acoustics

Article

Effects of Violin Back Arch Height Variations on Auditory Perception

Luca Jost [1,2,*], Mehmet Ercan Altinsoy [1] and Hannes Vereecke [2]

1 Faculty of Electrical and Computer Engineering, Dresden University of Technology (TUD), 01069 Dresden, Germany
2 Faculty of Applied Arts, Westsächsische Hochschule Zwickau (WHZ), 08056 Zwickau, Germany
* Correspondence: luca.jost.hwr@fh-zwickau.de

Abstract: One of the quintessential goals of musical instrument acoustics is to improve the perceived sound produced by, e.g., a violin. To achieve this, the connections between physical (mechanical and geometrical) properties and perceived sound output need to be understood. In this article, a single facet of this complex problem will be discussed using experimental results obtained for six violins of varying back arch height. This is the first investigation of its kind to focus on back arch height. It may serve to inform instrument makers and researchers alike about the variation in sound that can be achieved by varying this parameter. The test instruments were constructed using state-of-the-art methodology to best represent the theoretical case of changing back arch height on a single instrument. Three values of back arch height (12.1 , 14.8 and 17.5 mm) were investigated. The subsequent perceptual tests consisted of a free sorting task in the playing situation and three two-alternative forced choice listening tests. The descriptors "round" and "warm" were found to be linked to back arch height. The trend was non-linear, meaning that both low- and high-arch height instruments were rated as possessing more of these descriptors than their medium-arch height counterparts. Additional results were obtained using stimuli created by hybrid synthesis. However, these could not be linked to those using real playing or recordings. The results of this study serve to inform violin makers about the relative importance of back arch height and its specific influence on sound output. The discussion of the applied methodology and interpretation of results may serve to inform researchers about important new directions in the field of musical instrument acoustics.

Keywords: acoustics; violin; perception; back; arch height; hybrid synthesis; listening tests

check for updates

Academic Editor: Jian Kang

Received: 3 March 2025
Revised: 22 April 2025
Accepted: 6 May 2025
Published: 14 May 2025

Citation: Jost, L.; Altinsoy, M.E.; Vereecke, H. Effects of Violin Back Arch Height Variations on Auditory Perception. *Acoustics* **2025**, *7*, 27. https://doi.org/10.3390/acoustics7020027

1. Introduction

In the field of musical acoustics, the investigation of the relationship between geometrical properties and perceived sound is one of the most challenging tasks. However, its importance cannot be questioned, as it is only once the relation to perception has been uncovered that practitioners of the musical instrument making craft may draw practically applicable conclusions from a scientist's results. This article focuses on a single geometrical parameter of the violin—back arch height—and presents a series of investigations into its relation to perceptual outcomes. As such, only a small facet of a larger field can be hoped to be uncovered, contributing incrementally to the broader understanding of violin acoustic function. This study provides violin makers with practical advice that can be applied directly in their workshops. For researchers, methodologies for perceptual test design and inconsistencies in the status quo of acoustics research are discussed.

A body of literature has concerned itself with the connection between the physical properties of violins and their perceived sound. One factor often commented on by players is that of ageing. Many believe that a newly made violin first needs to mature and develop its sound. Inta et al. explored this phenomenon by comparing two violins [1]. One was played rarely and kept under controlled environmental conditions, while the other was regularly played by a professional violinist. At the onset of the experiment and again after three years, playing and listening tests were conducted. Ultimately, the study found no significant difference in preference between the two violins.

Another firmly held belief among many is that the old Italian instruments are superior [2]. In 2008, Bissinger categorized instruments into groups of "bad", "good" and "excellent" violins and rated the old Italian instruments at his disposal as "excellent" [3]. For these, detailed CT scans were available. No firm conclusions about the relationship between geometry and sound quality could be drawn. Examining the influence of the model on perceived sound, Fritz et al. conducted experiments in 2016 [4]. In a free sorting task, nine violins, two of them originals by Antonio Stradivari and one by Guarneri del Gesu, the rest modern copies, were evaluated by 21 violinists. While participants sorted the three del Gesu instruments in a group, the Stradivari models were not grouped together. No firm conclusions about the influence of the model on perceived sound output could be drawn.

The influence of strings and soundpost height were investigated by Fu in 2020 [5]. Furthermore, student- and performance-level instruments were compared in a series of perceptual tests. In general, the experiments faced large problems of inter-individual variability. When scrutinizing small physical changes such as changing soundpost height by a few tenths of a millimeter, these problems threaten to mask the actual signal that one is trying to detect. In 2019, Fritz et al. reported similar difficulties while trying to find the correlates of perceptual outcomes and construction parameters for a set of 25 violas [6].

More recently, in 2022, Nastac et al. reported on the outcomes of perceptual tests, conducted on a set of seven violins [7]. The instruments differed in terms of their plate thickness. It was found that instruments with thicker plates as compared to the chosen reference profile were preferred overall. Outside of the world of violins, similar methodology has been applied to the investigation of the steel-string guitar [8]. Here, the influence of bracewood and soundboard material properties on perceived sound was explored. The examined variables were density and Young's modulus. In a listening test using pairwise comparisons for preference ratings, low density and Young's modulus resulted in higher preference ratings. In a recent PhD thesis, Castrillo performed perceptual tests with string instrument bows [9]. It was found that players perceive changes in the mass distribution and adjust their playing accordingly.

For every perceptual test, appropriate stimuli are necessary. In the present case, six instruments were constructed experimentally using modern techniques (e.g., CNC machining and 3-D scanning) to best represent the theoretical case of changing back arch height on a single instrument without changing anything else. Similar methodology was applied previously by Fritz et al., in the so-called Bilbao Project [10], and Nastac et al. [7]. Only very limited research has concerned itself directly with acoustic outcomes of changing back arch height, and none has discussed perceptual outcomes [11,12]. As such, the presented investigation gives an initial insight into this parameter's influence on the instrument's perceived sound output.

The original research question was two-fold: Can participants distinguish between violins of different arch heights, and if so, how do they conceptualize the perceived differences? As such, the investigation was initially designed to test if participants could discriminate between different arch height violins in blind conditions and only then analyse their verbal descriptions of the differences they perceived. After a first test in the playing situation, it was decided to specify the test parameters further to allow participants to focus on individual properties of sound, in theory enabling higher task sensitivity. Therefore, relatively simple two-alternative forced choice (2AFC) protocols were implemented, in both live and headphone-based environments. Results provide a basis for arranging the six test instruments relative to each other in the perceptual domain. The test design was duplicated to ascertain if qualitatively similar outcomes would be observed in multiple independent instances.

Section 2 will present the test instruments. Section 3 details the applied method and obtained results for the first experiment, a free sorting task in the playing situation. Section 4 does the same for the self-experiment. In Section 5, the methodology of a live listening test and obtained results are discussed. Section 6 reviews the method and results of the online listening test using live recordings as stimuli. Section 7 does the same for the online listening test using stimuli obtained from hybrid synthesis. The content of these sections will then be discussed and summarized in Section 8.

2. Materials and Methods

Test Instruments

The six test instruments were designed and manufactured with the goal of acquiring a set of instruments that best represents the theoretical case of changing only back arch height on any given instrument without changing anything else. As such, they would have to be identical in all their properties except for back arch height, an impossibility outside of computer simulations. To achieve maximum similarity, backs, tops, necks, fingerboards and bridges were CNC-machined from wood selected specifically for similar material properties. Still, one prerequisite to obtaining useful data from the presented investigations is knowledge about error values associated with the methodology. As a solution to this, it was decided to produce the instruments as three matched pairs, representing three values of back arch height. In this way, residual differences between supposedly identical instruments could be quantified at each stage of the experiment. The procedure is illustrated in Figure 1. As can be seen, increments of 2.7 mm arch height change were chosen. The overall range from 12.1 to 17.5 mm well represents the practical range commonly encountered in violin making practice today. Some objective data about the material properties of top and back plates as well as outcomes of radiation measurements are included in Tables A1–A3 and Figures A1–A6 of the Appendix A and B. Realistically, a larger group of instruments would have been ideal to arrive at representative results. Within the constraints of this investigation, more instruments could not have been manufactured.

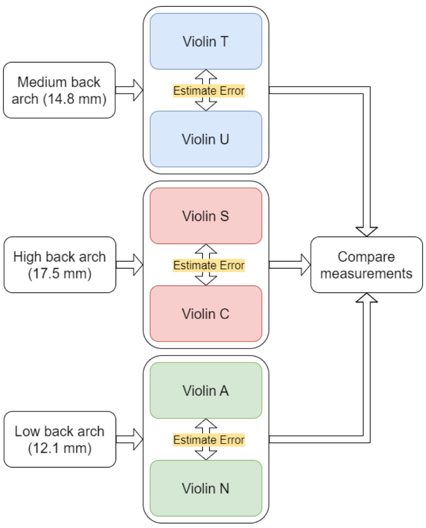

Figure 1. The procedure for constructing the six test instruments.

3. Experiment 1: Free Sorting Task

3.1. Method

For the free sorting task, a room of dimensions (4.12 × 5.59 × 2.83) m (volume 65.18 m^3, reverberation time 0.2 s, lowest axial mode 30.68 Hz) was available at the Erich Thienhaus Institute in Detmold, Germany. All six test instruments were placed on two tables covered in matte grey cloth and equipped with identical shoulder rests (Type Kun Original) adjusted to the same positions (see Figure 2). The instruments were tuned to a1 = 440 Hz, and a white piece of masking tape was stuck to the player-facing side of the chinrests, providing a number by which each instrument could be referenced during the test. Which violin received which label was randomized for each individual participant, while the experimenter could identify the instruments by way of pencil markings on the inside. Furthermore, the order in which the instruments were arranged on the table was randomized before each trial.

Figure 2. Test instruments as presented to the participants in the free sorting task.

All 11 individual sessions were recorded with the participant's express consent, and they were offered compensation of EUR 20 for their time, which three of them declined. Invitations for the experiment were sent out in German, but two participants preferred to communicate in English, which was accommodated by the experimenter. With the invitation, participants were instructed to bring their own instruments and bows. The instrument could serve as an absolute reference, for example, for the evaluation of room acoustics, which none of the participants were familiar with before the test. Personal bows were used to allow the musicians to focus on evaluating the instruments unfamiliar to them without also having to evaluate a new bow at the same time. Similar methodology was applied before by [13].

Before beginning each individual trial phase, some general data about the participants were collected. This included name, age, occupation, violin playing experience in years, violin practice per week in hours, hearing impairments as well as previous knowledge about the test and test instruments. Following this questionnaire, an as-far-as-possible neutral set of instructions was given to each participant for the actual trial phase. Here, a translation to English will be given, with the original German instructions being available in Appendix C:

- Please evaluate the instruments however you like and group them accordingly on the table.
- Similar instruments are to be grouped together; dissimilar ones are to be grouped apart.
- In general, you can create as many groups as you feel is necessary.
- Please try to verbalize your decision process during the experiment.

The actual assessment phase was conducted similarly to a typical buying decision in an instrument maker's workshop. The participant evaluated each instrument on their own terms while leading a dialogue with the experimenter. The experimenter supported the evaluation process by taking notes and on request reading them back to the participant. This was often necessary to remind the participant of assessments they had previously made. The experimenter carefully avoided influencing the actual outcome of the evaluation phase by providing only statements made by the participant themselves and not giving their own personal opinion.

The experiment was concluded once the participant was satisfied with the ordering of instruments they had achieved. Three final questions, aiming to quantify the perceived difficulty of the task, were addressed to each participant. Again, these are translated to English here, with the original German wording given in Appendix C:

- Did you encounter any specific difficulties in performing the experiment?
- On a scale from one to five, with one referring to very easy and five referring to very hard, how difficult was this task for you?
- We tried to constrain this task to the auditory dimension only. The other two possible dimensions that could influence the results are the visual and haptic dimensions. Please divide 10 points between these three according to how strongly you feel the result of the experiment was dependent on the individual dimension.
- In which way did you reach your result? Describe the groups you have created in your own words!

Adding to the data sheet collected for each participant, the duration of the assessment phase as well as the complete session and the final grouping result were recorded. Co-occurrence matrices will be used to represent these in the following, an example of which is given in Table 1. Each cell represents a possible pairing of two instruments from the set. A cell value of 1 is entered for an instrument pair, which was grouped by the participant; a cell value of 0 represents a pair of instruments that was not grouped together. The data entered in Table 1 represent the expected result, should the instruments be grouped perfectly by back arch height. In this case, three pairs of instruments, A and N possessing low back arch height, T and U possessing medium arch height and S and C possessing high arch height, would be created. As 11 individual co-occurrence matrices were collected during the live playing test, the overall result can be represented by a matrix in which each cell contains the average values over the 11 individual results.

Table 1. Co-occurrence matrix for expected result.

Pair	U	S	C	A	N
T	1	0	0	0	0
U	-	0	0	0	0
S	-	-	1	0	0
C	-	-	-	0	0
A	-	-	-	-	1

A total of 11 musicians participated in the free sorting task. Three of them were male and eight female, with an average age of 32 (Minimum 19, Maximum 58, Standard Deviation 13.8). Their self-reported playing experience was 26 years (Minimum 12, Maximum 50, Standard Deviation 13.1), and they reported playing the violin for an average of 11 h a week (Minimum 0, Maximum 35, Standard Deviation 12.3). On average, they completed the evaluation phase in 18 and a half minutes (Minimum 9, Maximum 28, Standard Deviation 6.7), and a single professional musician among them reported a slight hearing impairment for high frequencies in the left ear. The only selection criterion was their playing expertise. Participants adopted varied evaluation strategies. Some repeated the same phrase on each instrument; others played various pieces.

3.2. Results

The average co-occurrence matrix over all participants is shown in Table 2. Co-occurrence values for the expected pairs of the same arch height instruments have been highlighted. As can be seen, these are on or slightly below average, leading to no clear discrimination result linked to back arch height in this test. Overall, no clear groups of instruments emerge, which indicates that under these circumstances, the differences between instruments were too small to be reliably detected when considering sources of noise in the measurement such as inter-individual variability and stimulus variability due to the open test design, allowing players to evaluate the instruments on their own terms. Furthermore, a larger participant pool would be necessary to generate estimates that are more reliable. Individual co-occurrence matrices are given in Tables A4–A14 of the Appendix D.

Table 2. Average co-occurrence matrix for 11 participants. Those values representing same arch height pairs have been highlighted in red.

Pair	U	S	C	A	N
T	0.18	0.27	0.36	0.09	0.27
U	-	0.18	0.27	0.45	0.27
S	-	-	0.27	0.36	0.45
C	-	-	-	0.18	0.36
A	-	-	-	-	0.27

Semantic descriptors gathered from the recordings of all 11 individual sessions were categorized according to [14]. The 10 categories contained therein were supplemented with two additions of "preference" and "similarity". Overall, 268 comments divided into two groups were collected. These were relative statements (175) such as "instrument A is louder than instrument B" and absolute statements (93) such as "this instrument is loud". Results in the form of pie charts are shown in Figures 3–5 and show that most comments were related to texture and resonance of sound.

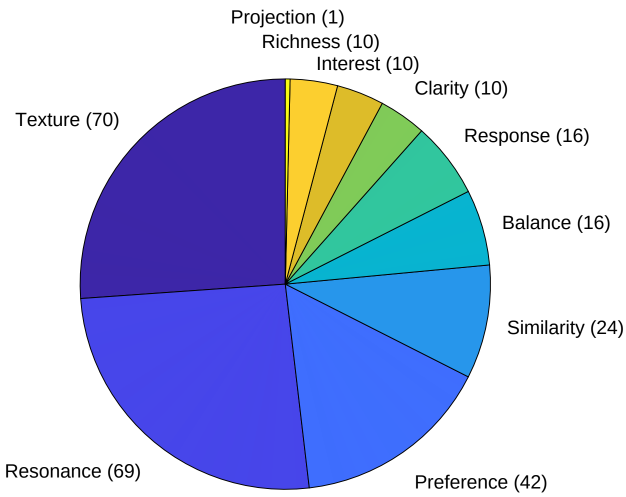

Figure 3. The distribution of semantic descriptors for all comments in the playing test categorized according to [14].

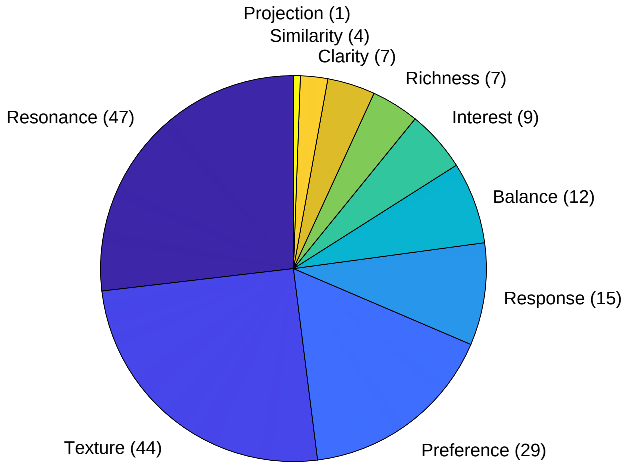

Figure 4. The distribution of semantic descriptors for absolute comments in the playing test categorized according to [14].

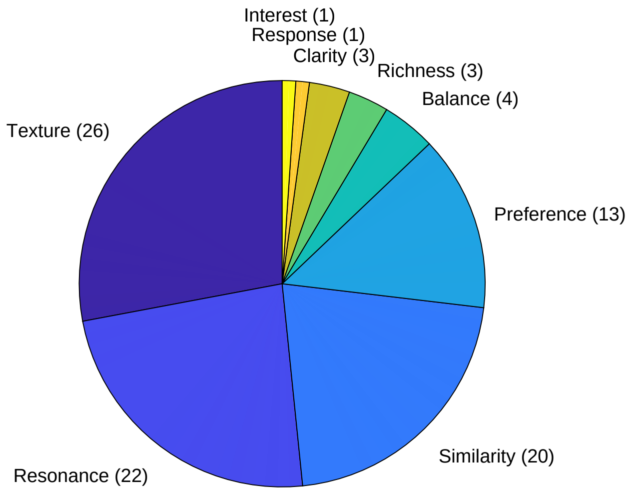

Figure 5. The distribution of semantic descriptors for comparative comments in the playing test categorized according to [14].

4. Experiment 2: Self-Experiment

4.1. Method

Scheduling 11 individual playing sessions with musicians resulted in a programme with many free slots between actual experiments. During this time, the principal author repeatedly performed the same sorting task as the participants and felt that, under specific conditions, he was able to categorize the instruments by arch height at a rate better than chance. A self-experiment was conducted to test this hypothesis. Playing only the open G-string and sorting the instruments on a scale from "boomy" to "flat", 27 individual orderings of instruments split into three pairs were collected. This procedure was chosen based on the answers from eight participants in the free sorting experiment. They adopted the strategy of first sorting the instruments on a scale according to some parameter and only then dividing this scale into groups. The scales adopted by them were round-pointed, round-sharp, dull-brilliant, soft-hard, warm-hard, round-bright (translated from German) and warm-bright, not bright-bright. The results of the self-experiment, while not immediately useful to the larger scientific community, were found to be valuable for further perceptual test design.

4.2. Results

The results of the self-experiment are presented in Table 3 as an average over 27 individual trials, the full results of which are summarized in Table A15 of the Appendix E. Again, the intersections of the same arch height instruments are highlighted and show the expected behaviour for discrimination by back arch height. This result provides supporting evidence for the idea that the principal author was able to distinguish between different arch height instruments under the specific conditions of the self-experiment. It therefore supports the hypothesis that a measurable signal exists, which was masked by test conditions in the free sorting task.

Table 3. The co-occurrence matrix in the self-experiment when playing only the open G-string. Those values representing same arch height pairs have been highlighted in red.

Pair	U	S	C	A	N
T	0.41	0.11	0.19	0.11	0.19
U	-	0.22	0.15	0.07	0.15
S	-	-	0.33	0.22	0.15
C	-	-	-	0.19	0.15
A	-	-	-	-	0.41

In contrast to the principal methodology of the free sorting task, in the self-experiment, the instruments were not simply grouped into pairs, but rather sorted on a scale, which was then divided into three pairs. This circumstance enables the application of some limited statistical analysis. Figure 6 summarizes the data as a box plot, quantifying the overall score attributed to each instrument. The three pairs are arranged in ascending order of back arch height on the X-axis, and a U-shape of the results can be observed. To test for a non-linear relationship between back arch height and the "boomy"-"flat" score acquired from the self-experiment, it was decided to calculate the distance correlation coefficient according to [15], which returned a value of 0.42. Ref. [15] recommends testing for significance by calculating the distance correlation coefficient for 10,000 random permutations of the data, which returned zero values exceeding the 0.42 threshold. Therefore, the result of the self-experiment can be seen as highly significant.

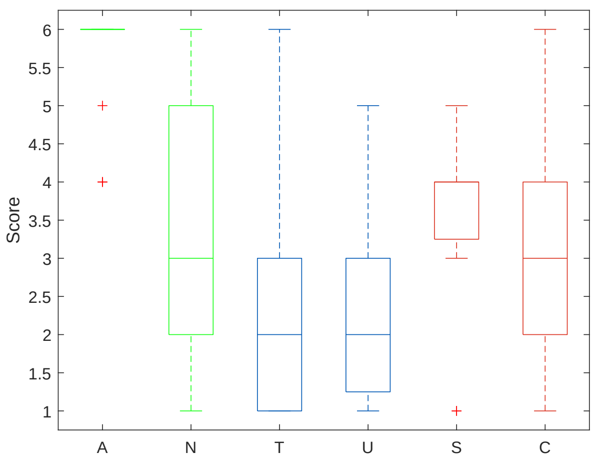

Figure 6. A box plot summarizing the results of the self-experiment, evaluating the boominess of the open G-string. Low arch height instruments are shown in green, medium arch height instruments in blue and high arch height instruments in red. Red plus symbols represent outlier values.

Clearly, the results of the self-experiment cannot be used as a valid assessment of the influence of back arch height on the sound of violins. First, there are several problems of the applied methodology, as the experimenter randomized the order between each individual trial themselves and possessed inappropriate prior knowledge of the instruments. Furthermore, a result produced by a single participant repeating the trial multiple times can only yield information about the perception of this individual person. Outcomes of this nature are of little practical usefulness. For these reasons, it should be clearly stated here that the self-experiment's sole use in this study was to inform the test design of the following perceptual experiments. Although proving the ability of a single person to differentiate instruments by a given semantic descriptor is not a directly useful result, it indicates that other participants may be able to do the same. Thus, it gives a well-defined hypothesis for testing in the following listening test.

5. Experiment 3: Live Listening Test

5.1. Method

In light of the results obtained from the free sorting task and subsequent self-experiment, it was decided to focus on a single perceptual descriptor in the live listening test. A two-alternative forced choice (2AFC) protocol was used. This methodology presents two stimuli to the participants and asks them to rate which of them possesses more of a given attribute. For comparative ratings of this nature, two avenues can be identified, which enable greater sensitivity as compared to pure discrimination tests. The first lies in constraining the task to a single perceptual attribute, which has been shown to increase inter-individual consistency [16]. The second is related to the specific strategy that participants may adopt to complete the task. According to Thurstonian modelling, during scale ratings, a more efficient "skimming" strategy may be employed, while for same-different ratings, a "comparison of distances" is required [17]. In this scenario, the "skimming" strategy is less likely to confuse two confusable stimuli, whose stimulus variability can be mapped as two overlapping normal distributions.

During the listening test itself, a professional musician played a short, slightly modified excerpt from the Glasunow Violin Concerto in A Minor Op. 82. (see Figure 7). She was situated on the stage of the Brahmssaal in the Hochschule für Musik Detmold, behind an acoustically transparent (100% cotton, $75\,\mathrm{g\,m^{-2}}$) curtain (see Figure 8). The room's acoustic properties are covered in detail in [18]. To impair her ability to adjust her playing to different instruments, she was wearing headphones emitting white noise throughout the test. This did not prevent her completely from hearing the violins she was playing. From six individual instruments, 15 possible pairs can be created, when excluding self-comparisons. All of them were tested in random order. Answer sheets in the German and English languages were provided to each participant, which they filled out during the test period. Along with general information (age, profession, instrument(s), experience in years), they answered a single question assessing their comprehension of the test protocol. For this, the musician played one of the instruments twice, the first time without and the second time with a heavy mute attached to the instrument. Participants were asked to rate which of the two sounded more muted to them. All 15 participants rated the second stimulus as sounding more muted, which indicated their understanding of the test protocol to the experimenter. Following this, eight comparisons were played by the musician, with the participants rating which of the sounds seemed more "round/warm" ("rund/warm" on the German answer sheets) to them. These descriptors were chosen since they had previously been used the most by participants in the playing test. Other descriptors used by the participants were as follows: hard, bright, clear, nasal, sharp, dull, small, dry, pointy, direct, scratchy, unnatural, brilliant, even, pleasant, tinny, closed, dark, introverted, loud, voluminous, warm, buzzing, blunt, musty, open, tense, strong, soft, full-bodied and tender. The variety of these is in agreement with similar lists previously discussed by [14]. After a break of approximately three minutes, the final seven comparisons were played and rated by the participants. The complete procedure was recorded using a condenser microphone (Type DPA 4011) connected to an interface (Type Focusrite Scarlett 2i2). Outcomes of comparative scale ratings were processed using a Bradley–Terry model, which was used to attribute ability scores to each instrument.

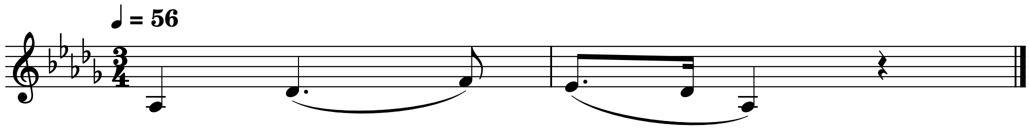

Figure 7. A slightly modified excerpt from the Glasunow Violin Concerto in A Minor Op. 82, used for the live listening test.

Figure 8. The setup for the live listening test with the musician situated behind an acoustically transparent curtain and wearing headphones playing white noise.

A total of 14 musicians (7 violin, 2 violoncello, 3 piano, 1 guitar and 1 trombone) and 1 non-musician participated in the live listening test, which was conducted in the Brahmssaal of the Hochschule für Musik Detmold. Their playing experience was on average 19 years (Minimum 15, Maximum 25, Standard Deviation 2.9). Musically inclined participants were invited. The single non-musician was a spouse of another participant. Each of them rated all 15 possible comparisons of the arch height instruments, deciding which of them sounded more "round/warm" to them. The decision to limit the evaluation to these semantic descriptors was based on the outcomes of the self-experiment, as described in the preceding paragraphs.

5.2. Results

Using the R statistics package "BradleyTerry2", an ability score was attributed to each instrument, representing the likelihood of it being rated as "warm/round" as compared to the rest of the group [19]. Here, for the reference required by the Bradley–Terry model, instrument U was chosen, to obtain positive scores of ability values for all instruments. Figure 9 visualizes the outcomes of this procedure as an error bar graph, while absolute values are presented in Table 4. Qualitative comparison to Figure 6 reveals the same

U-shaped trend, indicating a non-linear relationship between back arch height and the perceived "round/warm" score. However, as is to be expected with experiments of this kind, the attributed confidence intervals are quite large and do not allow for an entirely clear interpretation of the data. A larger number of participants would be required to achieve more reliable estimates. The inclusion of participants with no professional musical experience would be a worthwhile addition to future research. Full results of the live listening test are presented in Table A16 of the Appendix F.

Table 4. The ability scores with standard error derived from a Bradley–Terry model based on responses in the 2AFC listening test.

Instrument	Ability Score	Standard Error
T	0.10	0.31
U	0.00	0.00
S	1.20	0.33
C	0.47	0.31
A	0.80	0.32
N	0.14	0.31

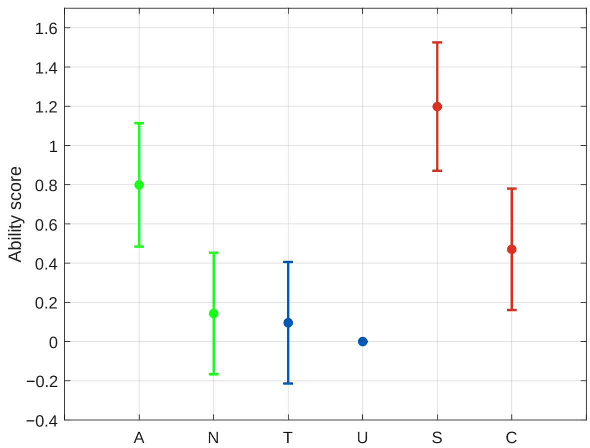

Figure 9. An error bar plot summarizing the results of a Bradley–Terry model based on the results of a 2AFC listening test protocol with 15 participants. Low arch height instruments are shown in green, medium arch height instruments in blue and high arch height instruments in red. Dots represent the ability score and lines the standard error.

6. Experiment 4: First Online Listening Test

6.1. Method

Since the results of the self-experiment and live listening test agreed qualitatively with each other but did not allow for firm conclusions on their own, it was decided to replicate the methodology of the live listening test in a further online listening test (see Figure 10). Refs. [20,21] previously showed the usefulness of this method. In this way, a larger number of participants could be included in the study. A total of 30 recordings of live playing were available as stimuli from the live listening test. These were set up in an online protocol using the listening test platform GoListen [22]. Stimuli were arranged in the same order as during the live listening test, and participants were instructed to rate the pairs of instruments according to the same question: "Which of these sounds more round/warm?".

Again, the test was set up in both the English and German languages, either of which could be chosen by each participant. They were asked to provide answers in quiet conditions and using high-quality headphones only. As before, information regarding age, profession, instrument(s) and playing experience was gathered before the evaluation phase began. Participants could listen to each of the two stimuli as often as they wanted. After the first eight trials, a pop-up message instructed them to take a short break. Invitations to the online listening tests were distributed among familiar groups of musicians, instrument makers and acoustics researchers.

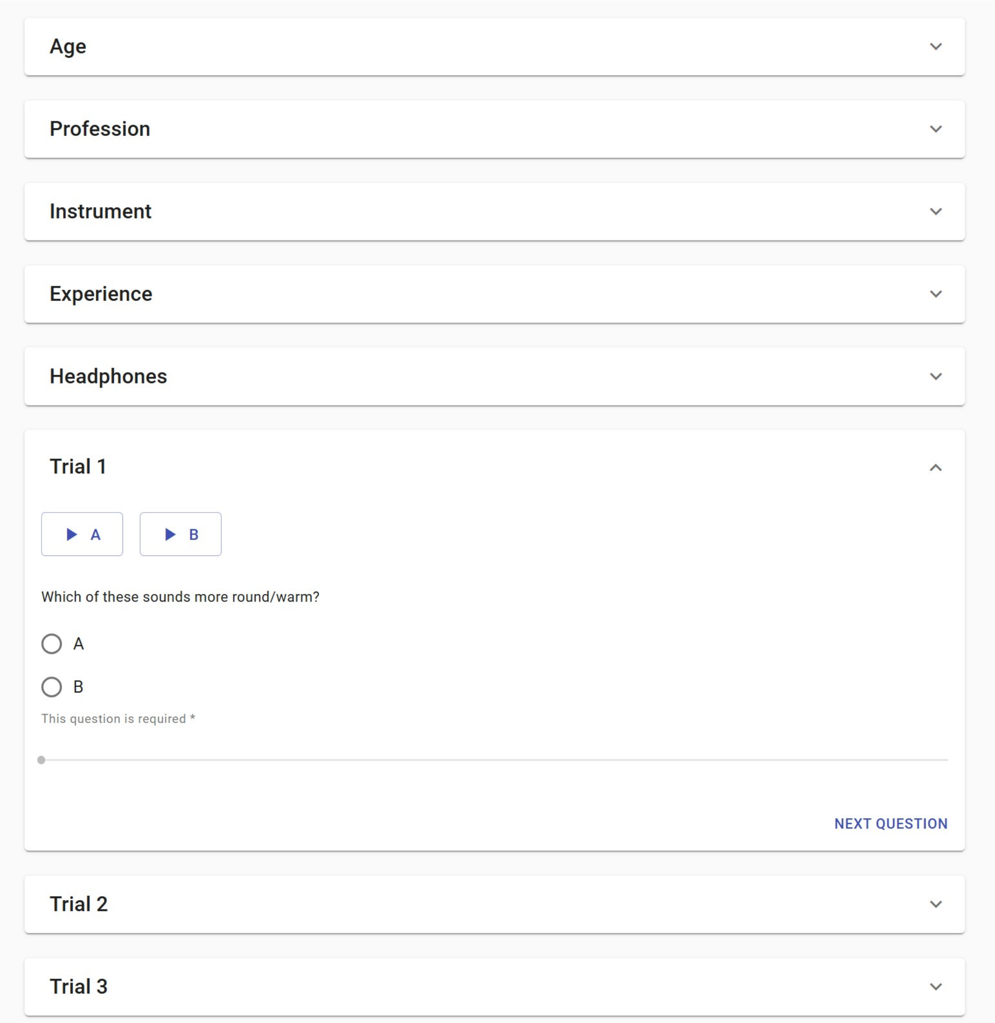

Figure 10. A screenshot of the interface of the online listening tests using the GoListen platform [22].

The first online listening test aimed to replicate the live listening test as closely as possible while reaching a larger number of total participants. A total of 26 listeners participated in the test, 25 of which reported playing an instrument (thirteen violin, five cello, three guitar, two viola da gamba, one viola and one trombone). Their average age was 46 (Minimum 24, Maximum 79, Standard Deviation 15.96) and average playing experience 27 years (Minimum 3, Maximum 58, Standard Deviation 16.89). As instructed, they used high-quality headphones, and 20 of them answered in English, with the remaining 6 participants preferring the German option.

6.2. Results

As before, Figure 11 shows the outcome of applying a Bradley–Terry model to the results of the listening test, and Table 5 provides absolute values of ability score and standard error. Here, instrument T was chosen as the reference, again to obtain positive ability score values for all instruments. A qualitative comparison to Figures 6 and 9 shows some agreement between all three individual results. Although each on their own is less than clear due to high values of standard error, the qualitative agreement between them provides a relatively strong basis for the assessment, in that back arch height and the perceived "roundness/warmth" of sound are correlated in a non-linear way. More reliable estimates would likely be obtained with a larger sample size of participants. Future studies would benefit from including evaluations of individuals without a professional music background. As before, full results of the first online listening test are summarized in Table A17 of the Appendix G. Comparing Figures 9 and 11, much smaller within-pair variations can be observed for the online test. With the available data, no clear explanation of this phenomenon can be given with any certainty. However, it could be speculated that the mono recordings used in the online listening test did not accurately reproduce some aspects of sound. These might have happened to be those that constituted the perceived differences within the pairs A-N and S-C in the live listening test.

Table 5. Ability scores with standard error derived from a Bradley–Terry model based on responses in the first online listening test, employing live recordings as stimuli.

Instrument	Ability Score	Standard Error
T	0.00	0.000
U	0.32	0.23
S	0.50	0.23
C	0.53	0.23
A	0.85	0.24
N	0.66	0.23

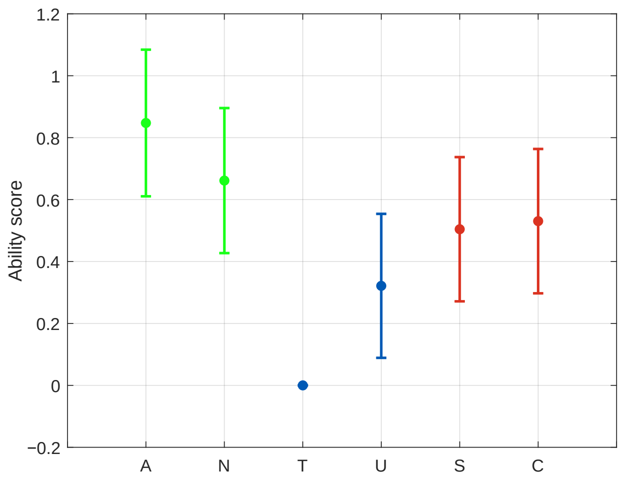

Figure 11. An error bar plot summarizing the results of a Bradley–Terry model based on the results of a 2AFC online listening test protocol with 26 participants utilizing live recordings as stimuli. Low arch height instruments are shown in green, medium arch height instruments in blue and high arch height instruments in red. Dots represent the ability score and lines the standard error.

7. Experiment 5: Second Online Listening Test

7.1. Method

The aim of the second online listening test was to generate comparative results using stimuli synthesized from radiation measurements of the experimental instruments. Usually, a human player is necessary to create stimuli for such experiments. However, this can introduce high levels of stimulus variability, since human performers are unable to repeat a phrase with perfect consistency. The hybrid synthesis method has been developed to address this problem [21,23]. Instead of a human player repeating the same phrase multiple times, a string input signal is recorded once. It is then played through filters that represent different acoustic measurement results. In principle, this should allow for strongly reduced stimulus variability and therefore improve the reliability of perceptual tests. However, few publications have tried to investigate the actual applicability of the method for this proposed purpose.

The second online listening test was set up in exactly the same way as the first, except for the actual stimuli, which were substituted with synthesized sounds. The basis for these were radiation measurements in semi-anechoic conditions. These were gathered, using a single front-facing microphone position at bridge height and horizontal excitation only. The measurement result was then convolved with a string input signal of the same phrase from the Glasunow violin concerto as used before and shown in Figure 7. While in the previous listening tests, individual signals were available for each instance of an instrument being played, resulting in 30 unique stimuli, only six individual signals were synthesized from the measurement results. Consequently, each of them was repeated a total of five times to arrive at the same 15 two-sided comparisons as before.

In the second online listening test, 19 participants of average age 44 (Minimum 24, Maximum 68, Standard Deviation 15.41) listened to the same comparisons of different back arch height instruments as before; however, the stimuli were generated using hybrid synthesis based on radiation measurements. A total of 18 of them were musicians (nine violin, two violoncello, two viola da gamba, two guitar, one viola, one trombone and one double-bass), and they reported, on average, a playing experience of 26 years (Minimum 3, Maximum 58, Standard Deviation 17.07). Most of them participated in the first online listening test and, as instructed, used the same high-quality headphones here.

7.2. Results

Resulting ability scores, obtained from a Bradley–Terry model, are shown in Figure 12, and the appropriate absolute values are given in Table 6. This time, instrument A was set to reference so that all ability scores were positive. When comparing the result to those of the two previous listening tests, no agreement of the methods can be concluded. For the specific application investigated here, hybrid synthesis cannot be considered a useful methodological approach. However, the shortcomings of the specific procedure followed here need to be considered when extrapolating from this outcome to the general viability of the hybrid synthesis method. As with all other experiments, increasing the number of participants would be necessary to arrive at more reliable estimates. Likewise, including assessments of individuals who have no professional background related to music would be a valuable addition in future investigations. All individual ratings of the second online listening test are presented in Table A18 of the Appendix H.

Table 6. Ability scores with standard error derived from a Bradley–Terry model based on responses in the second online listening test, employing synthesized sounds as stimuli.

Instrument	Ability Score	Standard Error
T	0.79	0.29
U	1.63	0.31
S	1.18	0.30
C	0.98	0.29
A	0.00	0.00
N	0.43	0.29

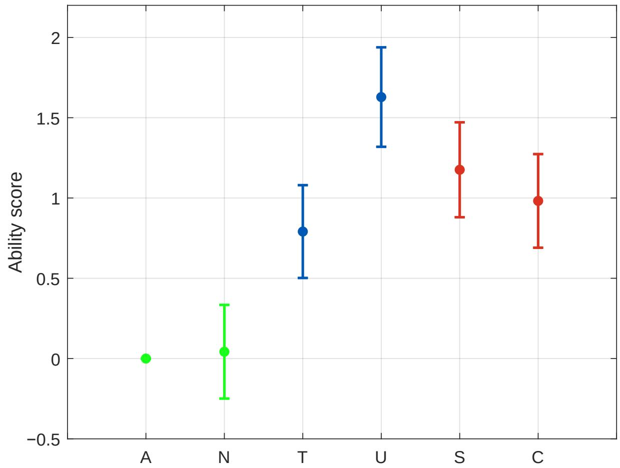

Figure 12. An error bar plot summarizing the results of a Bradley–Terry model based on the results of a 2AFC online listening test protocol with 19 participants utilizing synthesized sounds as stimuli. Low arch height instruments are shown in green, medium arch height instruments in blue and high arch height instruments in red. Dots represent the ability score and lines the standard error.

8. Conclusions

To evaluate the influence of back arch height changes on violin sound perception, six test instruments were constructed that represent the theoretical case of changing back arch height on a given violin, leaving all else untouched, as well as possible. Before discussing the individual perceptual tests, it is important to acknowledge that the measurements obtained in this study are relative, and individual differences in sound perception may influence the results. Furthermore, the room acoustics of the available spaces shape perceptual outcomes. The presented findings should be considered in this context and may not be generalizable to the broader population and independent of room acoustics.

In a first perceptual test, which was designed as a free sorting task in the playing situation, no discrimination by arch height could be observed. Some aspects of test design could be thought to explain this observation. The playing test did not restrict the musicians to evaluate a specific excerpt or semantic descriptor. While this prevents the experimenter's preconceived notions to influence the test results, it may also lead to large inter-individual variability. Additionally, when evaluating all aspects of a given violin's sound, one may miss a rather subtle difference in a specific area.

Using the principal author as a participant, a self-experiment was conducted to test if there was indeed a perceivable difference between the test instruments under specific conditions. Over the course of 27 individual trials, it could be clearly shown that the principal author was in fact able to—on average—sort the instruments by arch height. While this result cannot be used on its own, showing that one person could discriminate the instruments by arch height provides a basis for the assumption that under the right circumstances, others might be able to do so as well.

Testing this hypothesis, two perceptual tests following a 2AFC protocol in live and headphone-based listening situations were conducted. Totals of 15 and 26 participants, respectively, judged the instruments in pairwise comparisons in regard to the perceived "warmth/roundness" of sound. Both results were in qualitative agreement with each other and those from the self-experiment. While each of these on their own is hard to interpret due to relatively large standard error (see Figures 9 and 11 and Tables 4 and 5), their qualitative agreement provides some confidence in their common conclusions. Since, on average, the medium-arch height instruments were rated the lowest for "warmth/roundness" of sound, with the low- and high-arch height instruments achieving similar, much higher scores, a non-linear dependency on back arch height is suggested.

Two conclusions immediately useful to the violin maker may be drawn: when considering changing back arch height on a given model, the most significant expected change in sound perception is linked to the "warmth/roundness" of sound, and there may exist a turning point, where, e.g., further increasing arch height reverses its effect on sound. A third conclusion can be drawn from the circumstantial data of all the conducted trials. In none of the cases did participants rate the difficulty of presented tasks especially low, and none of the outcomes were entirely clear, indicating immediately noticeable and obvious differences between instruments. Since, in the workshop, more commonly, much smaller changes in back arch height than the 2.7 to 5.4 mm explored here are considered, this gives a much needed frame of reference. Based on the results presented here, a change in back arch height by, e.g., 0.3 mm, as may be practically necessary when working with limited back blank thickness, can be regarded as having no noticeable effects on perceived sound with some confidence.

Finally, another headphone-based listening test was conducted, in order to test the usefulness of the hybrid synthesis method for purposes such as those of the presented investigation. A total of 19 participants, most of which had participated in the first headphone-based listening test, rated the instruments in the same way as before. The only difference in the test setup was the stimuli, which had been synthesized from radiation measurements. These were conducted in the free–free boundary condition using a single microphone position and horizontal excitation. The outcomes of this procedure did not agree with any of the results of the other perceptual tests, leading to no clear applicability of the hybrid synthesis method for the purposes of this investigation. However, some clear shortcomings of the procedure as applied here, which are not intrinsic to the hybrid synthesis method and could well explain the presented results, need to be kept in mind. First, a single microphone position in nearly anechoic conditions does not represent the actual reverberant conditions of the listening tests used for comparison here well. Second, the free–free boundary condition of the radiation measurement affects the vibratory behaviour of the violin in a clearly different way than the boundary condition in the playing position as encountered in the listening tests. Third, simple horizontal excitation using an impact hammer might not sufficiently represent complex excitation by the bow.

Keeping all of these shortcomings in mind, no generalized conclusions about the usefulness of the hybrid synthesis method as a whole may be drawn from the presented data. Rather, it can be concluded that simplified methodology as applied here may not be sufficient to produce useful stimuli using convolution. To allow researchers to estimate the usefulness of this method for their work, more useful data input needs to be used for hybrid synthesis. For this, as alluded to before, measurements would need to be conducted in as close to the same acoustic situation as encountered during listening. The microphone would need to be placed in the listener's position in the same space (e.g., a concert hall) as used for the listening test. The musician would then be situated on the stage, holding the instrument in the same position as during playing. Now, the violin would simply need to be excited for a transfer function, representing as closely as possible the same circumstances as during the listening test, to be gathered. For this, an impact hammer could be mounted to the chinrest of the instrument, exciting the bridge at the bass side corner. Some preliminary tests of this procedure have been performed and delivered promising results for future investigations.

Author Contributions: Conceptualization, L.J.; methodology, L.J..; software, L.J.; validation, L.J.; formal analysis, L.J.; investigation, L.J.; resources, L.J.; data curation, L.J.; writing—original draft preparation, L.J.; writing—review and editing, L.J., M.E.A. and H.V.; visualization, L.J.; supervision, M.E.A. and H.V.; project administration, L.J.; funding acquisition, L.J. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for this research was provided through a full-time scholarship by the Westsächsische Hochschule Zwickau. The presented investigation was conducted in partial completion of a doctorate thesis at the Technische Universität Dresden. The Article Processing Charge (APC) were funded by the joint publication funds of the TU Dresden, including Carl Gustav Carus Faculty of Medicine, and the SLUB Dresden as well as the Open Access Publication Funding of the DFG.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: All related data are available from the author upon request.

Acknowledgments: Timo Grothe of the Erich Thienhaus Institute, Detmold, Germany, was instrumental in the organization of the presented perceptual tests. Claudia Fritz assisted the author by synthesizing sounds for the second headphone-based test.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
2AFC	Two-alternative forced choice

Appendix A. Material Properties of Spruce and Maple for Top and Back Plates

Table A1. The material parameters for spruce wood used to construct top plates. Velocities were deduced by Lucchi-Meter measurement and densities calculated by weighing the samples and approximating their volume from dimensional measures (relative humidity approx. 32 %). Values of standard deviation expresses as a percentage of the average have been highlighted in bold.

Name	ρ (kg m^{-3})	v_l (m s^{-1})	v_q (m s^{-1})	E_l (GPa)	E_q (GPa)
S1A	371	5897	2121	13.0	1.7
S1B	370	5915	2150	14.0	1.7
S2A	392	5958	2114	13.1	1.8

Table A1. *Cont.*

Name	ρ (kg m^{-3})	v_l (m s^{-1})	v_q (m s^{-1})	E_l (GPa)	E_q (GPa)
S2B	398	5958	2075	12.9	1.7
S3A	386	5910	2083	13.0	1.7
S3B	380	5921	2133	13.3	1.7
S4A	366	5867	2168	13.9	1.7
S4B	368	5827	2119	13.6	1.7
S5A	379	5931	2150	12.1	1.8
S5B	379	5910	2148	12.5	1.7
S6A	384	5921	2140	13.0	1.8
S6B	390	5933	2122	13.3	1.8
STDV	9.756	34.984	26.330	0.478	0.035
STDV %	**2.566**	**0.592**	**1.238**	**3.592**	**2.061**

Table A2. The material parameters for maple wood used to construct back plates. Velocities were deduced by Lucchi-Meter measurement and densities calculated by weighing the samples and approximating their volume from dimensional measures (relative humidity approx. 32 %). Values of standard deviation expresses as a percentage of the average have been highlighted in bold.

Name	ρ (kg m^{-3})	v_l (m s^{-1})	v_q (m s^{-1})	E_l (GPa)	E_q (GPa)
M1A	620	4586	2067	13	2.6
M1B	631	4714	2072	14	2.7
M2A	624	4584	2082	13.1	2.7
M2B	630	4520	2071	12.9	2.7
M3A	636	4528	2051	13	2.7
M3B	645	4546	2057	13.3	2.7
M4A	642	4654	1975	13.9	2.5
M4B	636	4622	2079	13.6	2.7
M5A	609	4460	2015	12.1	2.5
M5B	608	4538	2029	12.5	2.5
M6A	634	4535	2181	13	3.0
M6B	638	4571	2199	13.3	3.1
STDV	11.471	64.621	60.098	0.514	0.178
STDV %	**1.822**	**1.414**	**2.899**	**3.906**	**6.591**

Appendix B. Results of Radiation Measurements

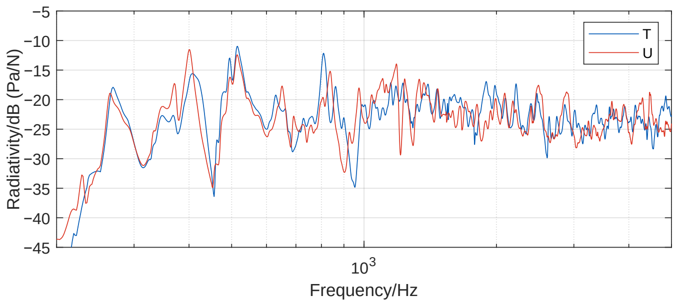

Figure A1. A comparison of instruments T and U (14.8 mm arch height) using the free–free boundary condition radiation measurement (0 dB $\widehat{=}$ 1 Pa N^{-1}).

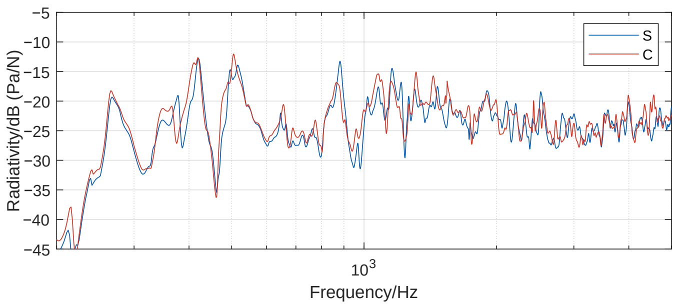

Figure A2. A comparison of instruments S and C (17.5 mm arch height) using the free–free boundary condition radiation measurement (0 dB $\widehat{=}$ 1 Pa N^{-1}).

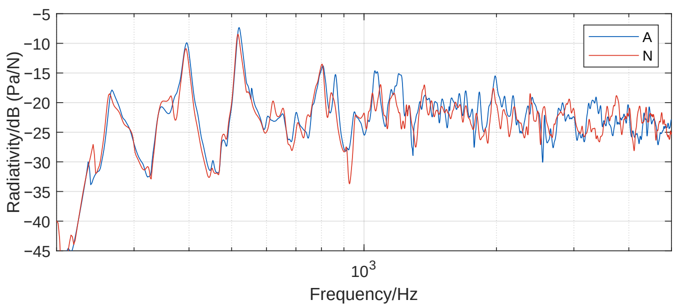

Figure A3. A comparison of instruments A and N (12.1 mm arch height) using the free–free boundary condition radiation measurement (0 dB $\widehat{=}$ 1 Pa N^{-1}).

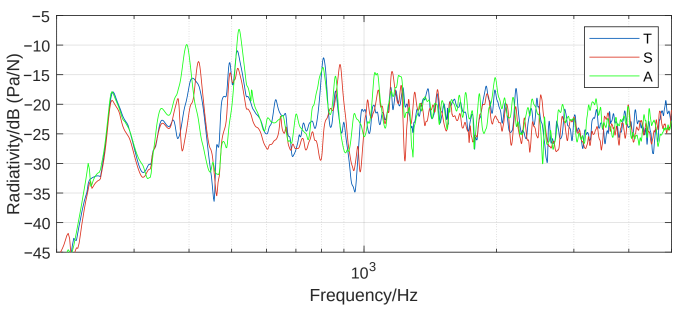

Figure A4. A comparison of instruments T (14.8 mm), S (17.5 mm) and A (12.1 mm) using the free–free boundary condition radiation measurement (0 dB $\widehat{=}$ 1 Pa N^{-1}).

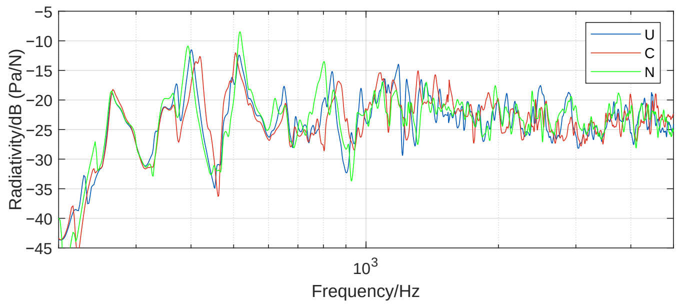

Figure A5. A comparison of instruments U (14.8 mm), C (17.5 mm) and N (12.1 mm) using the free–free boundary condition radiation measurement (0 dB $\widehat{=}$ 1 Pa N^{-1}).

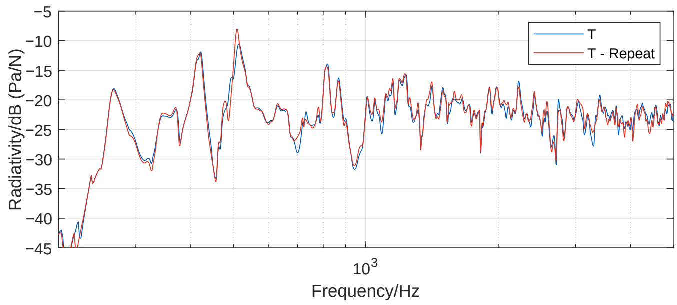

Figure A6. A repeatability measurement for instrument T three months apart, using the free–free boundary condition radiation measurement (0 dB $\widehat{=}$ 1 Pa N^{-1}).

Table A3. Selected features from [24,25] for the free–free boundary condition radiation measurements. Frequency values (denoted with f) are given in Hz and level values (L) in dB (0 dB $\widehat{=}$ 1 Pa N^{-1}).

#	f_{A0}	f_{CBR}	f_{B1-}	f_{B1+}	L_{A0}	L_{CBR}	L_{B1-}
T	268.8	346.9	407.8	515.6	−17.90	−22.70	−15.59
U	264.1	370.3	400.0	514.1	−18.92	−17.26	−11.54
S	267.2	378.1	420.3	515.6	−19.32	−19.07	−12.77
C	265.6	365.6	418.8	504.7	−18.22	−20.84	−12.64
A	267.2	346.9	395.3	520.3	−17.86	−20.70	−9.88
N	264.1	364.1	393.8	517.2	−18.50	−18.85	−10.87
#	L_{B1+}	$L_{M200-5000}$	ΔL_{Ausg}	ΔL_S	ΔL_N	ΔL_A	
T	−10.97	−21.89	3.34	−0.07	−1.93	1.47	
U	−12.41	−22.06	2.75	0.28	−0.87	2.67	
S	−13.93	−22.69	2.82	−0.14	−1.89	2.80	
C	−12.03	−22.13	3.71	−1.23	−2.56	2.33	
A	−7.34	−21.23	4.21	1.89	−0.91	2.04	
N	−8.46	−21.86	3.59	0.62	−0.10	1.14	

Appendix C. German Questionnaire for the Live Playing Test

- Ist es in Ordnung wenn ich die Aufnahme starte? Die Aufnahme an sich wird niemals mit jemandem geteilt werden, wenn ich etwas daraus wörtlich zitiere, dann nur mit Genehmigung und anoymisiert

- Können Sie einmal Ihren Namen, Alter und Beruf nennen?
- Seit wie vielen Jahren spielen Sie Geige?
- Wie viele Stunden pro Woche spielen Sie ca. Geige?
- Bestehen ärztlich diagnostizierte Beeinträchtigungen Ihres Hörsinns?
- Besitzen Sie Vorwissen über den Versuch?
- Können Sie dieses in eigenen Worten zusammenfassen?
- Evaluieren Sie die sechs Instrumente wie auch immer Sie möchten und gruppieren Sie sie dementsprechend auf dem Tisch
- Ähnliche Instrumente sind dabei in einer Gruppe zu platzieren, unterschiedliche Instrumente in verschiedenen
- Grundsätzlich können Sie so viele Gruppen bilden wie sie möchten
- Beschreiben Sie Ihren Entscheidungsprozess während dessen gerne verbal
- Hatten Sie Schwierigkeiten bei der Durchführung des Versuchs?
- Auf einer fünfstufigen Skala von 1 (sehr einfach) bis 5 (sehr schwer) als wie schwer empfanden Sie die durchgeführte Aufgabe?
- Wie sind Sie zu Ihrem Ergebnis gekommen? Beschreiben Sie die Gruppen die Sie gebildet haben in Ihren eigenen Worten!

Appendix D. Individual Results of Free Sorting Task

Table A4. Co-occurrence matrix for participant 1.

Pair	U	S	C	A	N
T	0	0	1	0	0
U	-	1	0	0	1
S	-	-	0	0	0
C	-	-	-	0	0
A	-	-	-	-	0

Table A5. Co-occurrence matrix for participant 2.

Pair	U	S	C	A	N
T	0	0	1	0	1
U	-	0	0	0	0
S	-	-	0	1	1
C	-	-	-	0	1
A	-	-	-	-	0

Table A6. Co-occurrence matrix for participant 3.

Pair	U	S	C	A	N
T	0	1	0	0	1
U	-	0	1	1	0
S	-	-	0	0	0
C	-	-	-	1	0
A	-	-	-	-	0

Table A7. Co-occurrence matrix for participant 4.

Pair	U	S	C	A	N
T	0	0	0	0	0
U	-	0	0	1	1
S	-	-	1	0	0
C	-	-	-	0	0
A	-	-	-	-	1

Table A8. Co-occurrence matrix for participant 5.

Pair	U	S	C	A	N
T	0	1	0	0	0
U	-	0	0	1	0
S	-	-	0	0	1
C	-	-	-	0	1
A	-	-	-	-	0

Table A9. Co-occurrence matrix for participant 6.

Pair	U	S	C	A	N
T	0	1	1	0	0
U	-	0	0	1	1
S	-	-	1	0	0
C	-	-	-	0	0
A	-	-	-	-	1

Table A10. Co-occurrence matrix for participant 7.

Pair	U	S	C	A	N
T	0	0	0	0	1
U	-	0	0	0	0
S	-	-	1	1	0
C	-	-	-	1	0
A	-	-	-	-	0

Table A11. Co-occurrence matrix for participant 8.

Pair	U	S	C	A	N
T	1	0	1	0	0
U	-	0	1	0	0
S	-	-	0	0	0
C	-	-	-	0	0
A	-	-	-	-	1

Table A12. Co-occurrence matrix for participant 9.

Pair	U	S	C	A	N
T	1	0	0	0	0
U	-	0	0	0	0
S	-	-	0	1	1
C	-	-	-	0	1
A	-	-	-	-	0

Table A13. Co-occurrence matrix for participant 10.

Pair	U	S	C	A	N
T	0	0	0	0	0
U	-	1	0	1	0
S	-	-	0	1	1
C	-	-	-	0	1
A	-	-	-	-	0

Table A14. Co-occurrence matrix for participant 11.

Pair	U	S	C	A	N
T	0	0	0	1	0
U	-	0	1	0	0
S	-	-	0	0	1
C	-	-	-	0	0
A	-	-	-	-	0

Appendix E. Full Results of Self-Experiment

Table A15. The individual results of 27 trials when sorting the six instruments in the main investigation on a scale from "boomy" (1) to "flat" (6), playing only the open G-string in the self-experiment.

#	1	2	3	4	5	6
1	A	N	S	T	C	U
2	A	C	N	T	U	S
3	A	T	S	C	N	U
4	A	N	C	S	T	U
5	A	C	S	N	U	T
6	A	U	S	N	C	T
7	N	A	U	S	C	T
8	A	S	C	N	T	U
9	A	N	S	U	C	T
10	A	N	S	C	T	U
11	A	S	T	C	U	N

Table A15. *Cont.*

#	1	2	3	4	5	6
12	A	S	T	N	U	C
13	T	U	A	S	C	N
14	A	C	S	U	N	T
15	A	N	S	T	U	C
16	A	U	S	C	N	T
17	A	S	U	C	N	T
18	C	S	A	N	T	U
19	A	T	S	C	N	U
20	A	N	U	S	C	T
21	A	S	N	C	U	T
22	A	N	C	S	U	T
23	A	N	S	C	U	T
24	A	C	S	N	U	T
25	A	N	S	C	U	T
26	A	T	N	C	U	S
27	A	C	S	T	U	N

Appendix F. Full Results of Live Listening Test

Table A16. Full results for live listening test. Rows show results for individual participants and columns those for individual trials. Bottom rows reveal identities of A and B as well as overall result for each trial.

#	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}
P_1	B	B	A	B	B	A	B	A	A	B	B	B	B	A	B
P_2	B	A	B	A	A	A	B	A	B	A	B	A	B	B	B
P_3	A	B	A	B	A	A	A	A	B	B	B	A	B	A	B
P_4	B	B	A	A	B	A	B	A	A	A	B	B	A	A	B
P_5	A	A	A	B	B	B	A	B	A	B	B	B	B	A	B
P_6	B	B	A	A	A	A	B	A	A	A	B	B	B	A	B
P_7	A	B	A	B	B	A	B	B	A	B	A	A	B	A	B
P_8	A	B	A	B	B	B	B	B	B	A	B	A	B	A	B
P_9	B	A	A	B	A	A	B	A	A	B	B	A	A	A	A
P_{10}	B	B	B	B	A	A	A	B	B	B	A	B	B	B	B
P_{11}	B	A	B	B	B	A	B	A	A	A	A	A	B	A	B
P_{12}	B	A	B	B	B	A	B	A	A	A	A	A	A	B	B
P_{13}	A	B	B	B	A	B	B	B	A	A	B	A	B	B	B

Table A16. *Cont.*

#	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}
P_{14}	B	B	A	A	A	A	B	B	B	A	B	A	A	A	A
P_{15}	B	A	B	B	A	A	A	A	A	A	B	A	A	A	B
I_A	U	T	A	U	U	S	A	C	C	S	T	N	N	C	C
I_B	S	A	U	T	N	N	S	N	U	C	S	T	A	T	A
R_A	5	6	9	4	8	12	4	9	10	9	4	10	5	11	2
R_B	10	9	6	11	7	3	11	6	5	6	11	5	10	4	13

Appendix G. Full Results of First Online Listening Test

Table A17. Full results for the first online listening test utilizing live recordings as stimuli. Rows show results for individual participants and columns those for individual trials. The bottom rows reveal the identities of A and B, as well as the overall result for each trial.

#	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}
P_1	B	B	A	A	A	B	B	B	A	B	B	A	B	A	B
P_2	B	B	B	A	B	A	B	B	B	B	B	A	A	A	B
P_3	A	B	A	A	B	B	A	B	B	B	A	A	B	A	B
P_4	B	A	A	B	A	A	A	A	B	A	A	B	A	B	B
P_5	B	A	A	B	B	A	A	A	A	A	A	B	A	B	A
P_6	B	B	B	A	B	A	B	A	A	B	B	A	B	A	A
P_7	B	B	A	A	B	A	B	B	B	A	B	A	A	A	B
P_8	B	B	B	B	B	A	A	B	B	B	A	A	A	A	A
P_9	A	A	B	A	A	A	B	B	B	A	A	B	B	B	B
P_{10}	B	A	A	A	B	A	A	A	B	B	B	B	A	B	B
P_{11}	A	B	A	B	A	B	B	A	B	B	B	A	B	A	B
P_{12}	B	B	B	A	B	B	A	B	A	B	A	B	B	A	A
P_{13}	A	B	B	B	A	A	A	B	B	B	B	B	B	B	B
P_{14}	A	B	A	B	A	B	B	A	B	B	B	A	B	A	B
P_{15}	B	B	A	B	B	A	A	B	B	B	B	A	A	A	A
P_{16}	A	B	A	A	A	A	A	B	B	B	B	A	B	A	B
P_{17}	B	B	B	B	B	B	A	B	B	A	B	A	A	A	B
P_{18}	B	B	A	B	B	B	B	B	A	B	A	B	B	B	B
P_{19}	A	A	A	B	B	B	B	B	A	B	A	B	A	B	A
P_{20}	B	B	A	B	A	B	A	B	A	B	A	A	A	A	A
P_{21}	B	B	A	A	B	A	A	B	B	B	B	A	B	A	A

Table A17. *Cont.*

#	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}
P_{22}	B	B	B	A	A	B	B	A	A	B	B	B	A	B	A
P_{23}	A	B	B	A	B	B	B	B	A	B	A	B	A	B	A
P_{24}	B	B	A	A	A	B	B	A	B	B	B	A	B	A	A
P_{25}	B	B	B	A	B	A	A	B	A	B	B	A	B	B	A
P_{26}	B	B	A	B	A	A	A	A	B	A	B	A	B	A	B
I_A	U	T	A	U	U	S	A	C	C	S	T	N	N	C	C
I_B	S	A	U	T	N	N	S	N	U	C	S	T	A	T	A
R_A	8	5	16	14	11	14	14	9	10	6	10	16	12	16	12
R_B	18	21	10	12	15	12	12	17	16	20	16	10	14	10	14

Appendix H. Full Results of Second Online Listening Test

Table A18. Full results for the second online test utilizing as stimuli sounds synthesized from free–free boundary condition radiation measurements. Rows show results for individual participants and columns those for individual trials. The bottom rows reveal the identities of A and B, as well as the overall result for each trial.

#	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}
P_1	B	A	B	B	A	A	B	A	A	A	B	B	A	B	A
P_2	B	B	B	A	A	A	B	A	B	A	B	B	A	A	B
P_3	A	A	B	B	B	A	B	A	B	A	A	B	A	A	B
P_4	A	A	B	B	A	B	B	B	B	B	A	B	B	B	A
P_5	A	A	B	A	A	A	B	A	A	A	B	B	A	A	A
P_6	A	A	B	A	A	A	B	A	A	B	A	B	B	A	A
P_7	A	A	B	A	B	A	A	B	B	A	A	B	B	B	B
P_8	B	A	B	B	A	B	B	A	B	B	B	B	B	B	B
P_9	A	B	A	B	B	B	B	B	A	A	B	A	B	B	A
P_{10}	A	B	B	B	A	A	B	A	A	A	B	A	B	B	B
P_{11}	A	A	B	A	A	A	B	A	B	A	B	A	A	A	A
P_{12}	A	A	B	B	A	B	B	A	A	B	B	A	B	B	B
P_{13}	A	B	A	A	A	A	B	A	B	B	B	B	B	A	A
P_{14}	A	B	B	A	A	A	B	A	A	A	B	B	B	A	A
P_{15}	B	A	B	A	A	B	A	A	A	A	B	B	B	B	A
P_{16}	A	B	B	A	A	B	B	A	B	B	A	A	A	A	A
P_{17}	A	A	B	A	A	A	B	B	B	B	B	B	B	A	A

Table A18. *Cont.*

#	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}
P_{18}	B	A	B	A	B	A	B	B	B	B	A	B	A	B	B
P_{19}	A	A	B	A	A	A	B	A	B	B	B	A	A	A	A
I_A	U	T	A	U	U	S	A	C	C	S	T	N	N	C	C
I_A	S	A	U	T	N	N	S	N	U	C	S	T	A	T	A
R_A	14	13	2	12	15	13	2	14	8	10	6	6	8	10	12
R_B	5	6	17	7	4	6	17	5	11	9	13	13	11	9	7

References

1. Inta, R.; Smith, J.; Wolfe, J. Measurement of the effect on violins of ageing and playing. *Acoust. Aust.* **2005**, *33*, 25–29.
2. Fritz, C.; Curtin, J.; Poitevineau, J.; Morrel-Samuels, P.; Tao, F.C. Player preferences among new and old violins. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 760–763. [CrossRef] [PubMed]
3. Bissinger G. Structural acoustics of good and bad violins. *J. Acoust. Soc. Am.* **2008**, *124*, 1764–1773. [CrossRef] [PubMed]
4. Fritz, C.; Curtin, J.; Dubois, D.; Verschoore, J. The violin—Studying the influence of the model by means of a free sorting task. In Proceedings of the International Symposium on Musical and Room Acoustics, Buenos Aires, Argentina, 11–13 September 2016.
5. Fu, L. Exploring the Perception of Violin Qualities: Student- vs. Performance-Level Instruments, Strings and Soundpost Height. Ph.D. Thesis, Mc Gill University Montreal, Montréal, QC, Canada, 2020.
6. Fritz, C.; Fraisse, V.; Dubois, D. The ObieAlto project: Looking for correlations between perceptual properties and constructional data. *J. Acoust. Soc. Am.* **2019**, hal-02362025.
7. Nastac, S.M.; Gliga, V.G.; Mihalcica, M.; Nauncef, A.M.; Dinulica, F.; Campean, M. Correlation between Acoustic Analysis and Psycho-Acoustic Evaluation of Violins. *Appl. Sci.* **2022**, *12*, 8620. [CrossRef]
8. Merchel, S.; Altinsoy, M.E.; Olson, D. Perceptual evaluation of bracewood and soundboard wood variations on the preference of a steel-string acoustic guitar. *J. Acoust. Soc. Am.* **2019**, *146*, 2608–2618. [CrossRef] [PubMed]
9. Castrillo, V.S. Influence of Bow Mechanical Properties on Violinist's Perception and Performance. Ph.D. Thesis, Sorbonne University, Paris, France, 2025.
10. Fritz, C.; Salvador, V.; Stoppani, G. The Bilbao project: Searching for relationships between sound and playing properties of violins with their construction parameters. In Proceedings of the Conference on Sound Perception, Poznan, Poland, 3–5 September 2021; p. hal-03446713.
11. Gough, C. Violin plate modes. *J. Acoust. Soc. Am.* **2015**, *137*, 139–153. [CrossRef] [PubMed]
12. Harris, N. On arching shape and violin tone Part 1: A qualitative analysis of the relationship between longitudinal string vibration, arching shape and violin sound. 2016.
13. Saitis, C.; Giordano, B.L.; Fritz, C.; Scavone, G.P. Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players. *J. Acoust. Soc. Am.* **2012**, *132*, 4002–4012. [CrossRef] [PubMed]
14. Saitis, C.; Fritz, C.; Scavone, G.P.; Guastavino, C.; Dubois, D. Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians. *J. Acoust. Soc. Am.* **2017**, *141*, 2746–2757. [CrossRef] [PubMed]
15. Székely, G.J.; Rizzo, M.L.; Bakirov, N.K. Measuring and testing dependence by correlation of distances. *Ann. Stat.* **2007**, *35*, 2769–2794. [CrossRef]
16. Saitis, C.; Scavone, G.P.; Fritz, C.; Giordano, B.L. Effect of task constraints on the perceptual evaluation of violins. *Acta Acust.* **2015**, *101*, 382–393. [CrossRef]
17. Lee, H.S.; O'Mahony, M. Sensory difference testing: Thurstonian models. *Food Sci. Biotechnol.* **2004**, *13*, 841–847.
18. Amengual, G.; ; Vincec, S. Investigations on the Influence of Acoustics on Live Music Performance Using Virtual Acoustic Methods. Ph.D. Thesis, Hochschule für Musik Detmold, Detmold, Germany, 2018.
19. Turner, H.; Firth, D. Bradley-Terry Models in R: The BradleyTerry2 Package. *J. Stat. Softw.* **2012**, *48*, 1–21. [CrossRef]
20. Atamer, S.; Altinsoy, M.E.; Zhang, Y. Advantages and Challenges of Online Listening Tests for Sound Quality Studies. In *Fortschritte der Akustik DAGA 2021*; DAGA: Wien, Austria, 2021; pp. 1103–1106.
21. Fritz, C.; Stoppani, G. Developing methodologies to correlate perceived sound qualities of violins with controlled construction parameters. In Proceedings of the Forum Acusticum 2023, Turin, Italy, 11–15 September 2023; pp. 236–240.
22. Barry, D.; Zhang, Q.; Wenyi, P.S.; Hines, A. Go Listen: An End-to-End Online Listening Test Platform. *J. Open Res. Softw.* **2021**, *9*, 20. [CrossRef]

23. Fritz, C.; Cross, I.; Moore, B.; Woodhouse, J. Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin. *J. Acoust. Soc. Am.* **2007**, *122*, 3640–3650. [CrossRef] [PubMed]
24. Ziegenhals, G. Subjektive und Objektive Beurteilung von Musikinstrumenten: Eine Untersuchung Anhand von Fallstudien. Ph.D. Thesis, TU Dresden, Dresden, Germany, 2010.
25. Jost, L. Establishing and Evaluating a Procedure for Measuring Violin Sound Radiation in Acoustically Challenging Spaces. Master's Thesis, Westsächsische Hochschule Zwickau, Zwickau, Germany, 2021.