MEASURING PERCEIVED DISTANCE OF VIOLINS – A DIRECT SCALING TEST

ROBERT MORES[1]

[1] *University of Applied Sciences Hamburg (HAW), Hamburg, Germany*
mores@mt.haw-hamburg.de

Specific violins are attributed to be acoustically intimate. A blind test is designed to answer the question, whether such intimacy can be measured in terms of perceived distance. The perceived distance is measured on some 24 subjects in a blind listening test while two violins are played which have already revealed some unspecific differences in terms of acoustical intimacy. A professional musician plays the violins on discrete positions of a physical scale, while subjects guess the sound origin in a blind test. To explore test design options and violins a few parameters are randomized such as the physical room, the loudness and the duration of samples. Additionally, intermediate voice references and continuous pink noise are investigated on whether these would possibly boost perceptual differences between violins. Subjects are screened and selected by quality measures for unreliability, discrimination and disagreement. The test delivers general results for human listening, as well as results for the usability of the test design. In terms of the investigated violins, there is little evidence to support the presumed differences. In conclusion, the perceived physical distance is not a prominent component of the acoustical intimacy of a violin.

INTRODUCTION

In the violin studio of Martin Schleske, Munich, we observed a few years ago an acoustical phenomenon: sounds from violins incorporating modified material where perceived more intimate than sounds from purely wooden instruments, both new and old valuable Italian violins. Observers reported that the sound, especially when played in pianissimo, was somehow touchable, located directly in front of oneself, whispering closely. The traditional instruments were also well audible but stayed in the back where they had been played. All violins had been played and listened to at identical locations in a 70qm² studio with some eight meters distance between musician and listeners. This observation was unexpected, because the focus in the ongoing innovation process has always been on maximizing acoustical output and on designing sound. However, the observed features are very desirable for soloist instruments, getting closer to the audience, while the orchestra stays behind. The question arises whether these features can be built into musical instruments in a determined way? In seeking such advantage for soloist instruments we investigated the acoustical properties causing such perception, but studies on radiation patterns, near-field - far-field transitions, energy distributions, phase coherence and other measures did not bring any conclusion, while further listening sessions still confirmed the perceptional distance. In order to guide the technical investigations, we started with studies on the perceptual process to understand causes for the observed perception. With proceeding listening sessions the most prominent cue reported was the difference in perceived distance while another prominent cue reported was the perceived difference of room size. Given these three aspects, perception of acoustical space, perception of a source location within the acoustical space, and various emotional attributes, using a general term such as acoustical intimacy seemed adequate to start with.

The question here is: Can the perceived acoustical intimacy of specific violins be measured in terms of physical distance?

Section 1 reviews relevant other studies while section 2 outlines the general approach of the study. Sections 3 and 4 describe two pretests, section 5 describes the main test and discusses results in the context of other work. Section 6 summarizes the findings.

1 DISTANCE PERCEPTION IN RESEARCH

Considerable research has been done in the field of distance perception, even though still much less than in the field of spatial hearing. Reviews of the literature can be found in [Bla83] and [Zah05]. There are also many studies on violins some of which get close to the subject while considering radiation patterns [Mey64]. No study has been found that addresses acoustical intimacy of violins or which cue would cause such perception.

In general, the research community concluded on several acoustic cues relevant for distance perception. These are: (i) intensity, usually investigated in environments, where loudness is the exclusive cue to distance (anechoic chamber of free-field), following the idea of the inverse-square law for sound propagation, (ii) direct-to-reverberant energy ratio, with conclusions for both, the fundamental relationship to distance perception in taking general measures of reverberant content [Sch54] as well as the relevance of the number of late reflections to distance perception [Bro99], (iii) spectrum, basically following the idea of absorption and diffraction of high frequency content, (iv) binaural cues as captured by measures of interaural time and level differences, ITD and ILD, concluding that these are most relevant only for distances up to 1 m, (v) other dynamic cues and non-acoustic cues.

Little work has been done in relating these cues to each other. There have been many studies on specific cues and only few studies on interaction. Intensity has long been considered the primary cue to distance [Tho92], today direct-to-reverberant energy measures are considered to be the prominent cue, and related modeling [Bro99] is the basis for sound design in recording studios. The outstanding study of Békésy considered both cues to understand human audition, concluding that measures of sound pressure and velocity are relevant to monaural distance estimations [Bek38]. Mershon concluded for these two prominent cues, that the direct-to-reverberant energy ratio provides absolute distance information, whereas intensity information must be compared relative to other presentations at different distances [Mer75]. Zahorik concludes that the direct-to-reverberant energy ratio might provide the rough absolute measure but does not necessarily facilitate discrimination, as he compared individual studies on this subject [Zah05].

This study does not intend to analyze relevant cues as such or to work on general models for distance perception. But it will keep in mind findings from other studies relevant to designing a test on perceptual distances between violins. These are:

a) Accuracy improves under reverberant conditions [Mer75][Nie93]. This finding suits well the intention of applying results to natural listening situations.
b) Intensity is not the most prominent cue [Nie93], and is a relative measure within sessions rather than an absolute measure [Mer75].
c) Intensity decreases the span of responses, regardless of accuracy [Nie93].
d) Interaction between intensity and reverberation content differs from room to room [Nie93].
e) Uncertainty grows with distance. This has been measured in terms of just noticeable differences [Edw55] but is also obvious with the reported span on responses [Nie93] [Loo98].
f) Large distances are generally underestimated and short distances are generally overestimated. See [Loo98] for converging results and [Zah05] for an extensive summary of past research on this.
g) Whereas it is clear that little head motions will improve spatial hearing [Bla83] it is not clear for distance location. Some studies reported improvements with head motion [Hol69] and some did not [Sim73].
h) Some methods include perceptually driven action and therefore facilitate responses based on every-day experience rather than asking for uncommon abstract responses [Loo98].
i) Noise in the listening process will bring sounds closer to the listener [Mer89].

Somehow disconnected from the cues reviewed above, the room acoustics expert Beranek only defines one cue if it comes to acoustical intimacy. In accordance to the visual cue, nearby walls would cause early reflections. More precisely, Beranek considers the initial-time-delay gap between direct sound and first reflection, ITDG, as the relevant cue [Ber04]. This is also relevant for the perceived distance, as ITDG will likely grow with distance and, surely, perceived distance will be biased by perceived room size. However, ITDG is not even mentioned in all the publications reviewed above.

2 TEST DESIGN

2.1 Method

This is the basic method used in both pretests and in the final test: a violinist plays at some of eight discrete positions on a line in a semi-reverberant. Listeners are asked in a blind test to guess the position of the sound origin, having the initial visual cue of the room and the numbered positions in mind. The direct scale ranges from a few meters to some 20 meters and positions are spaced with a potency function for several reasons. Variations to this basic method are (i) using a voice reference before and/or during the test, (ii) using a noise background in the listening area, (iii) using long and short passages of sound, and (iv) playing at two different intensities. After post selection, differences on the discrete scale are translated into a general factor for relative closeness.

2.2 Direct scaling – spacing and edges

In the test, listeners are asked to directly guess at which of the previously envisioned positions a sound originates. For this task, they keep the visual cue of the room and the numbered positions in mind. Such task suits the applicative context well as it is close to perceptually directed action, e.g. pointing at a source or approaching a source. The task therefore also challenges existing everyday orientation skills rather than specific estimation skills such as used in other studies where listeners are asked to respond by telling distances in meters. Even though literature recommends to use more than ten discrete entries on direct scales, here only eight entries are used in order not to overburden listeners. Statistical measures are used during pretests to develop a scale that is not too rough or that would avoid lock-in. For the spacing, a potency function is preferred against linear spacing. Distances are intentionally smaller between positions close to the listening area and are larger between positions at the far end, following approximately a quadratic relationship between discrete positions and physical distances:

$$d = d_0 + p \cdot (l+1)^q \qquad (1)$$

where d_0 is a few meters, $p < 1$ and $2 \le q \le 3$, are used to optimize spacing between the discrete positions l.

The study targets at a general difference in perceived distance across a wider scale, a general ratio of getting closer to the audience. Knowing from other publications, that uncertainty grows with distance, it is obvious that it would take more of a change in far locations than in close locations to trigger perceptual changes. That is why scaling is optimized over the sequence of pretests, such that far distant discrete steps would correspond to nearby discrete steps. This leads to the second reason: such correspondence also implies that the challenge for listeners is pleasantly homogenous across the entire scale. This will be confirmed by statistical results.

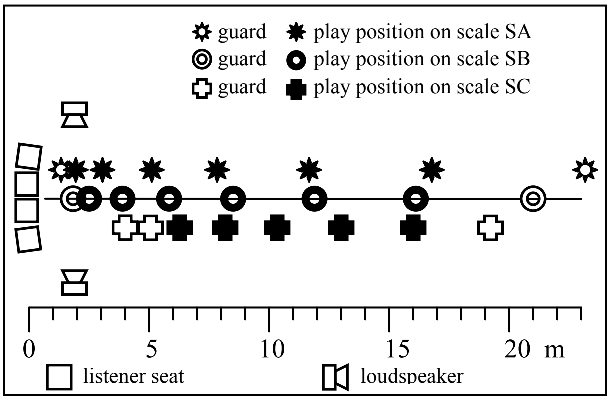

Figure 1: direct scales *SA*, *SB* and *SC* and their geometry

Edges on the direct scale imply problems. The limits of the scale imply restricted freedom for responding. This can be addressed by playing at inner positions only, leaving enough guarding space at the edges for individual response, see Figure 1. This, however, brings up the other problem of biasing, as listeners expect, that all positions will be used sooner or later. In order to balance between these two problems, the approach here uses (i) a small guard, (ii) random permutation, and (iii) sensible test instructions.

In this study, direct scales have been developed over pretests, see Table 1. Scale *SC* has finally been used in the main test and bases on d_o = 3.2 m, p = 0.2, q = 2 in formula (1).

scale	*SA*		*SB*		*SC*	
position l	d in m	type	d in m	Type	d in m	type
1	1.28	guard	1.80	guard	4.00	guard
2	1.92	play	2.52	play	5.00	guard
3	3.12	play	3.84	play	6.40	play
4	5.04	play	5.82	play	8.20	play
5	7.86	play	8.50	play	10.40	play
6	11.73	play	11.90	play	13.00	play
7	16.80	play	16.06	play	16.00	play
8	23.24	guard	21.00	guard	19.40	guard

Table 1: discrete positions of direct scales *SA, SB* and *SC* and their physical distances from listeners, all positions are eligible for response, but sounds only originate from positions of type play

The final translation from perceived differences on the discrete scale to a general factor for relative closeness is done along with an intermediate translation to the physical scale. Trusting the many findings of other studies that close by distances are usually overestimated and that far distances are usually underestimated, the generally applied potency function between perceived distance and physical distance is used here as well [Zah04],

$$d' = p \cdot d^q \qquad (2)$$

where the fit parameters p and q depend on listeners, stimulus and acoustical environment. Factor p ranges roughly from 1 to 3 and q from 0.2 to 1. These factors are found here by fitting the test results to the potency function with a least square means approximation. Factor p will finally be taken to represent the relative closeness, ignoring factor q.

2.3 Violins and other variables

Two violins are used. Violin *VA* is a nice but not superior German workmanship, 100 years old, conventional choice of wood, violin *VB* is an experimental work based on modified material. Both violins are well known in the research group, and are believed to embody the typical difference of perceived acoustical intimacy identified across several listening sessions and violins. Now, with the violin properties being the independent variable and the perceived distance being the dependent variable, the other variables to control or randomize are:

a) The acoustical room seems to be of great importance, as the various responses in earlier sessions included an influence to perceived room size. In earlier sessions the team also concluded that the differences were small in dry recording studio environments. Perceptual distances were larger in semi-reverberant rooms. However, remaining uncertainties did not allow for any conclusions on optimum criteria. Therefore, this study uses two rooms of different shape, but both well in the range of everyday listening experience and well suitable for listening to musical performance. Room *RA* is a hall, with a reverberation measure of $T_{60} = 1.2\ s$, and volume $V = 21\ m \times 16\ m \times 6\ m = 2016\ m^3$. The hall radius is therefore $r_{H,RA} = 2.34\ m$. Room *RB* is a laboratory, $T_{60} = 0.75\ s$, $V = 25\ m \times 13\ m \times 3\ m = 975\ m^3$, $r_{H,RB} = 2\ m$.

b) Intensity is set on two levels and controlled by instruction and experienced playing.
c) Directivity is controlled by instructing the musician.
d) Variations between sound samples are minimized by skilled playing.

2.4 Definitions

For clarity of discussion we denote the variables:

i = violins (total number $I = 2$)

j = individual listeners ($J = 4$ in a double session, and up to $J = 24$ across all double sessions)

k = sessions ($K = 2$ for a double session, $K = 2$ for the drift over a week)

l = positions ($L = 8$)

m = rooms ($M = 2$)

x_{ijklm} therefore represents a specific response from an individual listener j to violin i played in room m at position l in session k.

$x_{ij.lm}$ denotes the average response across sessions from an individual listener j to violin i played in room m at position l.

$x_{.j.lm}$ denotes the average response of listener j in room m given to all sounds originating at position l, across sessions and violins.

Likewise, replacement of a specific parameter by a dot defines the dimension across which the related data set is averaged.

The three additional variables background noise, intensity and length of sounds will be handled as session attributes during discussions.

The location of a violin serves as rough indicator for the perceived distance:

Location of violin i in room m $LOC_{i,m} = x_{i...m}$

Location of violin i across rooms $LOC_i = x_{i....}$

Position span PS_l denotes the standard deviation corresponding to the sum of squares of differences between individual responses and the average response $x_{i..lm}$ (responses of all subjects in all sessions to sounds from violin i at position l in room m), additionally averaged across violins and rooms.

$$PS_l = \frac{1}{I \cdot M} \sum_{i,m} \left[\frac{1}{J \cdot K - 1} \sum_{j,k} \left(x_{ijklm} - x_{i..lm} \right)^2 \right]^{1/2} \quad (3)$$

Likewise, the general position span *PS* for a chosen direct scale is the average PS_l across positions.

$$PS = \frac{1}{I \cdot L \cdot M} \sum_{i,l,m} \left[\frac{1}{J \cdot K - 1} \sum_{j,k} \left(x_{ijklm} - x_{i..lm} \right)^2 \right]^{1/2} \quad (4)$$

3 FIRST PRETEST

This pretest was done with two expert listeners from the research team, violins *VA* and *VB* in room *RB*, using the direct scaling set *SA* and a semi-professional musician, playing the first eight bars from the Ciaccona in Bach´s Partita in D minor for solo violin (BWV 1004). This piece of music provides plenty of broadband stimuli across all four strings, see Figure 2. Positions were permutated in a way as to investigate learning behavior and sensitivity to larger jumps along the scale. Additionally, in between pairs of violin sound samples, voice samples were presented by a speaker at nearby discrete positions on the same line, telling the true position such as to reference both violins against a stable scale. Voice is a useful type of sound source for this purpose because of its extensive training base in each of us. The impact of using a voice reference was additionally studied by using a voice-only learning phase prior to presenting sound samples in one of the sessions, similar as described in the second pretest, see Figure 3.

Figure 2: music for sound samples as used in the tests, top: Ciaccona from BWV 1004, bottom: Double from BWV 1002.

Results

Scaling – positions 1 and 8 are well represented among the responses even though they have never been played at. A guard interval seems necessary. For the general position span we find *PS* = 0.51. This indicates that the grid of scale *SA* is too widely spaced and response decisions are guided.

Voice reference - there is no significant difference in using a preliminary voice-based learning phase or not.

Violins – violin *VB* is located at LOC_{VB} = 4.29 whereas the conventional violin *VA* is located at LOC_{VA} = 4.50 (perfect responses would also average to *LOC* = 4.50). After approximation to the potency function (2), this location difference amounts to 8% (*p* factors for the two violins are 1.08 and 1.00). *VB* appears to be just a little bit closer to the audience. This result does not suit the expectation because the perceived difference of acoustical intimacy was much more obvious in earlier listening sessions. The voice anchor seems to be a too dominant cue, leaving little attention for listening to the process between violin and room. Factor q in (2) is 1.00 and 1.01 for the two violins, indicating that there is no such over- or underestimation effect across the scale.

4 SECOND PRETEST

This pretest was done with four expert listeners from the research team, using both violins *VA* and *VB* in both rooms *RA* and *RB*, using the direct scaling set *SB* and a professional musician, who played again the first eight bars from Bach´s Partita in D minor. Twelve sounds were presented per session, with a short break in between two consecutive sessions for a mandatory refresh of the visual cue. Positions were permutated for the different sessions using the MATLAB randperm function. The choice of the violin has also been randomized with MATLAB, only limiting the total number of changes between violins to eight per double session, see Figure 3. The voice anchor is now more sparsely used with only one voice reference for groups of four violin sounds. A double session is therefore preceded by the voice-based learning phase, Figure 3a), followed by the permutated sounds, sequence numbers 1 to 14 in Figures 3b), followed by a short break, and continuing with another short training, sequence number 15 to 19 in Figure 3a), and sound samples according to numbers 16 to 29 in Figure 3b) for the second half of the double session. Voice-less sessions do not use a training phase and have only the short break in between. A permutation example is given in Figure 3c). Each panel session is newly permutated.

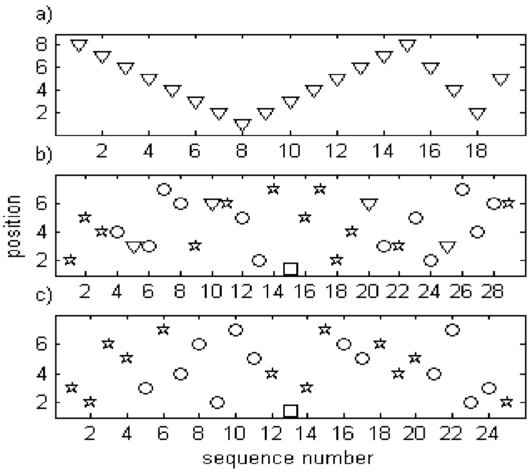

Figure 3: examples of permutated playing positions on scale *SB* and permutated choice of violin for the second pretest a) voice-based learning phase, b) permutated sounds with voice anchor, c) permutated sounds without voice anchor; triangle: voice, star: violin *VA*, circle: violin *VB*, square: silence for a minute break, effectively dividing a sequence into a double session

Results

Scaling – positions 1 and 8 are again well represented among the responses. The general positions span now is *PS* = 0.64, slightly more than in the first test due to the slightly denser direct scale *SB* versus *SA*. However, the grid still seems to be too widely spaced and response decisions might still be guided.

The voice reference slightly amplifies existing perceptual distances, see Table 2. However, the voice reference does also shift the location for both instruments. This shift is much larger than the perceptual distances becoming visible.

For the violins, the general location difference is still very small compared with earlier experience in listening sessions.

The team reported a shift of focus during the perceptual decision process. Spontaneous response during the onset of the sound would differ from a more rational response after listening for a while. This leads to variation of sound duration for the main test.

The team also reported biasing through special resolution for border seats when the sounds origins from nearby positions. In conclusion for the main test, the entire direct scale should shift further to the far end.

room	*RA*		*RB*	
voice anchor	no	yes	no	yes
LOC_{VA}	4.71	4.31	4.56	4.56
LOC_{VB}	4.69	4.48	4.56	4.75
Difference	0.02	-0.17	0.00	-0.19

Table 2: average locations LOC_{VA} and LOC_{VB} for violins *VA* and *VB*, normalized to the non-linear scale *SB*

5 MAIN TEST

5.1 Test outline

The main test was done with 24 untrained listeners from the department, 6 female and 18 male, 12 non-musicians and 12 musicians, all aged between 25 and 35 years. Again the two violins are compared in both rooms and played by the same professional musician. The direct scale *SC* has a denser spacing and a larger guard at its near-field end, see Figure 1 and Table 1. Sessions were repeated after one week with a permutation of listeners across tasks, panels and rooms.

Apart from the necessary changes to the test design, additional session attributes were introduced for variation. Pianissimo passages had to be used due to experiences in earlier listening sessions. For this purpose, another piece of music was chosen, the first eight bars from the Double following the Sarabande in Bach´s Partita in B minor for solo violin (BWV 1002), see Figure 2. Another variation is adding noise in the listening area. This variation is driven by other studies on grand piano sounds, where the team noticed that some of the perceptual distances were only significant when orchestral sound was present simultaneously. Yet another variation is the length of the musical passage used. For an overview of session attributes see Table 3, for the allocation of these attributes across sessions see Table 4.

session attribute	sound / noise	length in sec.	dB SPL at listener
ff long	Ciaccona, fortissimo	27	75
pp long	Double, pianissimo	20	60
pp short	2 octaves B-chord, pianiss.	1	60
pn40, pn45	analog pink noise	contin.	40 or 45

Table 3: session attributes used in the main test, the musical notations ff and pp are used to denote the fortissimo and the pianissimo play, for Ciaccona and Double see Figure 2

sound	pp long	pp long	pp long	pp long	pp short	pp long	ff long
noise	pn45	pn40	-	-	-	-	-
1st week	○	○		○	○	○	○
2nd week	○		○	○	○	○	○

Table 4: allocation of sessions attributes across panels for the main test, using sound and noise attributes according to Table 3, each connected pair of circles represents a comparison task for a panel of four individuals

Each entry of a connected pair of circles represents the task set for a panel of four listeners. This allocation allows to directly investigate individual session attributes within a panel, i.e., whether an attribute can amplify perceptual differences and whether locations of violins will shift. For instance, one panel investigates the impact of keeping sound samples short, another panel compares forte against pianissimo passages. In total, there were twelve panels, three for each room and each week. This allocation results in 24 double sessions and a total of 96 individual data sets.

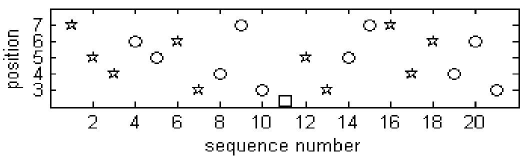

Figure 4: example of a permutated sequence of playing positions on scale *SC* and permutated choice of violin for double sessions in the main test; star: violin *VA*, circle: violin *VB*, square: silence for a minute break, effectively dividing a sequence into a double session

Figure 4 illustrates one example of the six machine permutations of sequences used. Due to the smaller set of positions in scale *SC*, the sequences become shorter, too. Again, a double session contains a short break for a refresh of the visual cue. In total, 480 sound samples resulted in 1920 individual responses.

5.2 Session notes

During the sessions, intensity was measured for control. Figure 5 summarizes the results for some of the measurements from the second week. Measurements were taken with Brüel & Kjaer 2237 at the listening position 0 m on axis, see Figure 1, at 1.80 m above ground. The target level was well met in the middle of the scale, and the level was kept constant over the sessions.

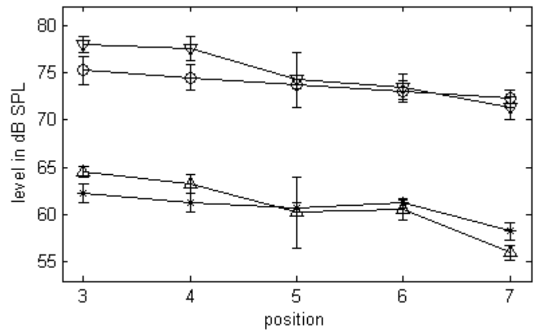

Figure 5: averaged sound levels and deviations of violin sounds in dB SPL across positions, measured in the 2nd week of the main test with B&K 2237 at listening position 0 m on axis, see Figure 1, circle: ff in room *RA*, triangle down: ff in room *RB*, star: pp in room *RA*, triangle up: pp in room *RB*

The optional noise level was set to 40 ± 1 dB or 45 ± 1 dB as measured at the four listening seats, see Figure 1. The noise source is an analogue noise generator (Sennheiser SmartNoise) set to pink noise and amplified by two speakers (Genelec 1029), positioned as illustrated in Figure 1.

Listeners were blindfolded and had to write down the position they believed a sound originates from. After the test they were interviewed on musical skills and remarks on the test. They reported from additional acoustical cues other than from the violin in only few cases. This is due to careful preparation and instruction. The musician was wearing specially designed shoes to prevent noise during the walks, and was instructed to avoid additional noise in general. The permutation seemed to work well as many listeners did not even notice that there were two violins involved. Listeners were instructed that the permutation is random with no influence from an instructor, supporting the idea that every sound is a new surprise to be responded to disregarding previous history.

Figure 6: room *RA*, volume V = 2016 m³

Figure 7: room *RB*, volume V = 975 m³

5.3 Postselection

For the purpose of post selection, three quality measures for listeners were considered according to the model of Schlich [Sch94]. Given the sum of squares across violins for individual listeners *j* in room *m*

$$SSV_{jm} = K \cdot \sum_{i,l} \left(x_{ij.lm} - x_{.j.lm} \right)^2 \qquad (5)$$

and the residuals across violins and double sessions for listener *j* in room *m*

$$SSR_{jm} = \sum_{i,k,l} \left(x_{ijklm} - x_{ij.lm} - x_{.jklm} + x_{.j.lm} \right)^2 \qquad (6)$$

and the span of responses from a listener

$$Span_{jm} = \frac{1}{K} \sum_{k,l} \left[\frac{1}{I-1} \sum_{i} \left(x_{ijklm} - x_{.jklm} \right)^2 \right]^{1/2} \qquad (7)$$

we obtain

$$UNRELIABILITY_{jm} = \frac{RMSR_{jm}}{Span_{jm}} \qquad (8)$$

where $RMSR_{jm}$ is the root mean square associated with SSR_{jm}, and

$$DISCRIMINATION_{jm} = \frac{MSV_{jm}}{MSR_{jm}} \qquad (9)$$

where MSV_{jm} and MSR_{jm} are mean squares corresponding to SSV_{jm} and SSR_{jm}. Likewise, the disagreement of individuals from the panel were measured by relating residuals across violins and sessions for individual listeners in a panel to residuals across violins, sessions and listeners of that panel.

However, a careful study of the raw data recommended not to consider disagreement as a selection criterion, because there were enough listeners with a general shift of the perceived scale but with excellent reliability and discrimination skills. Apart from that the panels of four were considered to be too small to trust a selection criterion based on panel agreement. Furthermore, most of the derived results from measurements are relative measures within panels, and there is no need for absolute measures in this study.

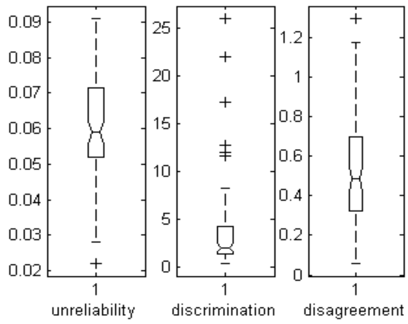

Figure 8: spread of unreliability, discrimination and disagreement for the 24 subjects involved in the main session

Figure 8 shows the spread of the three quality measures. Four out of 96 data sets exceeding 0.175 for a ratio of unreliability to discrimination have been excluded from further analysis.

5.4 Results

5.4.1 Results for the direct scale

Figure 9 illustrates the position span *PS* along the direct scale *SC* used in the main test, individually for the two rooms and averaged across rooms. Clearly, the much denser scale *SC* now delivers a wider span of responses than scales *SA* and *SB* did in the pretests. A span of one indicates that the associated sigma, expressing uncertainty, is about equivalent to one discrete step along the scale. Thus, a summation of the position span density functions would deliver an almost homogeneous function along the scale, or, the discrete scale feels like a continuous scale to the listener with obviously no polarity for decisions. The challenge for listeners is always about the same no matter whether a sound origins from the far end or from nearby positions.

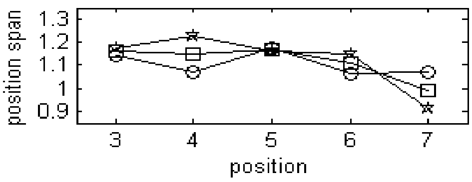

Figure 9: position span PS_L versus position l on the non-linear scale SC used in the main test, circle: room RA, star: room RB, square: average of rooms RA and RB

Note, that this homogeneous span is achieved over a scale that incorporates a potency function that would need only little adjustments for future reuse. The smaller span at the far end of the scale can be increased with the potency factor q in (1). A wider guard would also help, but the general downward trend beginning in the middle of the scale recommends further curve fitting rather than increasing the guard.

Also note, that one and the same scale worked fine in two quite different rooms. This is one of the positive results for reuse in the research community.

A few other observations encourage to use the proposed direct scale. Listeners reported that the challenge was perfectly leveled. One might argue that the scale is too short, or, too sparsely populated, effectively biasing decisions. However, in 17 out of 24 sessions the location difference between listeners of the same panel was larger than one discrete step. This indicates that listeners felt free to take appropriate decisions without being biased. Furthermore, with some of the session attributes the perceived scale was entirely shifted by more than one discrete step, again indicating, that listeners were guided by the sound and the task, rather than by an idea of nicely populating the response across a given scale.

The findings on position span agree very well with other work. Edwards measured just noticeable differences, JND, when moving sources towards or away from subjects [Edw55]. His JND is 1.16 m at 4 m distance and 1.78 m at 8 m distance. These values agree very well with our $PS_1 = 1.1$ and $PS_4 = 1.0$, see Table 1 for the corresponding physical distance.

5.4.2 Sensitivity to session attributes

The aim of introducing the session attributes was to boost the perception of possibly existing differences in terms of the location of violins. Table 5 summarizes the results for mean locations of the violins for each task and all pairs of sessions. On the basis of these panel results, several individual measures can be derived but also larger groups can be formulated for general results.

Sensitivity to noise

No significant difference in using noise levels at 40 dB or 45 dB SPL can be identified. However, using noise or not makes a change. The difference between the two noise levels across both rooms and both violins is $LOC_{45dB} - LOC_{40dB} = 5.33 - 5.73 = -0.40$ ($J = 8$, 1st week). This difference is much smaller than the difference between using noise or not, $LOC_{45dB} - LOC_{0dB} = 5.35 - 4.22 = 1.13$ ($J = 8$, 2nd week).

violin	room	sound	pp long	pp long	pp long	pp * long	pp short	pp ° long	ff long
		noise	pn45	pn40	-	-	-	-	-
VA	*RA*	1st week	5.40	5.43	-	5.03	4.85	5.30	4.03
VB			5.70	5.73	-	5.53	4.58	5.00	4.63
VA	*RB*		5.37	6.03	-	4.27	4.23	4.80	3.05
VB			4.83	5.76	-	4.17	4.25	4.75	3.78
VA	*RA*	2nd week	4.93	-	4.27	5.30	5.05	5.00	3.48
VB			5.67	-	3.93	5.18	5.13	5.25	3.95
VA	*RB*		5.40	-	4.45	5.35	4.83	4.98	4.43
VB			5.41	-	4.23	5.23	4.78	5.20	4.97

Table 5: panel results from the main test, locations LOC_i of violin i sounds averaged across positions $x_{i..lm}$, and normalized to the non-linear scale SC

A more general result is obtained when comparing all sessions under noise at any level ($J = 24$) to sessions without noise ($J = 24$, incl. columns marked * and °), both using long pianissimo sounds: $LOC_{40or45dB} - LOC_{0dB} = 5.47 - 4.86 = 0.61$. Therefore, using noise in the listening area will shift the entire scale further to the back about more than half a discrete step on scale SC. This contradicts the findings in [Mer89], where noise shifted the scale towards the listener.

Did noise boost any perceptible difference? The difference of perceived location is almost zero between violins and quite stable across panels and rooms. In a session with long pianissimo sounds ($J = 16$, marked *) the difference is $LOC_{VA,0dB} - LOC_{VB,0dB} = 4.99 - 5.03 = -0.04$. Another four panels concluded on the same difference ($J = 16$, marked °) $LOC_{VA} - LOC_{VB} = 5.02 - 5.05 = -0.03$. Using noise, the difference grows slightly ($J = 8$) $LOC_{VA,45dB} - LOC_{VB,45dB} = 5.28 - 5.40 = -0.12$. However, this difference is still much smaller than the shift caused by noise.

Sensitivity to intensity

With fortissimo sounds the entire scale shifts towards the listeners when compared with the pianissimo sounds. Considering the two violins across both rooms and both weeks the difference is $LOC_{pp} - LOC_{ff} = 5.01 - 4.04 = 0.97$. Such a shift agrees well with the general finding from other studies, that intensity is one of the main cues for distance perception.

On asking at which intensity level a potential location difference between violins is more likely to be perceived, the location difference across rooms and weeks is computed for the two violins. Playing loudly, we obtain $LOC_{VA,ff} – LOC_{VB,ff}$ = 3.75 – 4.33 = -0.58 (J = 16). This is more than for the pianissimo play (-0.04, see above) and agrees with the pretests. In conclusion, the pianissimo play did not help to bring differences.

The findings are in accordance with other work, as the intensity is found to be supporting discrimination, but it does not provide absolute measures. This becomes clear when considering Figure 5: the ff sound presentations at levels well above 70 dB SPL still facilitated locating at the far end whereas the pp sound presentations well below 65 dB SPL facilitated nearby locating.

Other than in [Nie93] we did not observe wider spans in the responses when intensity levels go down by 15 dB.

Sensitivity to sound duration

A result is that short sounds are located closer to the listener than longer sounds. Across both violins, both rooms and both weeks, the difference is $LOC_{long} – LOC_{short}$ = 5.03 – 4.71 = 0.32 (J = 16). Do the short sounds help to discover potential differences? The difference between violins across rooms and weeks, when played short, is $LOC_{VA,short} – LOC_{VB,short}$ = 4.74 – 4.68 = -0.06. This is practically no improvement when compared with the difference of only 0.03, when the violins are played long in front of the same panel (J = 16, marked *).

Sensitivity to room type

Asking similar questions in relation to the acoustical room, three measures are computed. The difference between the two rooms measured across both violins in both weeks and across all session attributes (but not the 40 dB column and not the unmarked pp long column, J = 24) is $LOC_{RA} – LOC_{RB}$ = 4.95 – 4.70 = 0.25. The difference between violins as measured in room *RA* across both weeks and across session variables (J = 24) is $LOC_{VA,\ RA} – LOC_{VB,\ RA}$ = 4.84 – 5.06 = -0.22. Likewise, the difference for room *RB* is $LOC_{VA,\ RB} – LOC_{VB,\ RB}$ = 4.67 – 4.74 = -0.07. In conclusion, there is only a slight difference in between rooms and none of the rooms is better suited than the other for discovering potential perceptual differences.

Session drift

The drift between weeks measured across both violins, both rooms and all session attributes is $LOC_{1st\_week} – LOC_{2nd\_week}$ = 4.68 – 4.97 = -0.30. This drift is about the same for all attributes.

Sensitivity to violins

Searching for the presumed difference between the violins, there is little evidence. The general difference across all parameters is $LOC_{VA} – LOC_{VB}$ = 4.75 – 4.90 = -0.14. This little value did not encourage translations to a general measure on the physical scale, using (2). The largest difference can be observed when played loudly (-0.58, see above). This difference is even stable across rooms and weeks. However, under these conditions violin *VA* is perceived closer than violin *VB*, which even contradicts pretest assumptions.

parameter / attribute	location difference	95% conf. interval	p-value
noise 45 dB vs. 0 dB	1.13	0.67 1.59	0.0009
intensity 60 dB vs. 75 dB	0.97	0.36 1.58	0.0043
duration long vs. short	0.32	0.03 0.62	0.036
room *RA* vs. *RB*	0.25	-0.14 0.63	0.200
violins *VA* vs. *VB*	-0.14	-0.53 0.24	0.451
1st vs. 2nd week	-0.30	-0.68 0.08	0.117

Table 6: mean difference of perceived locations *LOC* of violin sounds across positions $x_{i..lm}$ for specific session attributes, confidence intervals and p-value against null hypothesis

Overview of location differences

Table 6 summarizes the differences found for the location of violin sounds averaged across positions, panels and sessions, as discussed above. In addition, 95% confidence intervals these differences and p-values against null hypothesis are included from ANOVA analysis. Basically, the most significant location differences come from the noise additive and the intensity levels. The least significant differences are caused by the choices of violin and room type.

5.4.3 General human location abilities

Listeners were well able to locate even the true position of sound sources. Figure 10 illustrates the deviation of the perceived position from the true position. Surprisingly, the deviation is small considering the fact, that listeners memorized the visual cue for a few moments only. In accordance with the literature, there is a tendency to underestimate the far positions. However, the observed difference is much less than found by other studies. Overestimation of close-by positions cannot be observed, because scale *SC* starts at 4 m physical distance and the overestimation is reported for distances below 1 m. However, the tendency of the downward tendency for the deviation from position 4 to position 3 does not encourage to believe there will be such overestimation.

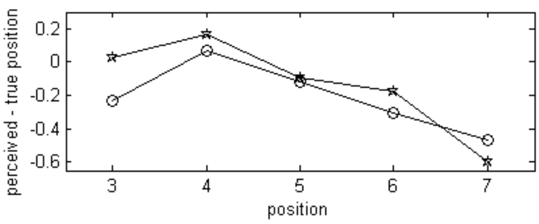

Figure 10: perceived position LOC_l minus true position l versus position l, across all subjects, violins, rooms and sessions, circle: across all sounds, star: across pp long sounds

With the experience of this study, the author trusts that the underestimation can even be further reduced. Denser spacing and a wider guard at the end will decrease the deviation between perceived and physical distance. Reviewing some of the earlier work, the underestimation was clearly caused by a limited scale [Nie93]. The same logic applies for overestimating nearby sources.

Another interesting observation is that the span of responses is much lower in this test when compared with other tests. Taking the gross of the best results in [Nie93], and normalizing the position span to the given scale - as has been done in this study - *PS* is roughly 1.5. This is much more than the 1.1 for scale *SC* or 0.51 for scale *SA*. This is surprising since (i) the scale used by Nielsen (1.0 1.71 2.92 5.0 m) is even wider than scale *SA*, (ii) the scale is much sparser populated, (iii) there is no guard around, and (iv) listeners used a permanent visual cue of a physical scale. One explanation may be that little head motion and violin motion facilitated additional cues in this study.

6 SUMMARY AND CONCLUSIONS

We played 480 sounds samples on violins in semi-reverberant rooms and analyzed 1920 responses from 24 individuals to understand perceptual differences in depth location. For this purpose, we developed a direct scale which proved to be well balanced between desirable resolution and challenge for subjects, but also well balanced over the depth of the scale. This scale worked fine in different acoustical environments and the community is encouraged to employ such scales.

Concerning general human abilities, subjects are able to accurately determine the true physical position without a visual cue. This is true for untrained listeners in rooms they are not even familiar with. This observation holds for different semi-reverberant rooms and is independent from the choice of the violin as well as from the choice of using a voice reference or not. In this test, a smaller span of responses has been observed compared with other tests.

Concerning the test design, the approximately quadratic scale proofs to be well balanced, since the statistics delivers likewise spans for all discrete positions. A grid with eight discrete positions seems to be a good choice between limiting the challenge for subjects and enhancing the resolution for direct scaling.

Pink noise in the listening area slightly increases the perceived distance for both types of violins, contradicting the findings from other studies. Short presentations of sound (~1 s) shifted the scale slightly towards the listeners when compared to long presentations (~20 s).

Concerning the violins, differences of perceived distance are only very small and are only observable with fortissimo passages. None of the other session attributes - duration of sound, noise additive, or voice anchor - helped to discover perceptual differences in distance.

With respect to the initial question, the perceived physical distance seems not to express what people perceive as an acoustical intimate sound. The small differences in terms of perceived distance do not explain the general difference between violins perceived in terms of acoustical intimacy. Obviously, the subjects in the test focused on solving the "room-position" task such that relevant binaural cues dominated response decisions rather than the sensation of tone or the feel of an emotional touch.

The often quoted underestimation of distant sources did not really appear here, as the scale was well balanced and had an additional guard at its ends.

ACKNOWLEDGEMENTS

The author thanks M. Schleske for valuable violins and the German Fed. Ministry of Education and Research, BMBF, for funding the violin project, no. AiF 1767X07.

REFERENCES

[Bec06] S. Bech, N. Zacharov: Perceptual Audio Evaluation – Theory, Method and Application. John Wiley & Sons, 2006.

[Bek38] Békésy, G., Über die Entstehung der Entfernungsempfindung beim Hören. Akustische Zeitschrift 3, 21-31, 1938.

[Ber04] L. Beranek: Concert Halls and Opera Houses - Music, Acoustic and Architecture. Springer New York, 2nd ed. 2004.

[Bla83] Blauert, J., Spatial Hearing. MIT Press Cambridge, MA, 1983.

[Bro99] Bronkhorst, A.W., Houtgast, T., Auditors distance perception in rooms. Nature, Vol.397, 517-520, Feb. 1999.

[Edw55] Edwards, A.S., Accuracy of auditory depth perception, J. of General Psychology 52, 327-329, 1955.

[Hol69] Holt, R.E., Thurlow, W.R., Subject orientation, and judgement of distance of a sound source. J. Acoust. Soc. Amer. 46, 1584-1585, 1969.

[Loo98] Loomis, J.M., et. al., Assessing auditory distance perception using perceptually directed action, Percept. Psychophys. 60, 966-980, 1998.

[Mey64] Meyer, J., Die Richtcharakteristiken von Geigen. Instr.-Bau.-Z. 18, 275, 1964.

[Mer75] Mershon, D.H., King, E., Intensity and reverberation as factors in the auditory perception of egocentric distance. Percept. Psychophasics 18, 409-415, 1975.

[Mer89] Mershon, D.H., et. al., Effects of room reflectance and background noise on perceived auditory distance, Perception 1, 403-416, 1989.

[Nie93] Nielsen, S.H., Auditory distance perception in different rooms. J. Audio Eng. Soc. Vol.41, No. 10, 755-770, 1993.

[Sch94] Schlich, P. GRAPES: A method and SAS program for graphical representation of assessor performance. Journal of Sensory Science 9 (1994), 157-169.

[Sim73] Simpson, W.E., Stanton, L.D., Head movement does not facilitate perception of distance of a source of sound, Am. J. Psychol. 86, 151-159, 1973.

[Tho92] Thompson, S.P., On the function of the two ears in the perception of space. Philosophical Magazine (5th series) 13, 320-334, 1892.

[Zah05] P. Zahorik, D. S. Brungart, A. W. Bronkhorst: Auditory Distance Perception in Humans: A Summary of Past and Present Research, Acta Acustica united with Acustica, Vol. 91, 409-420, 2005.