of instruments, chasing participants, finding a venue, etc.). So within a study mainly concerned with the physical concerns, acousticians generally try to address as many related issues as possible in one single test (where psychologists may produce a series of experiments). When different playing conditions need to be investigated like in [63] and when each instrument needs to be evaluated on different criteria like in [63, 58], one way to minimize the duration of the experiment is to ask for separate evaluations of the instruments (i.e. one at a time), which indeed avoids participants going back and forth between instruments. Such different procedures lead to different knowledge on the subjects' preferences that may even be contradictory (from an *a priori* scientific physicalist/objectivist point of view). For example, it has been shown in behavioral decision research that people may exhibit reversal preferences between a joint evaluation mode (JEM) and a separate evaluation mode (SEM) of different options (e.g. [73]). Hse et al. [74] suggest that such reversals are due to difficult-to-evaluate attributes having a greater impact, compared with easy-to-evaluate attributes, in JEM than in SEM. Indeed, JEM is considered a more direct way of comparing the options, and may emphasize small differences, otherwise difficult to evaluate independently or even undetectable (in SEM). In sensory sciences, this opposition has been investigated as well by contrasting monadic and comparative procedures. It has been shown that the first procedure actually induces a comparative processing between exemplars within the sequence of stimuli during the course of the experiment, in which participants focus on finding out what makes them different (discrimination) while the other procedure makes the participants focus on similarities/commonalities between the exemplars and thus process them at another (higher) level of categorisation (see [75] for a systematic comparison of these two types of procedures).

4.5. From real life to the lab: ecological validity of experimental settings

If the goal of investigating perceptual evaluations made by players is related to a real-life context of choosing an instrument to buy/replace their own, it may then seem more appropriate to use JEM. However, JEM is not unconditionally better: a musician may have to go to different shops and thus may not be able to compare the different instruments together. Furthermore, at the time of the decision, a musician is typically exposed to different possible alternatives (JEM) but at the time of experiencing the consequence of the instrument he/she has chosen, the musician is usually in SEM. And therefore, the attributes which were important in the decision phase may have a totally different relative impact in the consumption phase. As it is quite common for string players to borrow instruments for a few weeks before buying it, one may thus wonder what is the most ecologically valid situation. The answer then lies in the goal of the study. If the goal is to understand how players choose the most valuable instrument, then JEM is more relevant. If the goal is to understand how players choose the instrument that will optimise their "consumption" experience, then SEM may prove better. Time constraints make the issue even more complicated. While an individual may spend a limited amount of time in a shop, he or she will certainly play the instrument during an extended period of time if he or she has the possibility to borrow it. A laboratory experiment being necessarily constrained in time, one may thus wonder whether musical instrument evaluations done in SEM may be fully valid when performed in a limited duration. A partial answer is provided in [58] in which players were considerably less self-consistent in the experiment where they had to rate all criteria for a given violin compared with the experiment in which they were able to compare the various violins to determine ratings for a given criterion. In short, there is no single "correct" procedure inasmuch each procedure triggers and informs different processes; the question is consequently how much the processes triggered within laboratory settings are similar to the processes triggered in the real-life situation that the researcher aims to explore. This point can be seen as a contribution in defining the ecological validity of experimental setting in laboratories as introduced in acoustics by Gaver [76] and later Guastavino *et al.* [77].

4.6. Labelling perceptual and physical properties

Perceptual evaluations are very often asked to be done on certain criteria labeled by names such as *brightness, timbre, ease of playing* or *overall quality* which are supposed to correlate with physical properties. However, there are several issues we became aware of by working in a different language than English and that we could make more explicit by collaborating with linguists.

Non English-native researchers have all experienced the difficulties of translating the criteria from their language to English for publication. Indeed, while the translation in English of scientific terms is explicitly regulated within a scientific community as referring to a well defined (scientific) concept, it is not the case for the semantics of common sense words and particularly words referring to subjective experience diversely lexicalised in different languages and which "objective" reference is not explicitly negotiated in their use by different people.

However, this difficulty is the same, and therefore shared by all researchers (though they may not be aware of it), when translating the criteria as formulated in the musicians' everyday language and vocabulary into the terminology of science. The meaning of a word does not directly point to a physical property "in the world", but depends on the language itself and on its use by different types of speakers in different types of discourses. Therefore one needs to work at identifying the semantics of the participants in their own language and discourse, through tools borrowed from linguistic research, and relate it to the lexical semantics of physics used in scientific discourse.

Given this control on the semantics of the words used by musicians, it has to be further evaluated how the criteria used by the musicians match, even under a different label, the criteria under investigation from physical hypotheses/considerations by the acousticians. For example, while the *sustain* is, for guitar players, regarded as an essential characteristic of an electric guitar and is supposed to be affected by a change of wood, Paté *et al.* [68] showed that it actually does not appear to be discriminative for differentiating between four guitars that are identical except for their fingerboards, two being made with rosewood, two with ebony. On the other hand, the criterion labelled as *précision* (translated as *precision*), as a result of a linguistic analysis of the transcribed interviews of the players (10 professional guitarists) after they had freely played the four guitars allows to make a clear distinction between the two woods.

One way to avoid this inescapable variability/instability of natural lexical semantics is to work at establishing a consensus on the meaning of words as labels. This is achieved by negotiating an explicit definition as a definite and common shared reference [78] as it is the case when defining any terminology, in science as well as in sensory analysis. Indeed, if they are left undefined, the same words may mean different things to different players and even more different to the acousticians. But if the definition has to be negotiated, it should not be based exclusively on the acoustical *a priori* point of view/definition but should also account for the sense given by the musicians community as experts.

4.7. Correlating perceptual and physical properties

Apart from these wording issues, the next issue for establishing a correspondence between physical properties and psychological criteria is precisely related to what the words refer to as knowledge or concepts (if not physical realities). In the evaluation tests conducted in natural sciences laboratories (trained within the psychophysics tradition in acoustics as well as in sensory analysis) an objective description (given by a panel expert or scientist) is often contrasted (opposed) to the hedonic, subjective judgement of preference (given by any consumer or lay person), along with the idea that preference evaluations should be avoided because of being too subjective. First, both evaluation types are actually subjective, meaning provided by a subject; but as answers to questions, they are objective, meaning as observable as any physical measurements/recordings from an instrument. Moreover, preference of the players is the real object of psychological investigation, as real as the physical reality for the physicist. It is therefore fundamental to study both types of realities along with their specific characteristics in each scientific domain (psychology and physics respectively). The first one grounded on experiential knowledge is individual, variable from one person to another one, holistic and multimodal; the second one is universal, analytic along abstracted (conceptual) dimensions apart from sensory qualities and validated by the community of scientists as the true description of a unique (physical) reality. Indeed, within our experience, Wollman *et al.* [63] found weak correlations between how violinists assessed the overall quality of five violins and how they rated and ranked theses five violins in terms of preference, and this under three different playing conditions. While a simple combination of the criteria used – assigning them equal weight – was found to be well correlated with the overall quality rating, no consistent association between preference and these eight criteria was found across violins. It thus seems that the two types of ratings are at best weakly related for the musicians at a psychological level. This has been observed in many sensory analysis studies [57] in which global judgements do not correlate with analytical judgements on abstract properties issued from the physical description of the world. This therefore questions the usual methodology used by acousticians and indicates that the type of rating– overall quality vs. preference – should be further investigated referring to knowledge already acquired in other sensory domains and even in acoustics [64] as well as carefully considered in designing an experiment that aims to study instrument evaluation by players.

4.8. Statistical data processing: correlations between judgements

In some studies (e.g. [59, 63]), what is of interest is how judgements on a series of instruments are affected by different conditions/tasks (for instance with or without auditory feedback [63] or with different constraints on what could be played [59]). One common way to evaluate this is to compute the (Pearson) correlation coefficient between two judgements made in two different conditions about the same set of instruments (for instance a mark attributed to each instrument with or without constraint on the style of playing). This set of instruments being usually very small in size (for the reasons explained in section 4.1.), it is often argued that such correlations are not meaningful and cannot be used in a statistical analysis. This is due to a confusion regarding the aim of the statistical inferences, i.e. the population to which the conclusions are to be applied. Here, the population of concern is that of players, from which the participants are considered to have been randomly sampled. A (derived) measurement is associated with each participant: the correlation coefficient between two evaluations of a fixed set of a limited number of instruments. The important word is 'fixed': these instruments cannot be considered as a random sample of the whole population of such instruments and no inference will be drawn on the instruments. Of course, for a given player, if the instruments had been different, it is likely the correlation would have been different too, but that is not what is being studied here. So there is no issue with the correlation coefficients being calculated based on a limited number of instruments as the inferences will concern the distribution of these coefficients in the population of players, for this specific set of instruments.

5. Conclusion

While research in musical acoustics originally focused primarily on instruments and their mechanical behaviour, the player has been more and more included in the last couple of decades, from a mechanical point of view as a system interacting with an instrument to produce music (not reducible to sounds) as well as from a perceptual point of view (as the evaluator of this interaction (i.e. the way the sound is produced) and the result of this interaction (i.e. the sound that is produced). Studying how the latter is evaluated may be achieved by means of listening tests, but only if what is of interest is the listeners' evaluation. If it is the player's evaluation, such tests are inadequate as the control the player has on the sound production can affect his/her perception of the sound produced. Obviously, playing tests are necessary if what is of interest is the evaluation of the interaction. In most cases in recent musical acoustics research, this is what is at stake, as understanding better how musicians evaluate the quality of an instrument is an essential step in order, for instance, to search for physical quality parameters and potentially offer improvements and new directions for the design and manufacturing of instruments. To this end, experiments designed in acoustic laboratories would benefit from relying on psychological research directly involved in designing adequate experimental settings, as well as on the expertise of players and makers. A genuine pluridisciplinary collaboration between engineers, psychologists, makers and musicians is the main key to success.

Acknowledgement

The authors would like to thank Joseph Curtin, Jacques Poitevineau and Charalampos Saitis for fruitful discussions as well as Timothy Wofford for his careful reading.

References

[1] C. Hutchins, V. Benade: Research papers in violin acoustics 1975-1993 with an introduction essay 350 years of violin research - volume 1. Acoustical Society of America, 1997.

[2] F. Ablitzer, J.-P. Dalmont, N. Dauchez: Static model of a violin bow: Influence of camber and hair tension on mechanical behavior. J. Acoust. Soc. Am **131** (2012) 773–782.

[3] K. Guettler, H. Thelin: Bowed-string multiphonics analyzed by use of impulse response and the Poisson summation formula. J. Acoust. Soc. Am. **131** (2012) 2903–2913.

[4] G. E. Gough: A finite element approach towards understanding violin structural modes. J. Acoust. Soc. Am. **127** (2010) 1791.

[5] M. Demoucron: Measuring bow force in bowed string performance: Theory and implementation of a bow force sensor. Acta Acustica united with Acustica **95** (2009) 718–732.

[6] E. Schoonderwaldt: The violinist's sound palette: Spectral centroid, pitch flattening and anomalous low frequencies. Acta Acustica united with Acustica **95** (2009) 901–914.

[7] G. Bissinger: Structural acoustics model of the violin radiativity profile. J. Acoust. Soc. Am **124** (2008).

[8] J. Woodhouse: On the "bridge hill" of the violin. Acta Acustica united with Acustica **91** (2005) 155–165.

[9] J. Woodhouse, P. Galluzzo: The bowed string as we know it today. Acta Acustica united with Acustica **90** (2004) 579–589.

[10] J. Alonso-Moral, E. Jansson: Input admittance, eigenmodes and quality of violins. Tech. Rept. STL-QPSR 2-3, Department for Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden, 1982.

[11] C. Hutchins: A measurable controlling factor in the tone and playing qualities of violins. Catgut Acoust. Soc. J. (Series II) **1** (1989) 10–15.

[12] H. Dünnwald: Deduction of objective quality parameters on old and new violins. Catgut Acoust. Soc. J. series II **1** (1991) 1–5.

[13] A. Langhoff, A. Farina, L. Tronchin: Comparison of violin impulse responses by listening to convoluted signals. Proc. of International Symposium on Musical Acoustics, Paris, 1995.

[14] G. Bissinger: Structural acoustics of good and bad violins. J. Acoust. Soc. Am **124** (2008) 1764–1773.

[15] H. Fletcher, E. Donnel-Blackham, R. Stratton: Quality of piano tones. J. Acoust. Soc. Am **34** (1962) 749–761.

[16] R. Pratt, J. Bowsher: The subjective assessment of trombone quality. J. Sound and Vibration **57** (1978) 425–435.

[17] G. Plitnik, B. Lawson: An investigation of correlation between geometry, acoustic variables, and psychoacoustic parameters for French horn mouthpieces. J. Acoust. Soc. Am. **106** (1999) 1111–1125.

[18] R. Caussé, N. Misdariis, S. Thoréton: Caractérisation objective de la qualité d'un instrument de musique. Proc. of Congrès Français d' acoustique, Marseille, France, 1997.

[19] E. Poirson, J.-F. Petiot, J. Gilbert: Study of the brightness of trumpet tones. J. Acoust. Soc. Am. **118** (2005) 2656–2666.

[20] E. Poirson, J.-F. Petiot, J. Gilbert: Integration of users perceptions in the design process: Application to musical instrument optimization. J. Mechanical Design **129** (2007) 1–9.

[21] L. Placido, A. Mamou-Mani, D. Sharp: Investigating perceptual differences between two trumpets of the same model type. Applied Acoustics **72** (2011) 907–914.

[22] G. Widholm, R. Linortner, W. Kausel, M. Bertsch: Silver, gold, platinum - and the sound of the flute. Proc. of International Symposium on Music Acoustics, Perugia, Italy, 2001.

[23] D. Sharp, H. Wright, W. Ring: An acoustical investigation into the effect of the crook profile on the sound produced by the bassoon. Acta Acustica united with Acustica **89** (2003) 137–144.

[24] F. Pinard, B. Laine, H. Vach: Musical quality assessment of clarinet reeds using optical holography. J. Acoust. Soc. Am. **113** (2003) 1736–1742.

[25] B. Gazengel, J.-F. Petiot, E. Brasseur: Vers la définition d'indicateurs de qualités d'anches de saxophone. Proc. of Congrès Français d' acoustique, Lyon, France, 2010.

[26] B. Gazengel, J.-F. Petiot, M. Šoltés: Objective and subjective characterization of saxophone reeds. Proc. of Acoustics 2012, Nantes, France, 2012.

[27] J.-F. Petiot, P. Kersaudy, G. Scavone, S. McAdams, B. Gazengel: Objective and subjective characterization of saxophone reeds. Proc. of International Congress of Acoustics, Montreal, Canada, 2013.

[28] R. Caussé, B. Goepp, B. Sluchin: An investigation on "tonal" and "playability"' of eight digeridoos, perceived by players. Proc. of International Symposium on Music Acoustics, Nara, Japan, 2004.

[29] J. Smith, G. Rey, P. Dickens, N. Fletcher, L. Hollenberg, J. Wolfe: Vocal tract resonances and the sound of the Australian didjeridu (yidaki). III. Determinants of playing quality. J. Acoust. Soc. Am. **121** (2007) 547–558.

[30] J.-F. Petiot, R. Caussé: Perceptual differences between cellos: a subjective/objective study. Proc. of International Symposium on Music Acoustics, Barcelona, Spain, 2007.

[31] A. Mamou-Mani, D. Sharp, T. Meurisse, W. Ring: Investigating the consistency of woowind instrument manufacturing by comparing five nominally identical oboes. J. Acoust. Soc. Am. **131** (2012) 728–736.

[32] A. Paté, J.-L. Le Carrou, B. Navarret, D. Dubois, B. Fabre: A vibro-acoustical and perceptive study of the neck-to-body junction of a solid-body electric guitar. Proc. of Acoustics 2012, Nantes, France, 2012.

[33] W. Hill, A. Hill, A. Hill: Antonio stradivari, his life and work. Dover Publications, New York, 1902.

[34] J. Schelleng: Acoustical effects of violin varnish. J. Acoust. Soc. Am. **44** (1968) 1175–1183.

[35] L. Burckle, H. Grissino-Mayer: Stradivari, violins, tree rings, and the Maunder Minimum: a hypothesis. Dendrochronologia **21** (2003) 41–45.

[36] B. Stoel, T. Borman: A comparison of wood density between classical Cremonese and modern violins. PLoS ONE **3** (2008) e2554.

[37] C. Barlow, P. Edwards, G. Millward, R. Raphael, D. Rubio: Wood treatment used in Cremonese instruments. Nature **332** (1988) 313.

[38] J. Nagyvary, J. DiVerdi, N. Owen, H. D. Tolley: Wood used by Stradivari and Guarneri. Nature **444** (2006) 565.

[39] C. Hutchins, A. Hopping, F. Saunders: Subharmonics and plate tap tones in violin acoustics. J. Acoust. Soc. Am. **32** (1960) 1443–1449.

[40] G. Dubourg: The violin: Some account of that leading instrument and its most eminent professors, from its earliest date to the present time : with hints to amateurs, anecdotes, etc. Robert Cox and Co., London, England, 1852.

[41] J. Marchese: The violin maker: A search for the secrets of craftsmanship, sound, and stradivari. HarperCollins Publishers, New York, USA, 2007.

[42] A. Coggins: Blind testing. The Strad Magazine (February 2007) 52–55.

[43] F. Schwarze, M. Spycher, S. Fink: Superior wood for violins - wood decay fungi as a substitute for cold climate. New Phytologist **179** (2008) 1095–1104.

[44] R. Inta, J. Smith, J. Wolfe: Measurement of the effect on violins of ageing and playing. Acoustics Australia **1** (2005) 25–29.

[45] C. Fritz, I. Cross, E. Smith, K. Weaver, U. Petersen, J. Woodhouse, B. Moore: Perceptual correlates of violin acoustics. Proc. of 9th Int. Conf. on Music Perception and Cognition, Bologna, Italy, 2006.

[46] C. Fritz, I. Cross, J. Woodhouse, B. Moore: Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin. J. Acoust. Soc. Am. **122** (2007) 3640–3650.

[47] C. Fritz, J. Woodhouse, F.-H. Cheng, I. Cross, B. Moore, A. Blackwell: Perceptual studies of violin body damping and vibrato. J. Acoust. Soc. Am. **127** (2010) 513–524.

[48] C. Fritz, A. Blackwell, I. Cross, J. Woodhouse, B. Moore: Exploring violin sound quality: Investigating english timbre descriptors and correlating resynthesized acoustical modifications with perceptual properties. J. Acoust. Soc. Am. **131** (2012) 783–794.

[49] D. Dubois, A. Giboreau: Descriptors: attributes? labels? terms? names? a contribution of psycholinguistics to sensory evaluations in mac fie h.j.h workshop summary : Developing, comparing and using consumer and sensory vocabularies. Food Quality and Preference **17** (2006) 669–672.

[50] E. Rosch: Natural categories. Cognitive Psychology **4** (1973) 328–350.

[51] J. Bensa, D. Dubois, R. Kronland-Martinet, S. Ystad: Computer music modeling and retrieval, lecture notes in computer science, volume 3310. Springer Berlin Heidelberg, 2005, Ch. Perceptive and Cognitive Evaluation of a Piano Synthesis Model.

[52] A. Cardenas: Corrélation entre les propriétés acoustiques et perceptives du violon. Diploma Thesis. UPMC - Univ. Paris 06, France, 2009.

[53] C. Fritz, A. Muslewski, D. Dubois: A situated and cognitive approach of violin quality. 20th International Symposium on Music Acoustics, Sydney, Australia, 2010.

[54] R. Pratt, J. Bowsher: The objective assessment of trombone quality. J. Sound and Vibration **65** (1979) 521–547.

[55] R. Caussé, J. Maigret, C. Dichtel, J. Bensoam: Study of violin bow quality. Proc. of International Symposium on Music Acoustics, Perugia, Italy, 2001.

[56] G. Weinreich: What science knows about violins - and what it does not know. Am. J. Phys. **61** (1993) 1067.

[57] M. Meilgaard, B. Carr, G. Civille: Sensory evaluation techniques. CRC Press, USA, 2006.

[58] C. Saitis, B. Giordano, C. Fritz, G. P. Scavone: Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players. J. Acoust. Soc. Am. **132** (2012) 4002–4012.

[59] C. Saitis, G. P. Scavone, C. Fritz, B. Giordano: Effect of task constraints on the perceptual evaluation of violins. Acta Acustica united with Acustica **101** (2015) 382–393.

[60] G. Bissinger, F. Gearhart: A standardized qualitative violin evaluation procedure. J. Catgut Acoust. Soc. Series II **3** (1998) 44–45.

[61] I. Wollman, C. Fritz, J. Poitevineau: Influence of vibrotactile feedback on some perceptual features of violins. J. Acoust. Soc. Am. **136** (2014) 910–921.

[62] I. Wollman, C. Fritz, J. Frelat: On the characterization of vibrotactile feedback in violinists' left hand: a case study. Acta Acustica united with Acustica **101** (2015) 360–368.

[63] I. Wollman, C. Fritz, J. Poitevineau, S. McAdams: Investigating the role of auditory and tactile modalities in violin quality evaluation. PlosOne **9**(12) (2014).

[64] D. Dubois (ed.): Le sentir et le dire. L'Harmattan, Paris, France, 2009.

[65] D. Dubois, C. Guastavino, M. Raimbault: A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories. Acta Acustica united with Acustica **92** (2006) 865–874.

[66] C. Guastavino: Categorization of environmental sounds. Canadian Journal of Experimental Psychology **61** (2007) 54–63.

[67] P. Cheminée, C. Gherghinoiu, C. Besnainou: Analyses des verbalisations libres sur le son du piano versus analyses acoustiques. Proc. of Conference on Interdisciplinary Musicology, Montreal, Canada, 2005.

[68] A. Paté, J.-L. Le Carrou, B. Navarret, D. Dubois, B. Fabre: Influence of the electric guitar's fingerboard wood on guitarists' perception. Acta Acustica united with Acustica **101** (2015) 347–359.

[69] C. Saitis: Evaluating violin quality: Player reliability and verbalization. Dissertation. McGill University, Montreal, Canada, 2013.

[70] C. Saitis, C. Fritz, C. Guastavino, G. Scavone: Conceptualization of violin quality by experienced performers.

Proc. of Stochkolm Music Acoustics Conference, Stockholm, Sweden, 2013.

[71] C. Fritz, J. Curtin, J. Poitevineau, P. Morrel-Samuels, F.-C. Tao: Player preferences among new and old violins. Proc. Nat. Ac. Sci. **109** (2012) 760–763.

[72] C. Fritz, J. Curtin, J. Poitevineau, H. Borsarello, I. Wollman, F.-C. Tao, T. Ghasarossian: Soloist evaluations of six old italian and six new violins. Proc. Nat. Ac. Sci. (published ahead of print April 7, 2014).

[73] S. Lichtenstein, P. Slovic: Reversal of preferences between bids and choices in gambling decisions. J. Exp. Psychol **89** (1971) 46–55.

[74] C. Hsee, G. Loewenstein, S. Blount, M. Bazerman: Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. Psychological Bulletin **125** (1999) 576–590.

[75] A. Giboreau, C. Dacremont, S. Guerrand, D. Dubois: Le sentir et le dire. L'Harmattan, Paris, France, 2009, Ch. Décrire: Identifier ou catégoriser?, 211–232.

[76] W. Gaver: What in the world do we hear? An ecological approach to auditory event perception. Ecological Psychology **5** (1993) 1–29.

[77] C. Guastavino, B. Katz, J.-D. Polack, D. Levitin, D. Dubois: Ecological validity of soundscape reproduction. Acta Acustica united with Acustica **91** (2005) 333–341.

[78] B. Keysar, D. Barr, J. Balin, T. Paek: Definite reference and mutual knowledge: Process models of common ground in comprehension. Journal of Memory and Language **39** (1998) 1–20.

[79] D. Dubois: Categories as acts of meaning: the case of categories in olfaction and audition. Cognitive science Quaterly **1** (2000) 35-68.

[80] D. Dubois: From psychophysics to semiophysics: Categories as Acts of Meaning: a case study from olfaction and audition, back to colors. In: M. Plumacher, P. Holz (Eds.): Speaking of colors and odors. An interdisciplinary approach to cognitive and linguistic categorization of color vision and olfaction, Benjamins, Amsterdam. 45-119.