ARCHIVES OF ACOUSTICS
6, 1, 13-22 (1981)

PROBLEMS IN RECORDINGS FOR LISTENING EVALUATION OF THE QUALITY OF VIOLINS

ANDRZEJ SASIN, ANTONI KARUŻAS, TOMASZ ŁĘTOWSKI

Sound Recording Department, Chopin Academy of Music (00-368 Warszawa, ul. Okólnik 2)

This paper considers the possibility of performing a comparative evaluation of violin sound quality, based on sound recordings. The investigations dealt with the effect of the microphone techniques – mono, stereo *AB* and stereo *XY* – and the acoustics of the room, on the usefulness of the recordings made in given conditions for the evaluation of the sound quality of violins. The paper presents the investigation results and a comparison of these with the results of the "live" sound quality assessment performed by expert violinists.

1. Introduction

The ultimate test of the quality of any musical instrument is an evaluation made on the basis of auditory sensations. Investigations aimed at such an auditory evaluation can have the following forms;

(a) evaluation by an instrument player who can directly get to know an instrument under test, i.e. direct evaluation,

(b) evaluation of the sound quality of an instrument made on the basis of sound recordings, i.e. indirect evaluation.

In the previous investigations of the sound quality of musical instruments the so called listening tests containing music recorded for the purpose of underlining the characteristics of instruments investigated, have been widely used. It seems that for the evaluation based on such tests to be considered valid, it is necessary to pay more attention to the methods of sound recording being used. The aim of the present paper is to answer the question whether such a recording is possible that could be an index of quality of a given instrument, and possibly whether a set of principles could be established, which could be used in making recordings for the listening tests of the sound quality of instru-

ments. The investigation was carried out using the violin as the specimen instrument. The violin was selected because

(a) there is a great variety of relatively easily available instruments,

(b) there is a large interest in the investigations of the quality of the violin,

(c) the instrument is portable.

The investigations were carried out on three instruments. The sound qualities of all the instruments were previously evaluated by the direct method by three outstanding soloists employed at the Stringed Instruments Department of the Chopin Academy of Music in Warsaw. The results of this evaluation can be presented in the form of the following rank series:

1. Violin no. 1, unknown violin-maker, an instrument of very good quality;
2. Violin no. 2, made by NEUNERT, an instrument of good quality;
3. Violin no. 3, a mass made specimen, of very bad quality.

The experts at the same time pointed to the considerable difference between violins nos. 1 and 2 and violin no. 3.

The results of the experts' evaluation were unknown both to the workers carrying out the experiment and the listeners until the final results of the investigations were obtained.

2. Procedure

The first stage of the investigation consisted in a listening evaluation of selected instruments on the basis of sound recordings made in different ways. Established as a result of various discussions and tests, the sound material contained:

1. diatonic gamuts of one octave performed on each string;
2. three chords; *G* major, *D* major, *C* major, one following the other,
3. a part of Ysaye's Sonata;
4. a part of J. S. Bach's Chaconna.

All the musical pieces were recorded on the *mf* – *f* dynamic level.

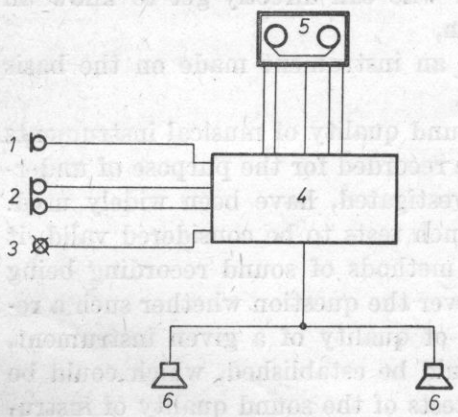

Fig. 1. A block diagram of the electroacoustic system used in the recordings

1 – a mono *U* – 67 Neumann microphone, *2* – two *U* – 67 microphones in the system *AB*, *3* – a stereo *SM* – 69 Neumann microphone in the system *XY*, *4* – a Siemens control panel, *5* – an eight-track Studer tape recorder, *6* – *ZG* – 60 *C* loudspeaker columns

The material mentioned above was recorded by three microphone techniques, i.e. mono, XY, and AB. A comparison of sound recordings obtained by these techniques was expected to provide an answer to the question whether the technique used does not matter in terms of sound quality evaluation, or otherwise whether the differences due to the different principles of microphone techniques are so essential with regard to this type of investigation that one of them should be preferred.

A block diagram of the system used for the implementation and monitoring of the recordings is shown in Fig. 1.

By using the multitrack technique a single performance of a musical piece performed on one instrument could be recorded at the same time by three independent recording systems. Thus employing one person only who reproduced the musical piece successively on three instruments, the minimum influence of external factors of the recording was assured.

All the recordings were made at studio $S-1$ and the concert hall of the Chopin Academy of Music in Warsaw. These rooms are distinctly different in terms of volume and reverberation time. Their parameters are the following:

– at studio $S-1$ the volume is 1080 m^3 and the reverberation time is 1.1s,

– at the concert hall the volume is 6000 m^3 and the reverberation time is 1.8s.

In both cases the reverberation time is given as a mean for the frequency range from 500 to 2000 Hz.

Thus the investigations also included the problem of the influence of the acoustics of the room both on the recording and on the sound quality of the instrument.

The final recordings were preceded by a selection of the position of the player in a given interior and of the optimum arrangement of microphones in each of the system used. In each case such a position of microphones was assumed as optimum that to a highest degree assured the condition of the similarity of a recording to the original sound of the violin recorded. This choice was made on the basis of a number of preliminary experiments.

The music material was recorded in one session for each of the rooms. In order to evaluate the quality of a given violin on the basis of the material recorded, listening tests were carried out. In all the tests the method of pairs comparison was employed [3].

The durations of the individual stimuli were the following:

1. diatonical gamuts – 7s (on one string);
2. chords – 4s;
3. a part of Isaye's sonata – 8s.

The test set consisted of 36 trials of the $A-B$ type arranged randomly and balanced in terms of succession. One-second intervals between stimuli in a trial and two and half second intervals between trials were used. The intervals between trials were intended to give the expert enough time to put his

Table 1

No.	mono						*XY*						*AB*					
	violin						violin						violin					
	I		II		III		I		II		III		I		II		III	
	S−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*
1	19	20	14	10	3	6	12	18	16	12	8	6	16	16	14	13	6	7
2	19	18	10	12	7	6	15	11	8	13	13	12	7	14	10	13	19	9
3	19	19	11	13	6	4	19	17	12	11	5	8	17	18	13	14	6	4
4	19	20	12	11	5	5	19	17	12	13	5	6	16	15	13	13	7	8
5	16	23	14	11	6	2	17	15	8	13	11	8	13	13	12	13	11	5
6	17	23	13	11	6	2	14	18	12	13	10	5	14	17	12	16	10	3
7	19	20	10	12	7	4	13	18	15	15	8	5	14	19	13	13	9	4
8	16	20	13	12	7	4	12	17	10	10	14	9	15	18	10	13	1	5
9	17	19	14	11	5	6	10	16	15	13	11	7	17	15	15	16	4	5
10	18	20	12	11	6	5	16	21	13	8	7	7	15	20	14	12	7	4
11	16	20	13	11	7	5	13	20	12	10	11	6	14	18	10	12	12	6
12	16	18	12	13	8	5	12	18	10	11	14	7	15	13	10	20	11	3
13	15	18	11	12	10	6	13	12	12	14	11	10	14	15	12	12	10	9
14	15	19	13	13	8	4	13	14	12	14	11	8	13	12	16	15	7	9
15	16	18	12	13	8	5	15	15	11	9	10	12	14	12	10	15	12	9
16	11	17	11	15	14	4	16	10	15	14	5	12	14	15	14	16	8	5
17	11	18	17	10	8	8	15	15	14	14	7	7	12	13	15	13	9	10
18	11	16	12	12	13	8	12	14	13	13	11	9	16	18	9	14	11	4
19	14	22	12	8	10	6	11	16	16	13	9	7	13	18	16	14	7	4
20	16	20	15	8	5	8	10	11	16	16	10	9	13	16	13	14	10	6
	320	388	251	229	149	103	277	313	252	249	191	160	282	315	251	281	177	119

S−1 – studio *S*−1, *SK* – concert hall of Chopin Academy of Music, I, II, III – violins.

mark in the estimation sheet. The whole test consisted of three sets of trials, each containing the same sound material recorded in one of the three selected microphone systems. Each set was intended to give an evaluation of a violin on the basis of recordings made in a given microphone system. The total duration of one set did not exceed 18′. The tests were performed on the basis of a 20-member expert group consisting of students from higher grades in the departments of sound engineering, of instruments (violins section) and of theory and composition (sections of theory and conducting). The listening test were carried out at the concert hall and at studio $S-1$ of the Chopin Academy of Music, both cases involving the whole expert group. The loudness level was 80 phons. 20′ rest intervals were introduced between the test sets. In order to give a clear distinction between instruments investigated, a preferential evaluation: worse-better, was used in the test [3]. The results of the experiment are shown in Table 1. It presents numbers of points given to a particular instrument by each expert, according to the instruction preceding the test. The evaluation marking was the following: choosing an instrument in a given pair as better — 1 point, lack of choice — 0 points. The table gives the sum totals of points assigned to the individual instruments in the comparison made by experts.

The points given to the particular violins and shown in Table 1, were the basis for a statistical analysis of the results, which used

(a) the Kolomogorov test — the investigation of the normalcy of the distribution of the results obtained [1].

(b) the t-Student test — the investigation of the statistical significance of differences [2].

The results of the Kolomogorov test permitted a hypothesis that the distribution is normal for the estimates of the whole listener group to be taken. The above investigation was performed in order to detect the possible cases of multimodal distributions or a rectangular distribution, which would suggest that it is impossible to obtain an "objectivized" qualitative judgment of the sound quality of violins under investigation.

The results from the t-Student test obtained for the individual pairs of instruments investigated with the three microphone systems and for the two rooms in which the recordings were made, are given in Table 2.

According to the rule frequently used in mathematical statistics, the following significance levels were taken for the hypotheses assumed [3]:

1. a highly significant result — the probability of the validity of the null hypothesis, $p \leqslant 0.1\,\%$;

2. a significant result — the probability of the validity of the null hypothesis within the limits from 0.1 to 1.0 %;

3. a probably significant result — the probability of the validity of the null hypothesis within the limits from 1.0-5.0 %;

4. an insignificant result — the probability of the validity of the null hypothesis $p \geqslant 5\,\%$.

The null hypothesis was the statement that the investigation method used does not permit to distinguish qualitatively between instruments under investigation.

Table 2

Violin		Technique		
		mono	XY	AB
I–II	concert hall	11.362	2.232	2.321
	studio $S-1$	4.232	1.277	2.867
I–III	concert hall	20.225	7.248	10.332
	studio $S-1$	7.423	4.011	4.164
II–III	concert hall	9.703	6.429	9.460
	studio $S-1$	6.158	2.939	2.783

The values of the t-Student test, which evidence a given significance level of differences, as read from the statistical tables for the t-Student distribution with $S = N-1$ degrees of freedom, where N is the number of listeners ($N = 20$), are the following [4]:

(1) $t \geqslant 3.883$ – highly significant;
(2) $t \geqslant 2.861$ – significant;
(3) $t \geqslant 1.729$ – probably significant;
(4) $t \geqslant 1.729$ – insignificant.

The values of the t-Student test obtained from the experiment are shown in Table 2. A comparison of the empirical values with the critical t-Student values taken from the tables shows a real difference between instruments investigated. Only once, i.e. the difference in evaluation between violin I and violin II, based on the recording by the XY technique at studio $S-1$, it appeared to be insignificant. A comparison of the empirical values of the t-Student test obtained for the recordings made at studio $S-1$ and at the concert hall, using the same microphone system, shows a significant difference between the results obtained. This suggests an effect of the acoustics of the room on the sound quality in the recording.

Since the values of the t-Student test obtained as a result of the above experiment suggest a distinction between instruments in each of the three systems used, the investigators determined to carry out another test in order to answer the question which of the sound recordings made in the three microphone systems was closest in terms of sound quality to the "live" performance (the assessment of the fidelity of a recording). A piece of J. S. Bach's Chaconna which was recorded previously was used as sound material. The test was based on the $A-W-B$ method with a presentation of the standard W ("live") performance between test stimuli. Each sound sample lasted 9 seconds. In view of the "live" W presentation it was difficult to take a precise duration of the intervals between the stimuli $A-W-B$ in the test. The real intervals between

the stimuli lasted about 2 seconds, while the interval between the test trials lasted 3 seconds. The whole test was divided into three trial sets, each consisting of six randomly arranged trials (three trials of the $A-W-B$ type and three trials of the $B-W-A$ type). The total duration of one set did not exceed 6′. The aim of each task set was to select such a microphone system that would give in its recording the most faithful representation of the quality of the individual instruments investigated. The test was carried out at the concert hall of the Chopin Academy of Music for a twelve-member expert group. The performer was situated half-way between loudspeakers.

The principle of point giving was the same as in the first test: choosing a given microphone system — 1 point, lack of choice — 0 points. Table 3 shows the sum total of points assigned to a given microphone system by each expert in each trial set. The principle of point summation was the same as in Table 1. The data in this table were then the basis for statistical processing of the results in the same way as the statistical elaboration of the results of the previous experiment.

Table 3

No.	Violin I			Violin II			Violin III		
	mono	*XY*	*AB*	mono	*XY*	*AB*	mono	*XY*	*AB*
1	4	0	2	3	1	2	4	1	1
2	4	0	2	3	0	3	2	3	1
3	3	3	0	3	1	2	1	1	4
4	3	1	2	3	1	2	1	2	3
5	3	2	1	2	3	1	4	1	1
6	3	1	2	2	1	3	3	2	1
7	3	0	3	2	3	1	1	2	3
8	2	3	1	2	2	2	2	1	3
9	2	2	2	2	2	2	3	2	1
10	2	2	2	2	1	3	2	1	3
11	2	0	4	2	2	2	0	2	4
12	1	2	3	1	2	3	0	2	4
	32	16	24	27	19	26	23	20	29

Under the assumption of significance levels according to the same principle as in the first test, there are now four critical values of the t-Student distribution read from the statistical tables for $S = N-1$ degrees of freedom (where $N = 24$, i.e. 12 listeners times two successions of presentation);

(1) $t \geqslant 3.496$ — highly significant;
(2) $t \geqslant 2.500$ — significant;
(3) $t \geqslant 1.714$ — probably significant;
(4) $t \geqslant 1.714$ — insignificant.

The empirical values of the t-Student test obtained for the individual pairs of microphone systems in the case of each instrument are given in Table 4.

Table 4

Technique	Violin		
	I	II	III
mono – XY	1.796	1.623	0.495
mono – AB	1.430	0.266	0.640
AB – XY	1.100	1.277	1.650

The results obtained show that the fidelity of a recording (similarity ot the original) was, according to the listeners, highest:

(a) in the case of violin I when the mono system was used;

(b) in the case of violin II, similar when the mono and the stereo AB systems were used;

(c) in the case of violin III when the stereo AB system was used.

In the case of all the three violins the representations in the stereo XY technique were found to be least faithful.

3. Discussion and conclusions

The results obtained in both experiments and a statistical analysis of them lead to the following conclusions and remarks.

The results of the evaluation of the sound quality of selected violions, made using the listening test, coincided with the results of direct evaluation made by experts (the members of the Department of Bow Instruments). The recordings made in each of the microphone systems used permitted a good estimation of differences between the instruments investigated. The credibility of the estimates made by the expert group, that is visible from a statistical processing of the results, was very high. A comparison of the values of the t-Student test obtained for the qualitative evaluation of recordings made in the particular microphone systems (Table 2) suggests the conclusion that difference between instruments was most conspicuous in the case of a recording made using a mono microphone, and successively in a recording made using the stereo techniques AB and XY. A comparison of the values of the t-Student test obtained for recordings made at studio $S-1$ and at the concert hall seems to indicate that each microphone system responded to a higher degree to the differences between instruments at the concert hall, i.e. a room with better acoustic conditions. An exception here is the t-Student test value for the comparison of violin I and violin II, based on recordings in the AB technique.

Since the aim of the investigations was not to solve the problem of the effect of the acoustics of a room, but only to show whether this effect occurs at all in terms of evaluation of an instrument recorded, it is difficult to draw speci-

fic conclusions on the basis of the results obtained. It seems useful, therefore, to perform investigations of the problem in nearest future, which may shed more light on it. This demand also applies to other variable conditions of the listening technique and to microphone types used in recording.

It follows from the results obtained from the second test that it was most troublesome for the listener group to select that microphone system which would most faithfully represent an instrument. The present results permit, however, the following conclusions to be drawn:

1. in the case of good instruments the mono technique assures highest fidelity;

2. in the case of poor instruments the stereo AB technique assures highest fidelity;

3. the stereo XY technique appears to be least useful in the case of auditory evaluation of the sound quality of the violin on the basis of microphone tests.

The present results also permit a conclusion that the better the violin is, the greater differences occur between recordings obtained by means of different microphone techniques. This leads to a further hypothesis that the poorer an instrument is, the less effect the recording technique has on the representation of the sound quality of an object investigated. This hypothesis requires, however, further and wider experimental evidence.

In summary it can be stated that it is useful to perform subjective investigation of the sound quality of the violin using the comparison method on the basis of sound recordings. Coincidence of estimates made by expert violinists in the so called direct evaluation with estimates made by a listener group by way of indirect evaluation indicates that a recording can sufficiently well evidence the quality of a given instrument. Since the most frequent aim of auditory quality evaluation is selecting the best of a group of objects tested, the present results are in favor of the recordings made in the mono system.

References

[1] A. GÓRALSKI, *Methods of description and statistical conclusion in psychology* (in Polish), PWN, Warszawa 1974.

[2] J. P. GUILDFORD, *Psychometric methods*, McGraw Hill Book Co., New York 1954.

[3] T. ŁĘTOWSKI, *Auditory assessment of electroacoustic devices* (in Polish), Reports of Research and Development Centre of Polish Radio and Television (1976).

[4] *Statistical tables* (in Polish) W. SADOWSKI, ed. PWN, Warszawa 1957.

Received on April 23, 1980.