

VSA Competitions—An Analysis

JOSEPH REGH

36 Sherwood Heights, Wappingers Falls, NY 12590
reghj@aol.com

Tuesday, November 4, 2008, 1:30 pm

Eric Chapman, Moderator: Joe Regh is obviously no stranger to anybody in the VSA, except maybe for those who are newcomers. Joe is a physicist by training who had a distinguished career at IBM. He's been on the VSA Board of Directors since 1983 and a vice president for the last 10 years. Apparently, he waited for me to retire as president before he became a member of the Board. Joe has made enormous contributions to this organization, particularly to the development of the statistics of the biannual competitions and its organization. And, believe me, it's quite a chore to do this kind of thing. Today Joe will talk about how to interpret all of the data for the last several competitions. It's my pleasure to introduce Joe Regh.

Joseph Regh: Let me begin by asking a question, which I assume is on all of your minds: Why aren't more gold medals awarded?

After each competition is over I get telephone calls from instrument makers who think they deserved a gold medal but didn't get one; even from makers whose instruments received very high scores but didn't receive one. They ask, "What is going on with this process?" Most of the time there's a relatively easy answer because in running a competition, we take a large amount of data. And we store the data so that we can go back to them year after year if we need to explain something. Sometimes it is not so obvious what happened in the process that prevents people from getting properly recognized—at least in their opinion.

About two years ago I had a discussion with Joan Miller, who used to be the treasurer of this Society, and who is a very competent computer programmer. She wrote into our competition software—used to collect and manipulate all the data the judges produce—a routine that allows us to do an almost instantaneous statistical analysis. So the first thing I want to do is acknowledge Joan's contribution to the development of our software.

It is something I had wanted to do for many years, but I didn't have the programming skills and the time to implement. What you see here today is an addendum to our software that was developed and exercised for the first time earlier this year. It allows you to access all of the databases, and at the click of a button, perform a statistical analysis with a display as to how these data came about.

Before I show you those charts, I have to set the stage for the most important aspect of the competition, and that is that the data are generated on each instrument by three judges who work independently. That is very important because it allows the committee to come to certain conclusions at the end of the competition when the winners have been determined. It is then possible to evaluate each judge's ability to predict the winner of the competition in the first round without consultation with any of his or her colleagues. That is a powerful way of measuring the ability of a judge to make the right predictions.

Let me go through the setup and the organization of the competition so you understand how it is conducted and what the important points are. Our competition is conducted in three rounds. In Round 1 the concentration is not as much on identifying the best instruments as it is on eliminating those instruments that have no possibility of being included in Round 2. It's almost like pre-screening all the instruments. That is not done haphazardly. Each judge in the workmanship category, for example, considers the merits of each instrument in five to six categories and ranks it numerically in a range of zero to 10 points. And that is tabulated. The tone judges do similarly.

Each of the tone judges plays every instrument and makes a decision as to whether the instrument should be going into Round 2, which would give it a point score of a 3, or a 2, which means there is some uncertainty, or a 1, which means it should not go on. The cumulative scores of the three judges

are then entered into the computer. The computer calculates a ranking, just as it does with scores from the workmanship judges, and the ranking is given back to the judges, who determine where to make the cut level for advancement into Round 2. There's obviously a gray zone, but there are usually very good markers that will allow the judges to make that cut level. It is possible that one of the judges will be impressed favorably by a particular instrument that was scored below their cut level. If all three judges agree to advance that instrument, then the competition rule is that all of the instruments that were scored above that new cut level will be advanced into Round 2.

In Round 2 the instruments are evaluated much more closely. All the data that were generated in Round 1 are wiped out and we start a whole new process. Each judge evaluates each of those instruments for workmanship again, independently, only now they have a smaller number of instruments that are of a generally better quality. The tone judges will have a separate room where the Round 2 instruments are transported to them, and each of them will play each instrument and also listen to the other two judges play them. Again, they do not talk to each other. They then form an opinion based on their data and again assign a 1, 2, or 3 to each instrument. If the score is a 1, it doesn't go to Round 3. If the score is a 2, maybe. With a score of 3, it definitely goes into the medal round.

In Round 3 the very few instruments that have survived are no longer tested in a way where we keep data. We put all the instruments in one room and we lock them up with the judges, who then can talk about them and play them for each other. When done, they may give us a recommendation for a gold medal, which has to be unanimous. If it's not unanimous, it will lead to a certificate of merit. Every instrument that goes into Round 3 will receive a certificate of merit.

This is the point I made in the very beginning, the uniqueness being that the judges work independently in Round 1. By analyzing the data after the end of the competition when we know which of the instruments are the award winners, we are able to assess how well each judge predicted the ultimate outcome. That is sort of a measure of the expertise that the judge brings to the judging process.

In Round 2 we do the same thing again, but with the tone judges evaluating a smaller number of instruments in a separate room. In Round 3 they work as a team, discussing the pros and cons of the highest-scoring instruments, and then we have the

awards designation.

Since we evaluate tone and workmanship separately in parallel, a gold medal instrument, by our definition, is one that excels in both. If one set of judges, either for tone or workmanship, gives a gold medal recommendation, and if the other set of judges gives it at least a certificate of merit in the medal round, then that instrument will get a gold medal. If the instrument does not meet the Round 3 requirements of the other set of judges, the instrument will get a silver medal. The silver medal is awarded specifically for tone if a gold medal recommendation is made by the tone judges, and likewise for workmanship if the gold medal recommendation is made by the workmanship judges.

Examination of the data produced in Round 1 is very revealing simply because of the independence in which the scores are generated by all these judges. That leads to the big question: Do the judges really know what they're doing? We have had some competitors who have had doubts about that. They didn't win and so the implication was that it was because of the limited capabilities of the judges that we hired. So the first question is: Do the judges reasonably well predict the ultimate winners in Round 1?

The listing in Table 1 gives you an idea about the number of data points (greater than 10,000) that we have analyzed for the hundreds of instruments in the three VSA competitions in 2002, 2004, and 2006. Of particular interest are the numbers of instruments that were advanced from Round 1 to Round 2.

Taking the 2006 competition as an example, for workmanship, 21 of the 184 violins made it into Round 2. However, for these same 184 violins, the tone judges advanced 40 into Round 2. That is very understandable because it is significantly more difficult to come to a tonal consensus on an instrument than for quality of craftsmanship. As an extra precaution to make sure that we do identify the best instruments, we increase the Round 2 population whenever there is some uncertainty. To advance 40 of the 184 violins, we included all those violins that were identified by some of the judges as being worthy of being included in Round 2.

Table 2 lists the violin workmanship scores in descending order. The first 52 violins of the total of 184 entered in the competition are listed on this sheet. There were three more sheets just like this one to complete the tabulation. It was difficult to do that graphically without having a scanning tool. The medals or certificates that were ultimately

Table 1. Numbers of instruments evaluated in Rounds 1 and 2 for VSA Competitions in 2002, 2004, and 2006.

	Workmanship or Tone	2006	2006	2004	2004	2002	2002
Round		1	2	1	2	1	2
Violin	Workmanship	184	21	182	22	164	40
	Tone	184	40	182	56	164	34
Viola	Workmanship	96	34	85	23	64	17
	Tone	96	54	85	21	64	17
Cello	Workmanship	48	20	48	20	31	13
	Tone	48	35	48	17	31	16
Data Points		2952	1227	2835	672	2331	621

awarded were added to this listing at the end of the competition. Every judge independently generated his/her list of scores, so there was no fudging possible. Out of nearly 200 instruments, all were identified in the first 28 spaces. Obviously, these judges really knew what they were doing.

Table 3 is a similar list of scores given by the violin tone judges for the same 184 violins. You can see that there's a somewhat larger distribution. In fact, some of these violins, after the fact, were scored below the designated level. Because of the larger spread, it was necessary to increase the population in Round 2 to make sure that no matter how far down the list an instrument got, it was included in that final population that would be examined in Round 2.

The results of the violin tone judging in Round 2 are listed in Table 4. Forty instruments were included in Round 2. The scores were generated by each of the three judges independently, but with more care and time, because every instrument was played in an isolated environment. Each one was played and listened to three times. These judgments are more accurate and more functional. You will see that the conclusions, although independently generated, were remarkably identical. We can conclude that the tone judges were equally competent. These people knew what they were doing.

The results of the viola workmanship judging in Round 1 are shown in the listing of Table 5. There is a difference in the distribution here that accounts for different interpretations of whatever the judges were looking at. There are personal tastes involved. This is the first of four pages identical to this. You have to look for this long tail underneath with no data entered and no predictions made.

The results of the viola tone judging in Round 1 are shown in the listing of Table 6. The scores are more indicative of different tastes, and when we got to Round 2 (Table 7), the distribution tightened up significantly and the best instruments were identified.

The results of the cello workmanship judging in Round 1 are listed in Table 8 and for cello tone in Round 1 in Table 9.

If the judges had selected 20 instruments for Round 2 (Table 10), but one of them considered instrument number 30 also to be really nice and interesting, then the judges would be required to go to 30, they have to include every instrument above. Then the population would be increased, and so these instruments that would be above that cut level that did not qualify would also be dropped out here. This proves beyond any argument that the judges that we hire, while they do have different tastes, do correctly identify the best instruments. You can see the personal tastes of the judges in going about their job. Understandably, those are more pronounced for the tone judges than it is for workmanship judges.

Now I will show you a series of graphs that are the outcome of the statistical analysis on some of the data. Some of it is more interesting than others, but with one set for violins I want to set the stage and give you a comprehensive picture. Later on I'll show you the major charts so you can see the consistency from one instrument category to another, and from one competition year to another.

This chart shows a typical workmanship data distribution (Fig. 1). It has 184 data points and lists all the various aspects that the judges look at in evaluating the instruments. One of the instructions we give the judges is that, for this population

Table 2. 2006 VSA Intl. Competition. Violins, Workmanship, Round 1. Each judge's 52 highest scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	88	79.2	SW	171	86.8	G	88	79.2	SW
2	17	75.2	SW	88	79.2	SW	60	76.8	2C
3	60	72.4	2C	192	71.6	CW	197	74.4	SW
4	77	70.0	.	45	71.2	CW	192	67.6	CW
5	55	69.2	.	164	70.4	.	171	66.8	G
6	79	69.2	.	136	66.0	.	187	66.0	CW
7	121	69.2	.	90	65.2	.	68	64.0	.
8	68	68.4	.	43	63.6	CW	160	62.0	.
9	43	67.6	CW	17	63.6	SW	17	60.4	SW
10	197	67.6	SW	100	63.6	.	87	58.8	.
11	178	66.8	.	93	62.8	.	43	58.4	CW
12	100	66.8	.	87	62.0	.	77	58.4	.
13	44	66.8	.	24	56.0	.	84	58.4	.
14	187	66.0	CW	57	55.2	.	44	57.6	.
15	84	66.0	.	91	55.2	.	178	57.6	.
16	171	64.0	G	8	54.4	.	79	56.8	.
17	152	61.6	.	102	54.0	.	109	56.0	.
18	164	61.6	.	58	54.0	.	5	56.0	.
19	61	60.0	.	139	53.6	.	6	54.0	.
20	57	60.0	.	84	53.6	.	183	53.6	.
21	76	60.0	.	178	53.6	.	164	52.8	.
22	87	59.2	.	13	53.6	.	132	52.8	.
23	45	59.2	CW	187	53.6	CW	136	52.8	.
24	109	58.4	.	44	53.2	.	57	52.4	.
25	136	58.4	.	50	52.8	.	45	52.0	CW
26	192	58.4	CW	125	52.4	.	100	52.0	.
27	160	58.4	.	142	52.4	.	90	51.6	.
28	183	58.4	.	197	51.6	SW	121	51.6	.
29	93	58.4	.	77	51.2	.	34	50.8	.
30	179	57.6	.	132	50.4	.	74	50.0	.
31	139	57.6	.	94	50.4	.	179	49.2	.
32	4	57.6	.	73	50.0	.	200	49.2	.
33	27	57.6	.	114	49.2	.	182	49.2	.
34	94	56.8	.	79	48.4	.	8	49.2	.
35	129	56.8	.	133	48.4	.	135	48.4	.
36	130	56.4	.	200	47.6	.	24	48.4	.
37	185	56.0	.	5	47.2	.	125	48.4	.
38	42	56.0	.	150	46.8	.	185	48.4	.
39	24	56.0	.	66	46.4	.	177	47.6	.
40	142	56.0	.	49	46.4	.	180	46.0	.
41	90	55.6	.	129	46.0	.	142	45.6	.
42	150	55.6	.	169	46.0	.	130	45.6	.
43	54	55.6	.	68	46.0	.	129	45.2	.
44	39	55.2	.	195	46.0	.	26	45.2	.
45	34	54.0	.	119	46.0	.	102	44.4	.
46	91	53.2	.	184	46.0	.	61	44.4	.
47	92	52.8	.	183	45.6	.	54	43.6	.
48	71	52.8	.	151	45.2	.	93	43.6	.
49	108	52.4	.	168	45.2	.	58	43.6	.
50	73	52.0	.	109	45.2	.	169	43.6	.
51	74	52.0	.	185	45.2	.	150	43.6	.
52	18	51.6	.	152	45.2	.	37	43.2	.

Table 3. 2006 VSA Intl. Competition. Violins, Tone, Round 1. Each judge's 52 highest scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	6	3.0	.	3	3.0	.	3	3.0	.
2	20	3.0	.	34	3.0	CT	6	3.0	.
3	26	3.0	.	37	3.0	ST	9	3.0	.
4	34	3.0	CT	42	3.0	.	12	3.0	.
5	37	3.0	ST	52	3.0	.	16	3.0	.
6	42	3.0	.	60	3.0	2C	18	3.0	.
7	44	3.0	.	61	3.0	.	21	3.0	.
8	84	3.0	.	67	3.0	.	23	3.0	.
9	95	3.0	.	76	3.0	.	31	3.0	.
10	110	3.0	.	86	3.0	.	34	3.0	CT
11	118	3.0	.	100	3.0	.	37	3.0	ST
12	119	3.0	.	110	3.0	.	46	3.0	.
13	121	3.0	CT	130	3.0	.	49	3.0	.
14	130	3.0	.	146	3.0	.	55	3.0	.
15	135	3.0	.	161	3.0	.	60	3.0	2C
16	136	3.0	ST	171	3.0	G	67	3.0	.
17	138	3.0	.	186	3.0	.	73	3.0	.
18	146	3.0	.	195	3.0	CT	76	3.0	.
19	157	3.0	CT	2	2.0	.	78	3.0	.
20	158	3.0	CT	4	2.0	.	83	3.0	.
21	169	3.0	.	6	2.0	.	118	3.0	.
22	171	3.0	G	7	2.0	.	125	3.0	.
23	14	2.0	.	16	2.0	.	130	3.0	.
24	27	2.0	.	21	2.0	.	134	3.0	.
25	36	2.0	.	22	2.0	.	135	3.0	.
26	43	2.0	.	25	2.0	.	138	3.0	.
27	45	2.0	.	26	2.0	.	141	3.0	.
28	53	2.0	.	27	2.0	.	145	3.0	.
29	60	2.0	2C	33	2.0	.	146	3.0	.
30	88	2.0	.	36	2.0	.	160	3.0	.
31	89	2.0	.	43	2.0	.	161	3.0	.
32	92	2.0	.	45	2.0	.	162	3.0	.
33	94	2.0	.	46	2.0	.	171	3.0	G
34	96	2.0	.	51	2.0	.	4	2.0	.
35	120	2.0	CT	54	2.0	.	17	2.0	.
36	125	2.0	.	55	2.0	.	22	2.0	.
37	139	2.0	.	58	2.0	.	25	2.0	.
38	161	2.0	.	62	2.0	.	26	2.0	.
39	168	2.0	.	65	2.0	.	27	2.0	.
40	175	2.0	.	66	2.0	.	29	2.0	.
41	1	1.0	.	68	2.0	.	42	2.0	.
42	2	1.0	.	72	2.0	.	43	2.0	.
43	3	1.0	.	73	2.0	.	44	2.0	.
44	4	1.0	.	78	2.0	.	48	2.0	.
45	5	1.0	.	83	2.0	.	51	2.0	.
46	7	1.0	.	87	2.0	.	53	2.0	.
47	8	1.0	.	90	2.0	.	54	2.0	.
48	9	1.0	.	94	2.0	.	56	2.0	.
49	10	1.0	.	95	2.0	.	58	2.0	.
50	11	1.0	.	112	2.0	.	61	2.0	.
51	12	1.0	.	118	2.0	.	62	2.0	.
52	13	1.0	.	120	2.0	CT	65	2.0	.

Table 4. 2006 VSA Intl. Competition. Violins, Tone, Round 2. Each judge's scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	37	3.0	ST	37	3.0	ST	26	3.0	.
2	60	3.0	2C	60	3.0	2C	34	3.0	CT
3	121	3.0	CT	120	3.0	CT	37	3.0	ST
4	136	3.0	ST	136	3.0	ST	60	3.0	2C
5	157	3.0	CT	157	3.0	CT	120	3.0	CT
6	158	3.0	CT	158	3.0	CT	121	3.0	CT
7	171	3.0	G	171	3.0	G	136	3.0	ST
8	195	3.0	CT	34	2.0	CT	158	3.0	CT
9	6	2.0	.	42	2.0	.	171	3.0	G
10	26	2.0	.	44	2.0	.	195	3.0	CT
11	34	2.0	CT	61	2.0	.	6	2.0	.
12	42	2.0	.	76	2.0	.	27	2.0	.
13	44	2.0	.	78	2.0	.	42	2.0	.
14	76	2.0	.	95	2.0	.	44	2.0	.
15	84	2.0	.	118	2.0	.	46	2.0	.
16	120	2.0	CT	125	2.0	.	61	2.0	.
17	125	2.0	.	145	2.0	.	73	2.0	.
18	161	2.0	.	161	2.0	.	76	2.0	.
19	3	1.0	.	169	2.0	.	78	2.0	.
20	16	1.0	.	195	2.0	CT	125	2.0	.
21	21	1.0	.	3	1.0	.	138	2.0	.
22	27	1.0	.	6	1.0	.	145	2.0	.
23	43	1.0	.	16	1.0	.	146	2.0	.
24	46	1.0	.	21	1.0	.	157	2.0	CT
25	55	1.0	.	26	1.0	.	161	2.0	.
26	61	1.0	.	27	1.0	.	169	2.0	.
27	67	1.0	.	43	1.0	.	3	1.0	.
28	73	1.0	.	46	1.0	.	16	1.0	.
29	78	1.0	.	55	1.0	.	21	1.0	.
30	83	1.0	.	67	1.0	.	43	1.0	.
31	86	1.0	.	73	1.0	.	55	1.0	.
32	95	1.0	.	83	1.0	.	67	1.0	.
33	110	1.0	.	84	1.0	.	83	1.0	.
34	118	1.0	.	86	1.0	.	84	1.0	.
35	130	1.0	.	110	1.0	.	86	1.0	.
36	135	1.0	.	121	1.0	CT	95	1.0	.
37	138	1.0	.	130	1.0	.	110	1.0	.
38	145	1.0	.	135	1.0	.	118	1.0	.
39	146	1.0	.	138	1.0	.	130	1.0	.
40	169	1.0	.	146	1.0	.	135	1.0	.

Table 5. 2006 VSA Intl. Competition. Violas, Workmanship, Round 1. Each judge's 52 highest scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	8	74.0	CW	70	75.6	CW	82	59.6	.
2	101	73.6	CW	55	75.2	G	16	58.4	SW
3	45	73.2	G	16	74.8	SW	101	57.6	CW
4	41	70.4	.	13	73.6	CW	65	56.8	2C
5	105	69.6	2C	29	72.8	SW	45	56.0	G
6	50	69.2	.	8	71.6	CW	70	56.0	CW
7	85	68.8	.	65	71.2	2C	12	56.0	CW
8	27	68.4	.	1	69.2	CW	85	55.6	.
9	12	67.6	CW	12	67.6	CW	75	55.2	.
10	1	66.8	CW	45	66.8	G	38	55.2	.
11	76	65.6	.	27	66.4	.	76	55.2	.
12	70	64.8	CW	101	66.4	CW	1	54.4	CW
13	102	64.8	.	73	66.0	.	49	54.0	.
14	103	64.0	.	9	65.2	.	19	53.2	.
15	90	64.0	.	49	65.2	.	37	53.2	.
16	98	64.0	.	19	64.4	.	99	53.2	.
17	13	63.6	CW	37	63.6	.	55	53.2	G
18	19	63.2	.	11	62.8	.	18	53.2	.
19	73	63.2	.	105	62.8	2C	7	53.2	.
20	75	62.8	.	75	62.8	.	98	53.2	.
21	65	62.8	2C	41	61.6	.	105	52.4	2C
22	82	62.8	.	94	61.6	.	9	52.0	.
23	15	62.4	.	15	61.2	.	79	51.6	.
24	37	62.4	.	104	61.2	.	60	51.6	.
25	55	61.6	G	98	59.6	.	36	50.0	.
26	32	61.6	.	38	59.6	.	2	50.0	.
27	29	61.2	SW	53	59.6	.	8	49.2	CW
28	16	61.2	SW	88	58.8	.	33	48.4	.
29	36	60.8	.	68	58.4	.	103	48.4	.
30	94	60.8	.	90	58.0	.	29	48.4	SW
31	99	60.0	.	7	58.0	.	11	48.4	.
32	14	58.8	.	32	58.0	.	4	48.4	.
33	38	58.4	.	60	57.6	.	91	48.4	.
34	40	58.0	.	10	56.0	.	106	48.4	.
35	9	57.2	.	85	54.8	.	28	48.4	.
36	91	56.8	.	96	54.4	.	89	48.4	.
37	49	56.4	.	51	54.4	.	72	48.4	.
38	104	56.0	.	82	53.6	.	13	48.4	CW
39	83	56.0	.	83	53.2	.	41	48.0	.
40	7	55.6	.	28	52.8	.	15	47.6	.
41	2	54.8	.	2	52.4	.	83	47.6	.
42	53	53.6	.	14	52.0	.	35	46.8	.
43	51	53.2	.	31	51.6	.	32	46.8	.
44	74	53.2	.	18	51.6	.	24	46.8	.
45	96	53.2	.	5	51.2	.	31	46.8	.
46	88	53.2	.	103	51.2	.	14	46.8	.
47	71	52.8	.	21	51.2	.	104	46.8	.
48	59	52.4	.	57	50.0	.	100	46.8	.
49	60	52.0	.	102	50.0	.	109	46.8	.
50	79	50.8	.	74	49.6	.	17	46.4	.
51	69	50.8	.	33	49.2	.	96	46.0	.
52	68	50.8	.	4	48.8	.	30	46.0	.

Table 6. 2006 VSA Intl. Competition. Violas, Tone, Round 1. Each judge's 52 highest scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	4	3.0	.	8	3.0	.	11	3.0	CT
2	9	3.0	.	17	3.0	.	15	3.0	.
3	14	3.0	.	33	3.0	CT	24	3.0	.
4	28	3.0	.	34	3.0	.	27	3.0	.
5	30	3.0	.	35	3.0	CT	33	3.0	CT
6	34	3.0	.	38	3.0	.	37	3.0	.
7	40	3.0	.	39	3.0	.	69	3.0	.
8	46	3.0	G	40	3.0	.	70	3.0	.
9	48	3.0	.	45	3.0	G	73	3.0	CT
10	66	3.0	.	50	3.0	.	74	3.0	.
11	69	3.0	.	69	3.0	.	94	3.0	ST
12	83	3.0	.	84	3.0	CT	96	3.0	.
13	103	3.0	.	88	3.0	.	98	3.0	.
14	106	3.0	.	94	3.0	ST	1	2.0	.
15	5	2.0	.	106	3.0	.	2	2.0	.
16	7	2.0	.	1	2.0	.	4	2.0	.
17	8	2.0	.	7	2.0	.	5	2.0	.
18	10	2.0	.	9	2.0	.	7	2.0	.
19	11	2.0	CT	10	2.0	.	8	2.0	.
20	12	2.0	.	11	2.0	CT	9	2.0	.
21	13	2.0	.	13	2.0	.	12	2.0	.
22	15	2.0	.	15	2.0	.	13	2.0	.
23	16	2.0	.	16	2.0	.	14	2.0	.
24	17	2.0	.	19	2.0	.	16	2.0	.
25	19	2.0	.	27	2.0	.	19	2.0	.
26	20	2.0	.	29	2.0	.	21	2.0	.
27	21	2.0	.	31	2.0	CT	28	2.0	.
28	24	2.0	.	32	2.0	.	30	2.0	.
29	27	2.0	.	36	2.0	.	31	2.0	CT
30	29	2.0	.	37	2.0	.	32	2.0	.
31	31	2.0	CT	41	2.0	.	35	2.0	CT
32	32	2.0	.	43	2.0	.	39	2.0	.
33	33	2.0	CT	46	2.0	.	40	2.0	.
34	35	2.0	CT	48	2.0	.	41	2.0	.
35	36	2.0	.	53	2.0	.	45	2.0	G
36	37	2.0	.	54	2.0	.	48	2.0	.
37	39	2.0	.	55	2.0	G	50	2.0	.
38	41	2.0	.	57	2.0	.	51	2.0	.
39	43	2.0	.	62	2.0	.	54	2.0	.
40	46	2.0	.	65	2.0	2C	55	2.0	G
41	50	2.0	.	66	2.0	.	57	2.0	.
42	55	2.0	G	68	2.0	CT	62	2.0	.
43	58	2.0	.	70	2.0	.	65	2.0	2C
44	60	2.0	.	71	2.0	.	66	2.0	.
45	62	2.0	.	72	2.0	.	68	2.0	CT
46	64	2.0	.	73	2.0	CT	72	2.0	.
47	65	2.0	2C	74	2.0	.	75	2.0	.
48	68	2.0	CT	75	2.0	.	76	2.0	.
49	70	2.0	.	76	2.0	.	83	2.0	.
50	71	2.0	.	77	2.0	.	84	2.0	CT
51	72	2.0	.	79	2.0	.	85	2.0	.
52	73	2.0	CT	83	2.0	.	88	2.0	.

Table 7. 2006 VSA Intl. Competition. Violas, Tone, Round 2. Each judge's 52 highest scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	31	3.0	CT	11	3.0	CT	31	3.0	CT
2	33	3.0	CT	68	3.0	CT	33	3.0	CT
3	35	3.0	CT	84	3.0	CT	94	3.0	ST
4	46	3.0	G	94	3.0	ST	8	2.0	.
5	65	3.0	2C	107	3.0	.	9	2.0	.
6	94	3.0	ST	7	2.0	.	11	2.0	CT
7	55	2.0	G	8	2.0	.	14	2.0	.
8	66	2.0	.	9	2.0	.	28	2.0	.
9	68	2.0	CT	13	2.0	.	48	2.0	.
10	73	2.0	CT	14	2.0	.	50	2.0	.
11	84	2.0	CT	15	2.0	.	55	2.0	G
12	88	2.0	.	27	2.0	.	65	2.0	2C
13	105	2.0	2C	28	2.0	.	66	2.0	.
14	106	2.0	.	31	2.0	CT	68	2.0	CT
15	4	1.0	.	32	2.0	.	73	2.0	CT
16	7	1.0	.	33	2.0	CT	105	2.0	2C
17	8	1.0	.	35	2.0	CT	4	1.0	.
18	9	1.0	.	39	2.0	.	7	1.0	.
19	11	1.0	CT	40	2.0	.	13	1.0	.
20	13	1.0	.	45	2.0	G	15	1.0	.
21	14	1.0	.	55	2.0	G	16	1.0	.
22	15	1.0	.	65	2.0	2C	17	1.0	.
23	16	1.0	.	69	2.0	.	19	1.0	.
24	17	1.0	.	72	2.0	.	24	1.0	.
25	19	1.0	.	73	2.0	CT	27	1.0	.
26	24	1.0	.	74	2.0	.	30	1.0	.
27	27	1.0	.	75	2.0	.	32	1.0	.
28	28	1.0	.	76	2.0	.	34	1.0	.
29	30	1.0	.	88	2.0	.	35	1.0	CT
30	32	1.0	.	91	2.0	.	37	1.0	.
31	34	1.0	.	98	2.0	.	39	1.0	.
32	37	1.0	.	101	2.0	.	40	1.0	.
33	39	1.0	.	103	2.0	.	41	1.0	.
34	40	1.0	.	105	2.0	2C	45	1.0	G
35	41	1.0	.	106	2.0	.	62	1.0	.
36	48	1.0	.	4	1.0	.	69	1.0	.
37	50	1.0	.	16	1.0	.	70	1.0	.
38	62	1.0	.	17	1.0	.	72	1.0	.
39	69	1.0	.	19	1.0	.	74	1.0	.
40	70	1.0	.	24	1.0	.	75	1.0	.
41	72	1.0	.	30	1.0	.	76	1.0	.
42	74	1.0	.	34	1.0	.	83	1.0	.
43	75	1.0	.	37	1.0	.	84	1.0	CT
44	76	1.0	.	41	1.0	.	85	1.0	.
45	83	1.0	.	48	1.0	.	88	1.0	.
46	85	1.0	.	50	1.0	.	91	1.0	.
47	91	1.0	.	62	1.0	.	93	1.0	.
48	93	1.0	.	66	1.0	.	96	1.0	.
49	96	1.0	.	70	1.0	.	98	1.0	.
50	98	1.0	.	83	1.0	.	101	1.0	.
51	101	1.0	.	85	1.0	.	103	1.0	.
52	103	1.0	.	93	1.0	.	104	1.0	.

Table 8. 2006 VSA Intl. Competition. Cellos, Workmanship, Round 1. Each judge's scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	19	70.0	CW	2	79.2	SW	2	62.4	SW
2	2	68.4	SW	51	79.2	G	15	62.4	.
3	6	67.2	CW	28	74.8	CW	3	61.6	.
4	46	66.8	.	44	71.2	CW	43	61.6	.
5	13	64.8	.	43	71.2	.	38	61.2	.
6	44	64.4	CW	50	71.2	.	9	60.8	.
7	32	62.8	.	21	70.4	G	22	59.2	.
8	51	62.4	G	19	69.6	CW	51	58.4	G
9	38	62.4	.	47	69.6	.	19	58.4	CW
10	22	62.4	.	10	68.0	CW	50	57.6	.
11	36	62.0	.	22	67.2	.	40	57.2	.
12	43	61.2	.	12	65.6	.	35	57.2	.
13	10	60.8	CW	9	65.2	.	21	56.8	G
14	49	58.8	.	36	64.4	.	28	56.4	CW
15	12	58.4	.	6	63.6	CW	47	56.4	.
16	21	58.4	G	38	63.2	.	30	56.0	.
17	52	58.0	.	32	62.4	.	46	55.6	.
18	14	57.2	.	30	61.2	.	10	55.6	CW
19	46	57.2	.	33	61.2	.	20	55.6	.
20	47	56.8	.	16	58.8	.	16	55.2	.
21	15	56.8	.	49	58.4	.	52	55.2	.
22	50	56.4	.	15	58.4	.	12	54.8	.
23	3	56.0	.	42	57.6	.	44	54.8	CW
24	1	55.6	.	39	57.6	.	32	54.8	.
25	33	54.8	.	52	55.2	.	49	54.8	.
26	9	54.8	.	35	55.2	.	33	54.0	.
27	16	54.4	.	27	52.8	.	17	54.0	.
28	18	54.4	.	3	52.8	.	39	54.0	.
29	7	54.0	.	17	52.8	.	36	53.2	.
30	42	54.0	.	46	52.8	.	6	53.2	CW
31	23	52.4	.	18	51.6	.	27	53.2	.
32	28	51.6	CW	34	50.8	.	45	53.2	.
33	35	51.6	.	5	50.8	.	42	51.6	.
34	5	50.8	.	40	50.8	.	18	50.8	.
35	40	49.2	.	7	49.2	.	25	49.2	.
36	41	49.2	.	23	48.4	.	7	49.2	.
37	30	48.8	.	13	46.8	.	34	48.4	.
38	27	48.4	.	4	44.8	.	41	48.4	.
39	25	48.4	.	25	44.8	.	23	47.6	.
40	17	48.0	.	1	44.4	.	13	47.6	.
41	8	47.2	.	20	43.6	.	24	47.6	.
42	20	46.4	.	24	43.6	.	53	47.2	.
43	39	45.6	.	45	43.6	.	1	46.4	.
44	34	43.6	.	14	42.8	.	55	46.4	.
45	24	43.2	.	8	42.8	.	8	45.6	.
46	53	42.8	.	41	40.4	.	14	44.8	.
47	4	29.2	.	53	32.0	.	5	42.0	.
48	55	28.0	.	55	18.8	.	4	40.0	.

Table 9. 2006 VSA Intl. Competition. Cellos, Tone, Round 1. Each judge's scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	2	3.0	.	15	3.0	ST	1	3.0	.
2	3	3.0	.	28	3.0	.	7	3.0	.
3	7	3.0	.	30	3.0	.	15	3.0	ST
4	8	3.0	ST	38	3.0	CT	17	3.0	.
5	15	3.0	ST	39	3.0	.	23	3.0	.
6	19	3.0	.	41	3.0	.	28	3.0	.
7	21	3.0	G	42	3.0	.	35	3.0	CT
8	23	3.0	.	43	3.0	.	51	3.0	G
9	24	3.0	.	45	3.0	CT	2	2.0	.
10	27	3.0	.	46	3.0	.	3	2.0	.
11	28	3.0	.	47	3.0	ST	4	2.0	.
12	32	3.0	.	50	3.0	ST	6	2.0	.
13	33	3.0	.	51	3.0	G	8	2.0	ST
14	34	3.0	.	52	3.0	.	9	2.0	.
15	35	3.0	CT	1	2.0	.	10	2.0	.
16	38	3.0	CT	2	2.0	.	12	2.0	.
17	39	3.0	.	7	2.0	.	13	2.0	.
18	47	3.0	ST	8	2.0	ST	16	2.0	.
19	53	3.0	.	9	2.0	.	18	2.0	.
20	1	2.0	.	10	2.0	.	19	2.0	.
21	4	2.0	.	13	2.0	.	21	2.0	G
22	6	2.0	.	19	2.0	.	22	2.0	CT
23	9	2.0	.	21	2.0	G	25	2.0	.
24	10	2.0	.	22	2.0	CT	27	2.0	.
25	13	2.0	.	32	2.0	.	30	2.0	.
26	17	2.0	.	36	2.0	.	32	2.0	.
27	20	2.0	.	40	2.0	.	33	2.0	.
28	22	2.0	CT	49	2.0	.	34	2.0	.
29	25	2.0	.	53	2.0	.	36	2.0	.
30	36	2.0	.	3	1.0	.	38	2.0	CT
31	40	2.0	.	4	1.0	.	39	2.0	.
32	44	2.0	.	5	1.0	.	40	2.0	.
33	45	2.0	CT	6	1.0	.	42	2.0	.
34	46	2.0	.	12	1.0	.	43	2.0	.
35	49	2.0	.	14	1.0	.	44	2.0	.
36	50	2.0	ST	16	1.0	.	45	2.0	CT
37	51	2.0	G	17	1.0	.	46	2.0	.
38	52	2.0	.	18	1.0	.	47	2.0	ST
39	5	1.0	.	20	1.0	.	49	2.0	.
40	12	1.0	.	23	1.0	.	50	2.0	ST
41	14	1.0	.	24	1.0	.	52	2.0	.
42	16	1.0	.	25	1.0	.	53	2.0	.
43	18	1.0	.	27	1.0	.	5	1.0	.
44	30	1.0	.	33	1.0	.	14	1.0	.
45	41	1.0	.	34	1.0	.	20	1.0	.
46	42	1.0	.	35	1.0	CT	24	1.0	.
47	43	1.0	.	44	1.0	.	41	1.0	.
48	55	1.0	.	55	1.0	.	55	1.0	.

Table 10. 2006 VSA Intl. Competition. Cellos, Tone, Round 2. Each judge's scores are listed in descending order.

	JUDGE 1			JUDGE 2			JUDGE 3		
	Black No	Score	Award	Black No	Score	Award	Black No	Score	Award
1	2	3.0	.	1	3.0	.	8	3.0	ST
2	3	3.0	.	15	3.0	ST	15	3.0	ST
3	8	3.0	ST	21	3.0	G	17	3.0	.
4	19	3.0	.	22	3.0	CT	21	3.0	G
5	21	3.0	G	35	3.0	CT	35	3.0	CT
6	22	3.0	CT	38	3.0	CT	38	3.0	CT
7	35	3.0	CT	45	3.0	CT	45	3.0	CT
8	38	3.0	CT	47	3.0	ST	47	3.0	ST
9	39	3.0	.	50	3.0	ST	51	3.0	G
10	47	3.0	ST	51	3.0	G	2	2.0	.
11	50	3.0	ST	2	2.0	.	3	2.0	.
12	51	3.0	G	3	2.0	.	7	2.0	.
13	15	2.0	ST	7	2.0	.	19	2.0	.
14	17	2.0	.	8	2.0	ST	22	2.0	CT
15	23	2.0	.	10	2.0	.	33	2.0	.
16	27	2.0	.	17	2.0	.	36	2.0	.
17	28	2.0	.	19	2.0	.	42	2.0	.
18	33	2.0	.	28	2.0	.	43	2.0	.
19	36	2.0	.	38	2.0	.	46	2.0	.
20	42	2.0	.	39	2.0	.	50	2.0	ST
21	43	2.0	.	42	2.0	.	1	1.0	.
22	45	2.0	CT	43	2.0	.	9	1.0	.
23	52	2.0	.	46	2.0	.	10	1.0	.
24	53	2.0	.	9	1.0	.	13	1.0	.
25	1	1.0	.	13	1.0	.	23	1.0	.
26	7	1.0	.	23	1.0	.	27	1.0	.
27	9	1.0	.	27	1.0	.	28	1.0	.
28	10	1.0	.	30	1.0	.	30	1.0	.
29	13	1.0	.	32	1.0	.	32	1.0	.
30	30	1.0	.	33	1.0	.	34	1.0	.
31	32	1.0	.	34	1.0	.	39	1.0	.
32	34	1.0	.	40	1.0	.	40	1.0	.
33	40	1.0	.	49	1.0	.	49	1.0	.
34	46	1.0	.	52	1.0	.	52	1.0	.
35	49	1.0	.	53	1.0	.	53	1.0	.

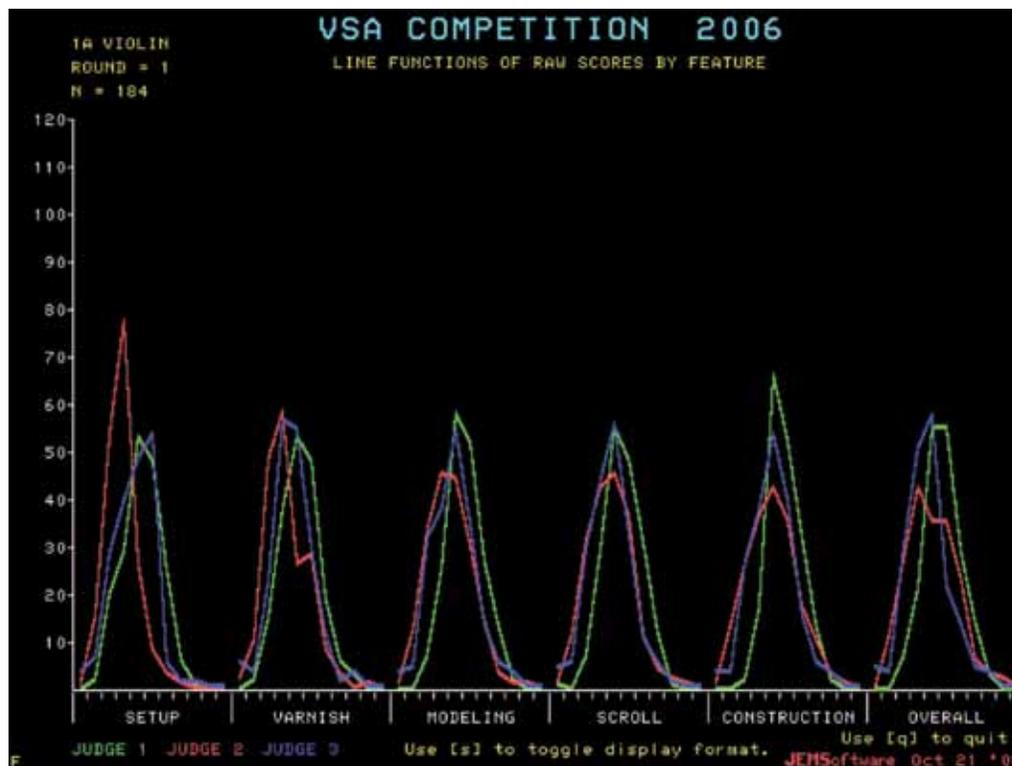


Figure 1. 2006 VSA Intl. Competition, Violins, Round 1. Workmanship scores (line functions of raw scores, overall and for five categories) by three judges for 184 violins.

of instruments, the average score for any category should be 5.0. Out of a point spread from 0 to 10, a 5 would be in the middle, so that would be average. Anything above got a score between 5 and 10. Anything below received a score from 0 to 5. You can see that the distributions pretty much fall into the center. The color designates the judge.

Two judges apparently were in closer agreement than the third judge, whose scoring mean was way below 5.0, but it's a consistent mean. It indicates that that judge was not so impressed with the general quality of the population. From a statistical evaluation, it really didn't make any difference, because that judge in red would penalize every instrument identically. Lowering the point score would not affect the outcome of the competition.

This is another way of looking at the distribution, and we're still looking at workmanship (Fig. 2). Again, the colors indicate the different judges. They all have slightly different ideas as to what an average instrument looks like. Plotting the data differently, these are the total scores, and the white line is the average score from all three judges combined (Fig. 3). You can see that there are very consistent trends. The judges separated out instruments equally, again independently.

These are the distributions of violin tone scores (Fig. 4): 3s for the better instruments, then 2s and the 1s. For the basses we didn't have a large enough statistical base to do any meaningful analysis. There were different ideas as to how to evaluate these instruments. There were a large number of instruments in the violin section that simply did not meet the criteria of the judges, and so they got eliminated rather early. There was a different kind of distribution for the violas. There were many more "undecideds." For the cellos, it was some of everything.

This next chart (Fig. 5) is the most significant one in this presentation, and I'll show you several more like it. The top is a plot of the tone scores of all the violins as evaluated by the tone judges. The bottom numbers are workmanship scores, including all 184 violins. The colored lines are the respective placement of a particular instrument on one scale and its corresponding ranking on the other. Look how tightly these instruments are distributed here. These are the top instruments identified by the workmanship judges. These up here are the top instruments identified by the tone judges, and the lines in between. One particular violin was highly ranked by the tone judges, but the workmanship

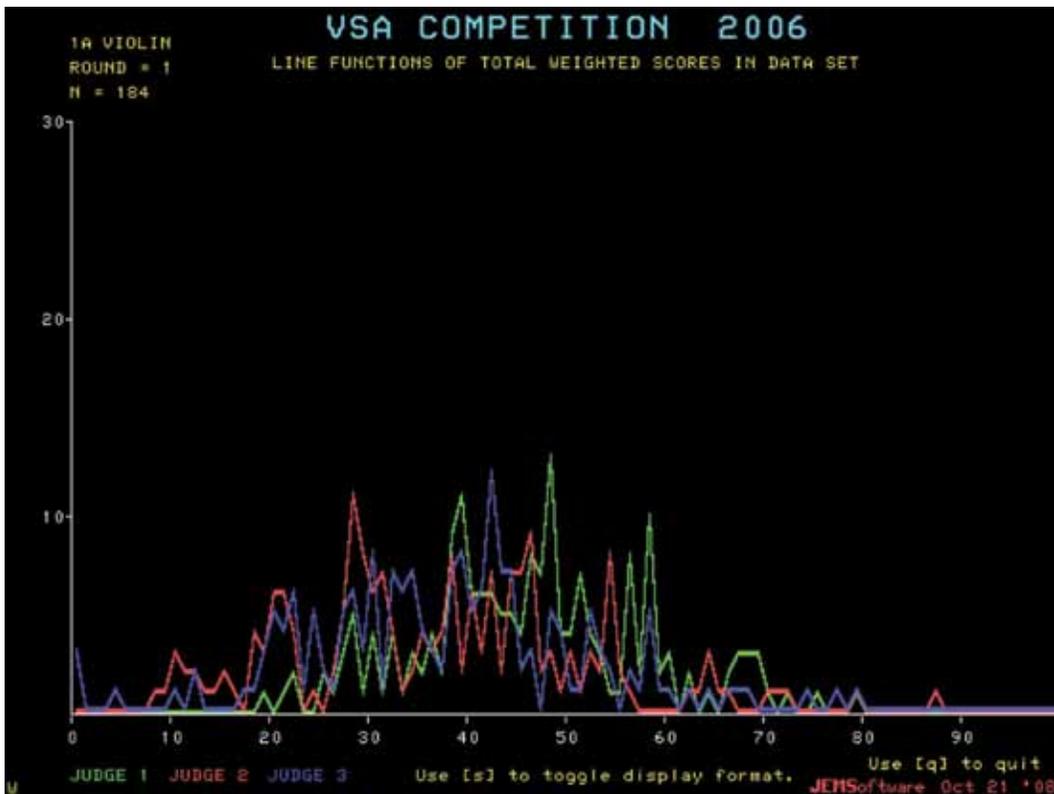


Figure 2. 2006 VSA Intl. Competition, Violins, Round 1. Workmanship scores (line functions of total weighted scores in data set) by three judges for 184 violins.

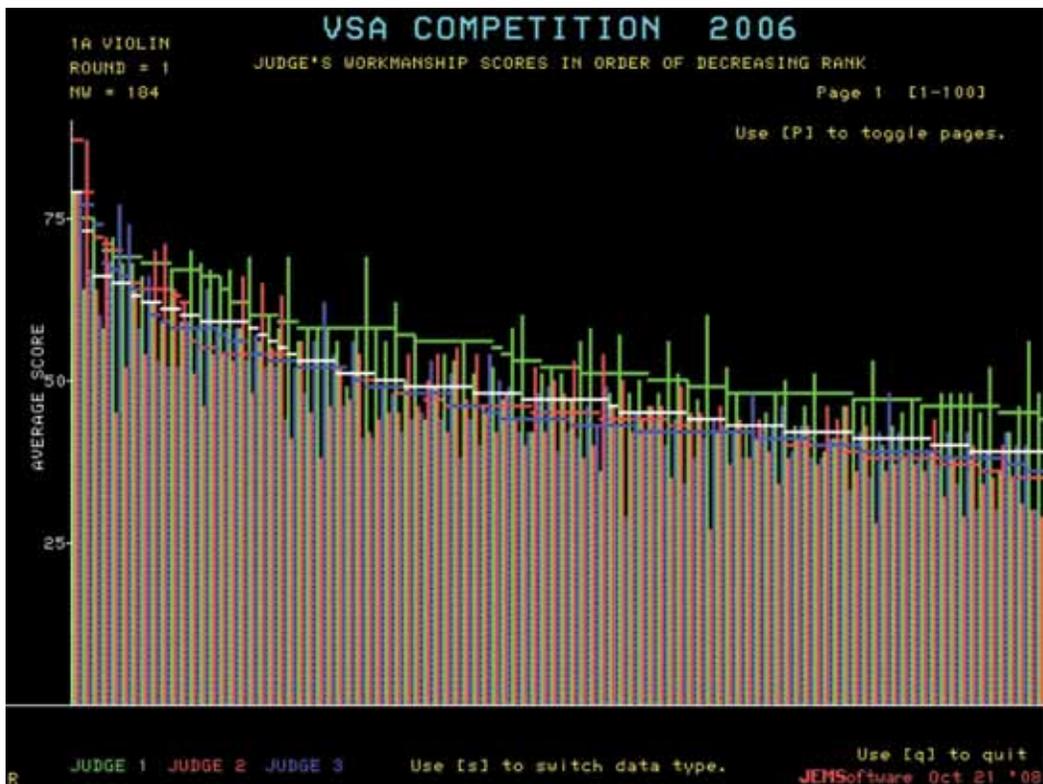


Figure 3. 2006 VSA Intl. Competition, Violins, Round 1. Judges' Workmanship scores, plotted in order of decreasing rank, for 184 violins. (The cascading white horizontal bars are the average scores of the three judges.)

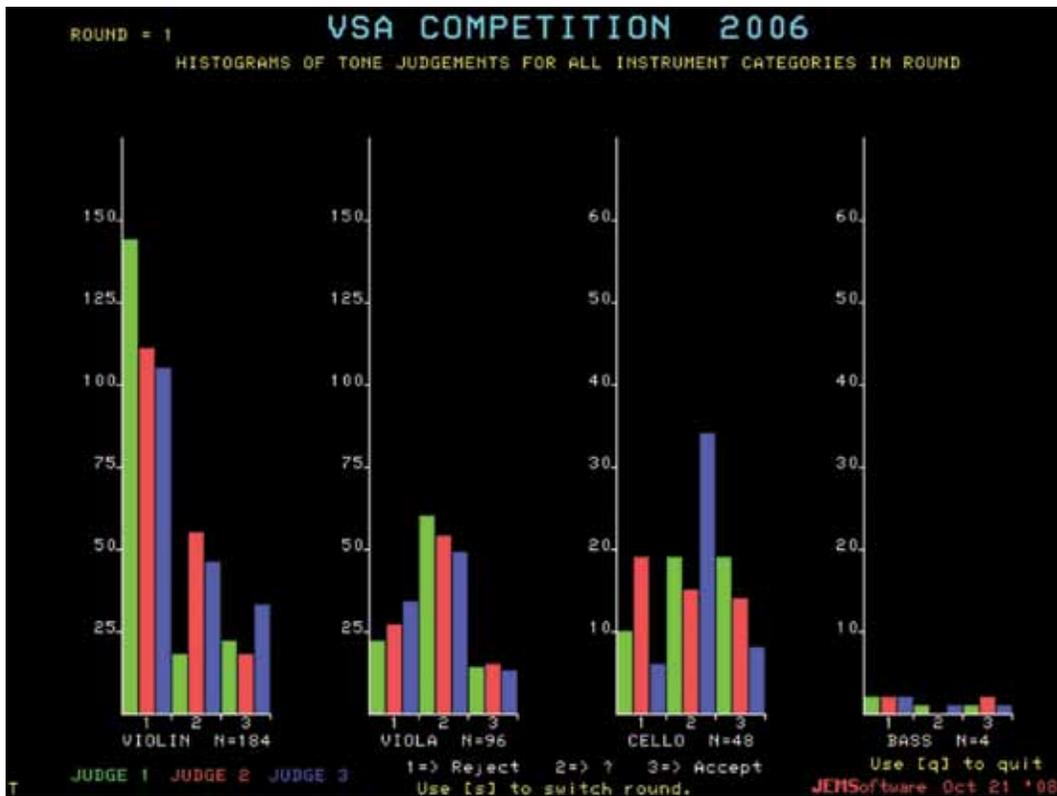


Figure 4. 2006 VSA Intl. Competition, Round 1. Histograms of Tone judgments by three judges for 184 violins, 96 violas, 48 cellos, and 4 basses.



Figure 5. 2006 VSA Intl. Competition, Violins, Round 1. Ranked average Workmanship scores (average of three judges) related to Tone Levels (medals and certificates of merit).

judges scored it almost at the very bottom. Look at all these lines. They fall all over the scale. A gold medal-winning violin is indicated.

Why don't we have more gold medals? Because the instruments don't qualify to receive gold medals. There is something fundamentally lacking that is the responsibility of the maker that needs to be fixed to increase the number of medals that we can give out.

Another way of looking at this composite data is shown in Fig. 6. The scale here is the ranking of the tone, with high being good, and this is the ranking of the workmanship, where 100 percent is best. You see that this is just a random distribution of points. There's only one really good instrument, and it's right up there. That's the gold medal winner.

This next chart is for violins, Round 2 with a population of 21 instruments (Fig. 7). If you look at the center of the distribution on all of these, it is significantly higher than 5.0, the average value that we had set before. The judges were instructed to erase everything that they had seen before and make new evaluations. For this population of 21 instruments, the average good instrument is a 5. So you use the whole tone spread. Well, the judges were more impressed, or maybe they didn't understand that mission correctly, so they gave much higher scores than they did in Round 1. I don't have any problems with that because every one of them did that equally well.

This is the same data distribution (Fig. 8). You have the white cascading line, sort of being the average of the three, and you can see that there are sometimes vastly different approaches to evaluating an instrument. The green judge was much more generous than the red and blue judges, but he advantaged everybody equally. So when you look at averages, it does not play into any kind of preferential or detrimental treatment.

This is Round 2 of the tone (Fig. 9). And here you have more instruments eliminated. These are the final ones that made it into Round 3. For the viola section, a very large number of instruments were eliminated and you end up with a few, but these are agreed on. And for the cellos, the instrument distribution is almost the same for each one of the three categories.

This is violin, Round 2 (Fig. 10). Now this is an interesting chart. You see all these little pimples up here? Each one of these pimples is an instrument that got into Round 2. But most of these don't have tails, as most over there do. The reason for that is

that these instruments never leave the counter part, never made it to Round 2. And this is only Round 2 data. They're not even on the chart. And, again, these two instruments are the ones that we're looking for. These are the gold medal quality instruments. The distribution and the performance here are exactly identical to that in the violins.

This is Round 1 for violas (Fig. 11). Look at that instrument here. It's at the very bottom of the workmanship score, but it's way up there on the tonal score. The tone judges don't evaluate the appearance of the instruments, they just play them, and if they sound good, they give them a good rating. The workmanship judges do look at them. If they look beautiful, they give them a high rating. But the correlation between the two is very poor, if nonexistent.

This is cellos, Round 1 (Fig. 12). There is the gold medal instrument. This is the cello, the data distribution in Round 1. You can see that one of the judges was more generous than the other two. The other two sort of fell on top of one another. But once you populate your average, everybody is equally influenced by that difference.

Going back to the previous competition in 2004, I will not show you all the in-between charts now. I'm only giving you the results to give you another level of confirmation. It is real and it is supported by large databases. This is now the competition in 2004, violins, Round 1 (Figs. 13 and 14). Look at that tight distribution and see what happens to that on the total scale. And you have the same thing up here. Same violin data, 182 data points. You can see that the average well satisfies everybody.

These are the results for the violas, Round 1 in 2004 (Figs. 15 and 16). We awarded gold medals to two instruments shown here. Some of the other violas were very close.

Now, let me give you a different aside, because this is a puzzling phenomenon. You have all these good makers making instruments, and yet so few of them satisfy both tone and workmanship requirements. The curious part is that largely the same people win these medals. The conclusion I come to is that there are some intelligent violin-makers who really know how to make good looking and good sounding instruments. These are the makers we should look for and emulate and pick their brains as to what it is that they do so much better than anybody else. That is a message that I give to makers who enter competitions and are trying to get better awards.

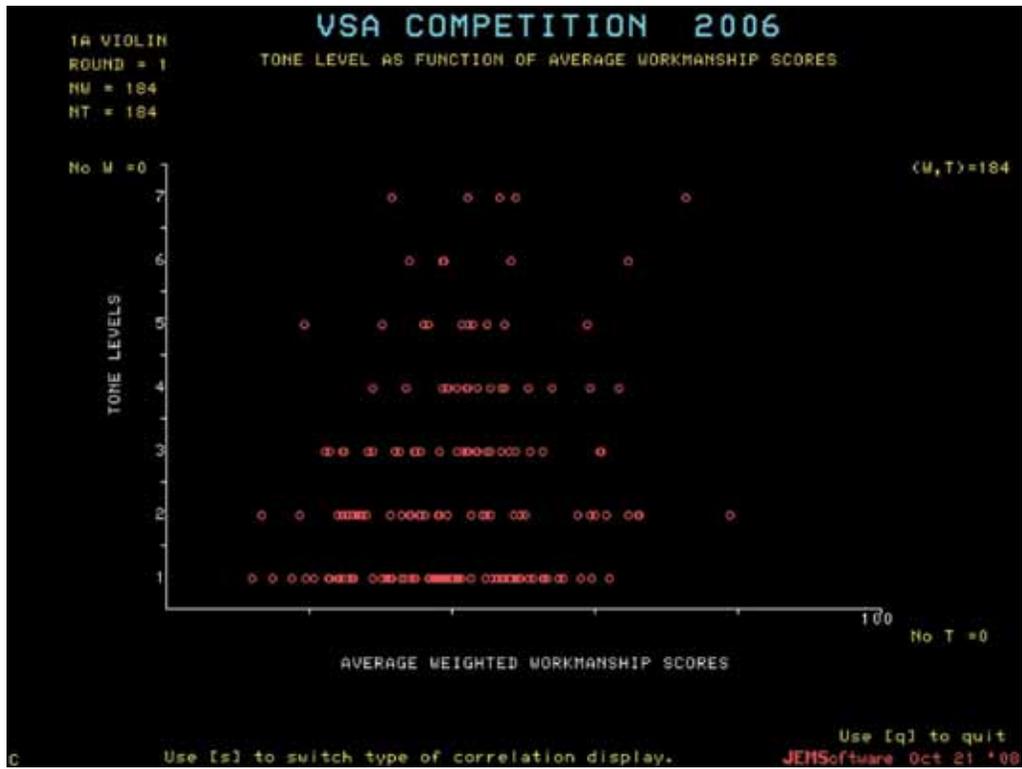


Figure 6. 2006 VSA Intl. Competition, Violins, Round 1. Tone Level versus average Workmanship score (average of three judges) for 184 violins.

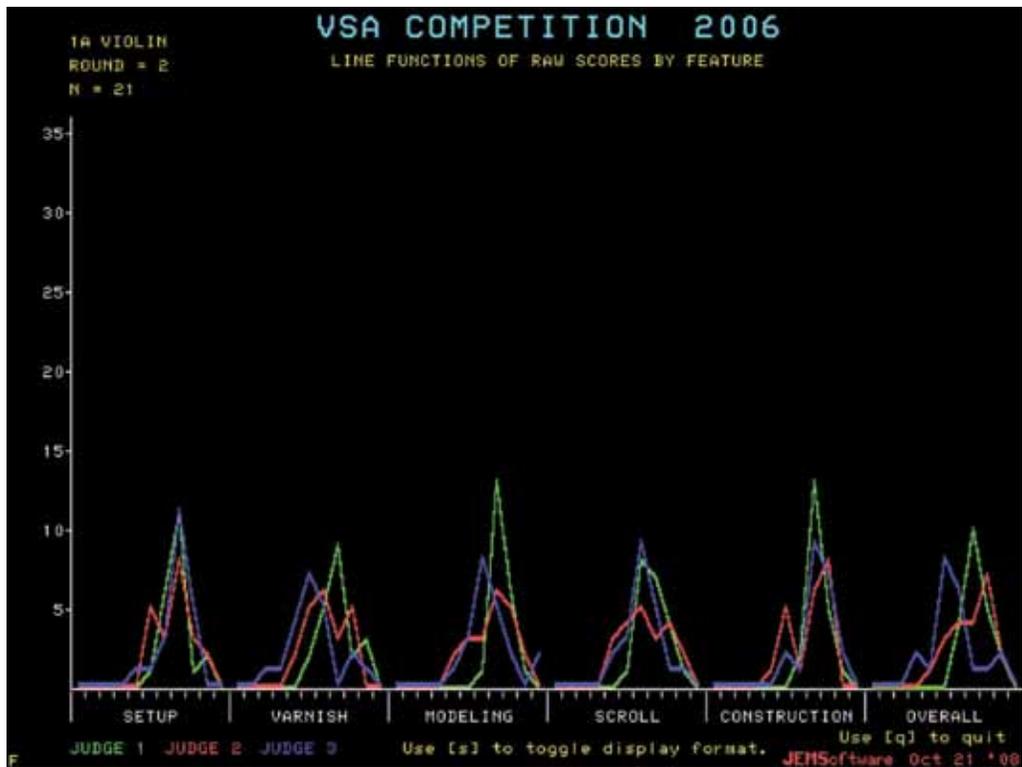


Figure 7. 2006 VSA Intl. Competition, Violins, Round 2. Workmanship scores (line functions of raw scores, overall and for five categories) by three judges for 21 violins.



Figure 8. 2006 VSA Intl. Competition, Violins, Round 2. Judges' Workmanship scores, plotted in order of decreasing rank, for 21 violins. (The cascading white horizontal bars are the average scores of the three judges.)

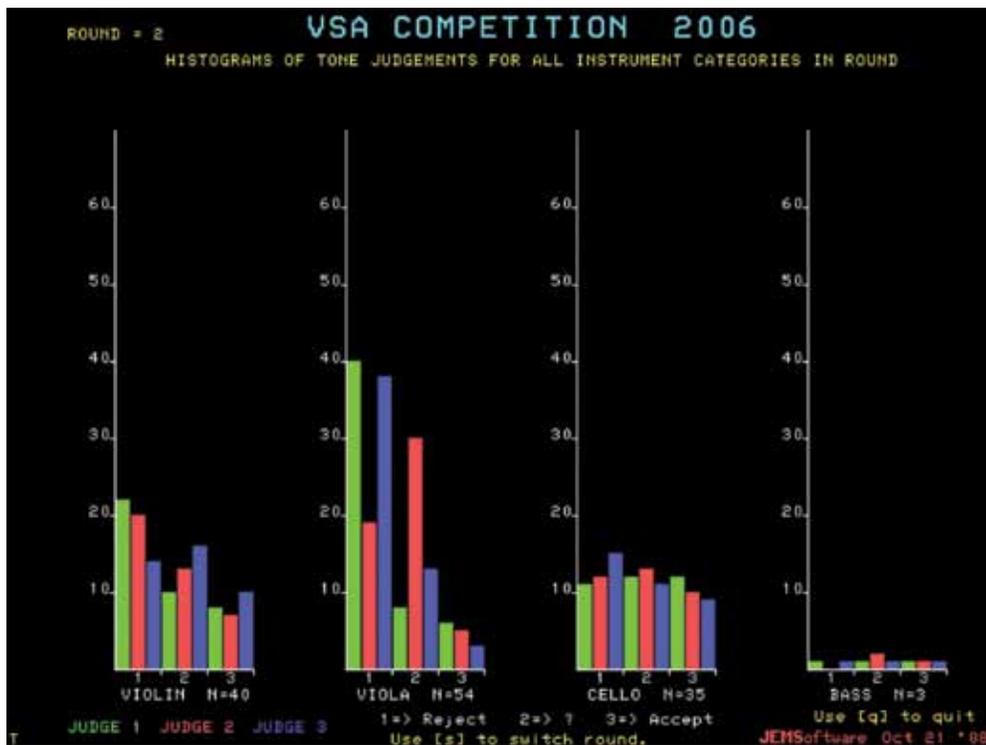


Figure 9. 2006 VSA Intl. Competition, Round 2. Histograms of Tone judgments by three judges for 40 violins, 54 violas, 35 cellos, and 3 basses.



Figure 10. 2006 VSA Intl. Competition, Violins, Round 2. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).

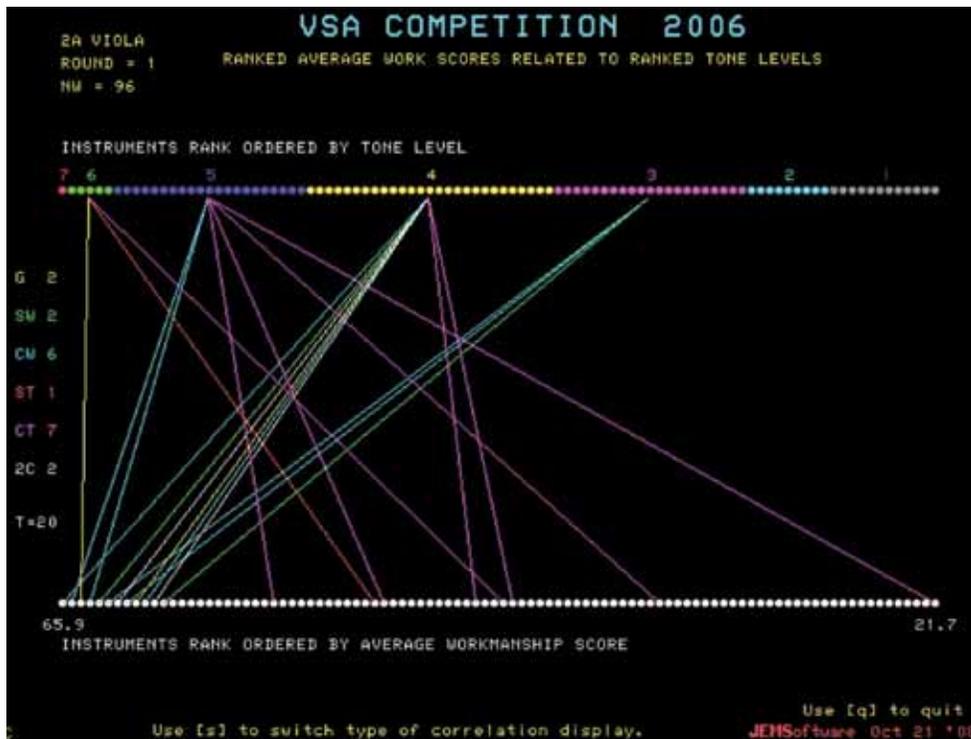


Figure 11. 2006 VSA Intl. Competition, Violas, Round 1. Ranked average Workmanship scores (average of three judges) related to Tone Levels (medals and certificates of merit).

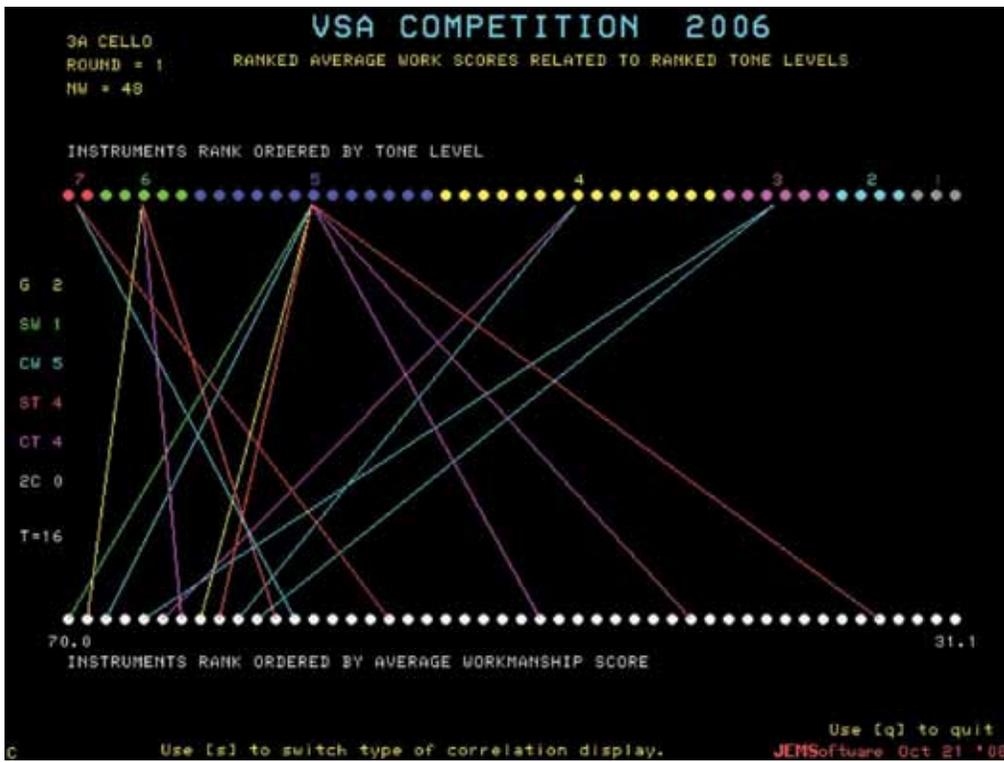


Figure 12. 2006 VSA Intl. Competition, Cellos, Round 1. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).

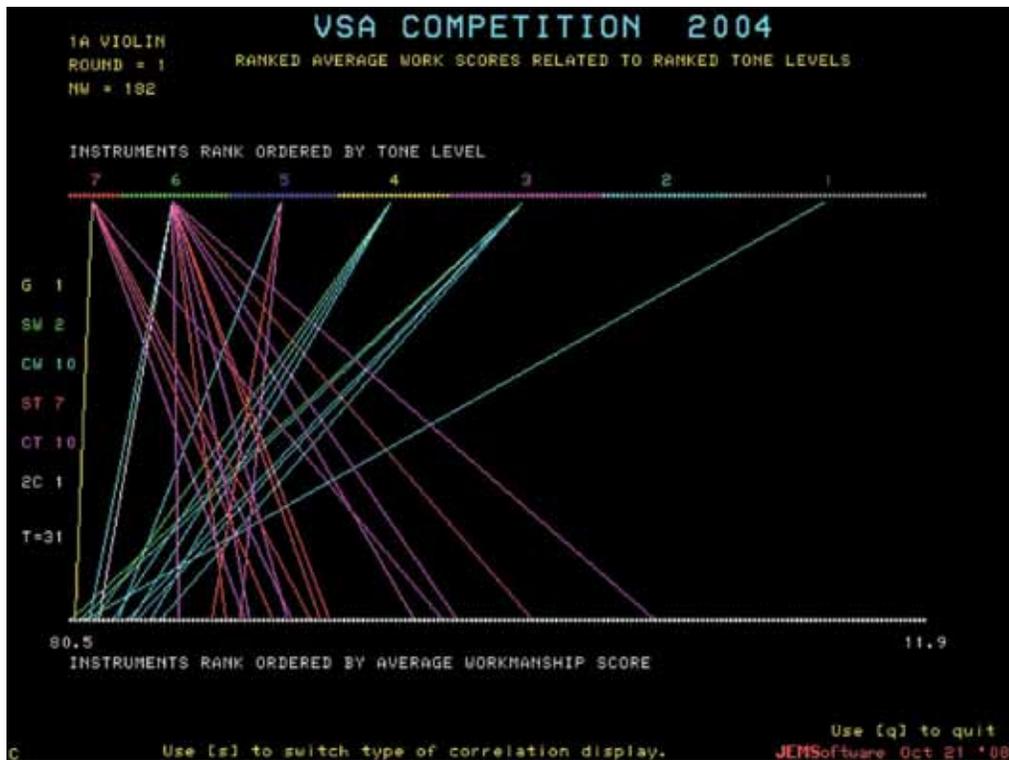


Figure 13. 2004 VSA Intl. Competition, Violins, Round 1. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).

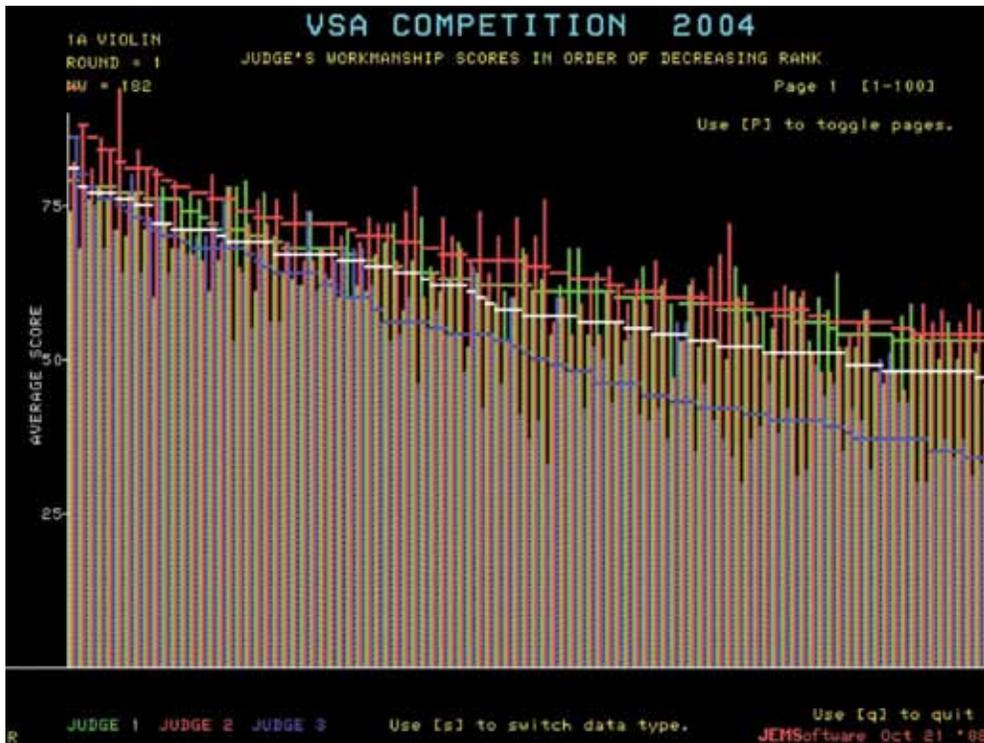


Figure 14. 2004 VSA Intl. Competition, Violins, Round 1. Judges' Workmanship scores, plotted in order of decreasing rank, of 182 violins. (The cascading white horizontal bars are the average scores of the three judges.)



Figure 15. 2004 VSA Intl. Competition, Violas, Round 1. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).



Figure 16. 2004 VSA Intl. Competition, Violas, Round 1. Judges' Workmanship scores, plotted in order of decreasing rank, for 85 violas. (The cascading white horizontal bars are the average scores of the three judges.)

Figure 18 shows the cello results in Round 1 of 2004 with the gold medal cello indicated. For comparison, take a look at the data for 2002. The data for violin, Round 1, has the same distribution (Figs. 19 and 20). The data look regular.

Next are the results for viola in 2002 when no gold medal was awarded. I recall the uproar from the audience when that was announced at the awards banquet. No gold medal? Boo! Boo! Who was blamed for that? Initially, the competition committee was blamed for not knowing how to run a competition. But when the judges' scores were examined, it was understandable why a gold medal wasn't awarded. The judges made those decisions, not the computer or the competition committee that massages all these data points. Here are the 2002 results for viola, Round 1 (Figs. 21 and 22), and cello, Round 1 (Figs. 23 and 24). There is the gold medal cello and the distribution of the data.

Conclusions: First, all the judges that we have tested in three competitions were competent and independently identified the ultimate winning instruments within a very small space of the distribution. There is no or little correlation between

what an instrument looks like and what it sounds like. And some makers have won gold medals repeatedly. They know something that we don't know. Let's pick their brains and find out what that is.

My final conclusion is that the lack of gold medals has nothing to do with the competition process or the qualification of the judges. The makers have to work with players and, of course, conversely too, to increase the tonal performance of their instruments. The competition design is correct and identifies the correct instruments.

Tom King: Can we see one of those charts again for Round 1? What I wanted to see were the lines that show the workmanship on the bottom line and the tone on the top line. Your conclusion there was that the judges correctly identified the instruments in the first round. I didn't notice that myself, so I'm presuming that there is a yellow line for the first round results that would be high for workmanship and high for tone in that round.

Dr. Regh: Tell me to stop when you see what you're looking for.

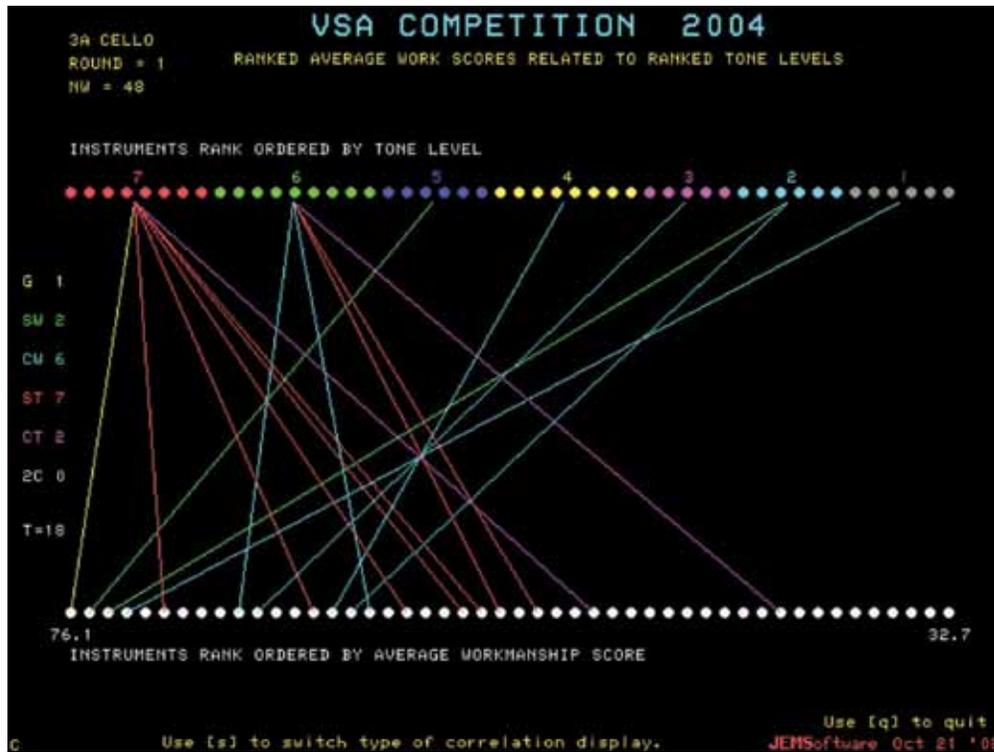


Figure 17. 2004 VSA Intl. Competition, Cellos, Round 1. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).



Figure 18. 2004 VSA Intl. Competition, Cellos, Round 1. Judges' Workmanship scores, plotted in order of decreasing rank, for 48 cellos. (The cascading white horizontal bars are the average scores of the three judges.)



Figure 19. 2002 VSA Intl. Competition, Violins, Round 1. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).

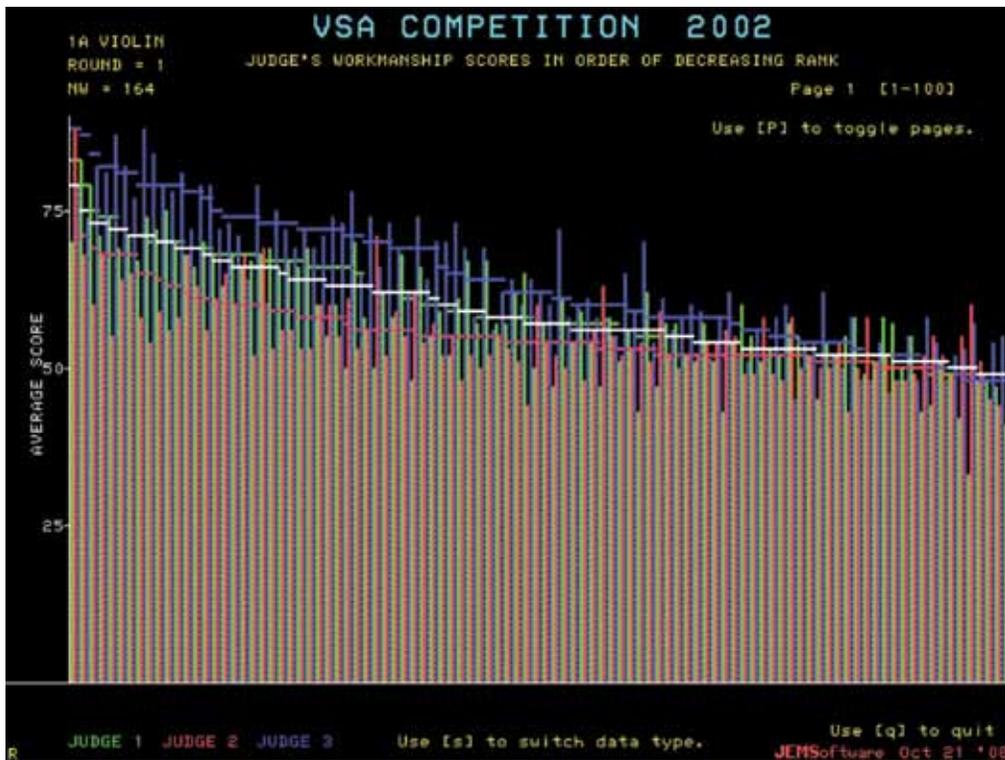


Figure 20. 2002 VSA Intl. Competition, Violins, Round 1. Judges' Workmanship scores, plotted in order of decreasing rank, for 164 violins. (The cascading white horizontal bars are the average scores of the three judges.)

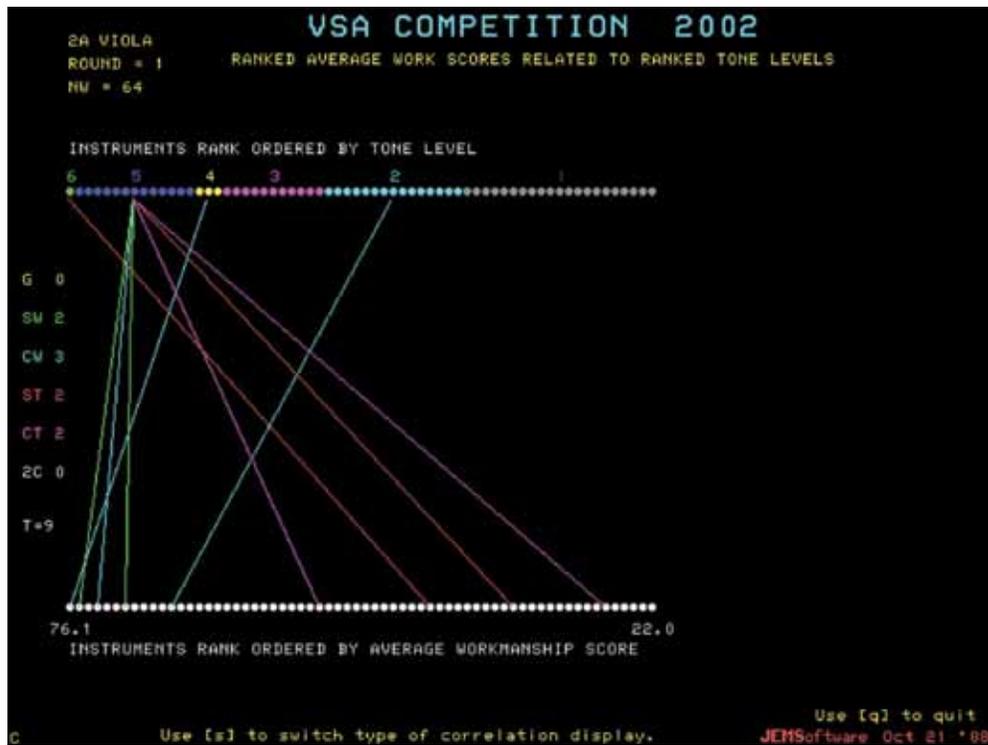


Figure 21. 2002 VSA Intl. Competition, Violas, Round 1. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).

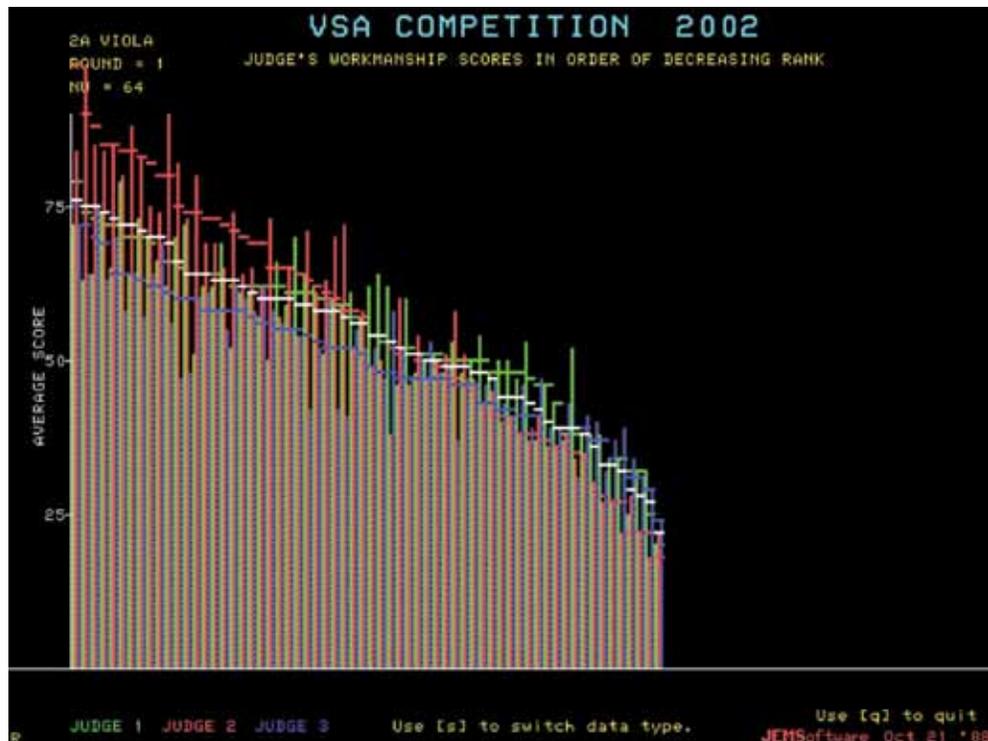


Figure 22. 2002 VSA Intl. Competition, Violas, Round 1. Judges' Workmanship scores, plotted in order of decreasing rank, for 64 violas. (The cascading white horizontal bars are the average scores of the three judges.)

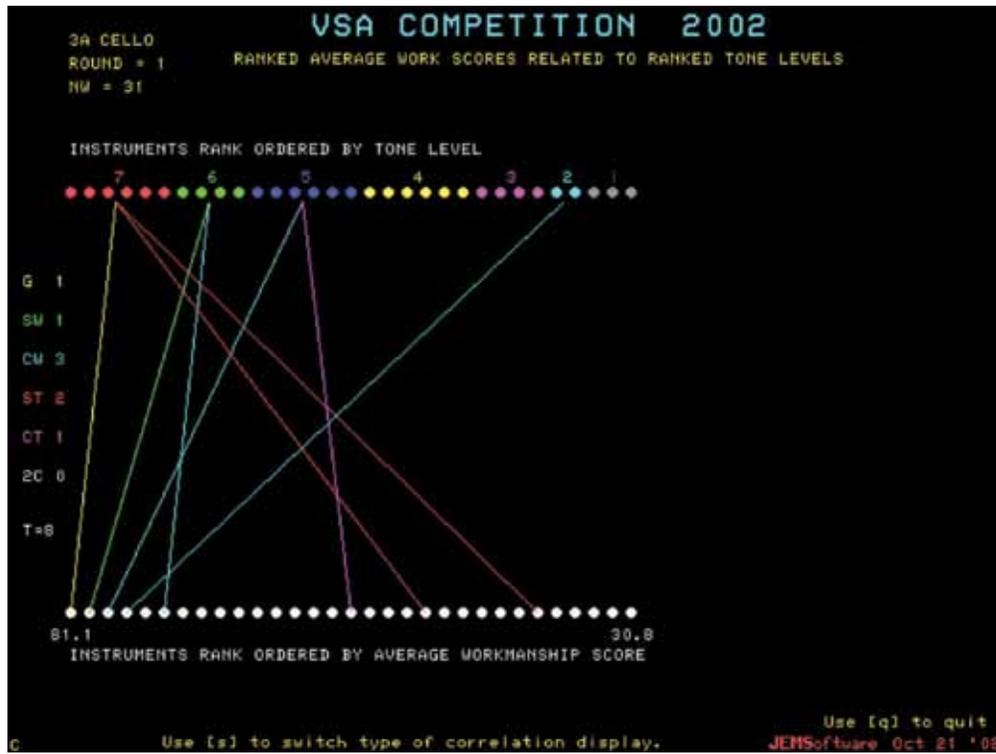


Figure 23. 2002 VSA Intl. Competition, Cellos, Round 1. Ranked average Workmanship scores (average of three judges) related to ranked Tone Levels (medals and certificates of merit).

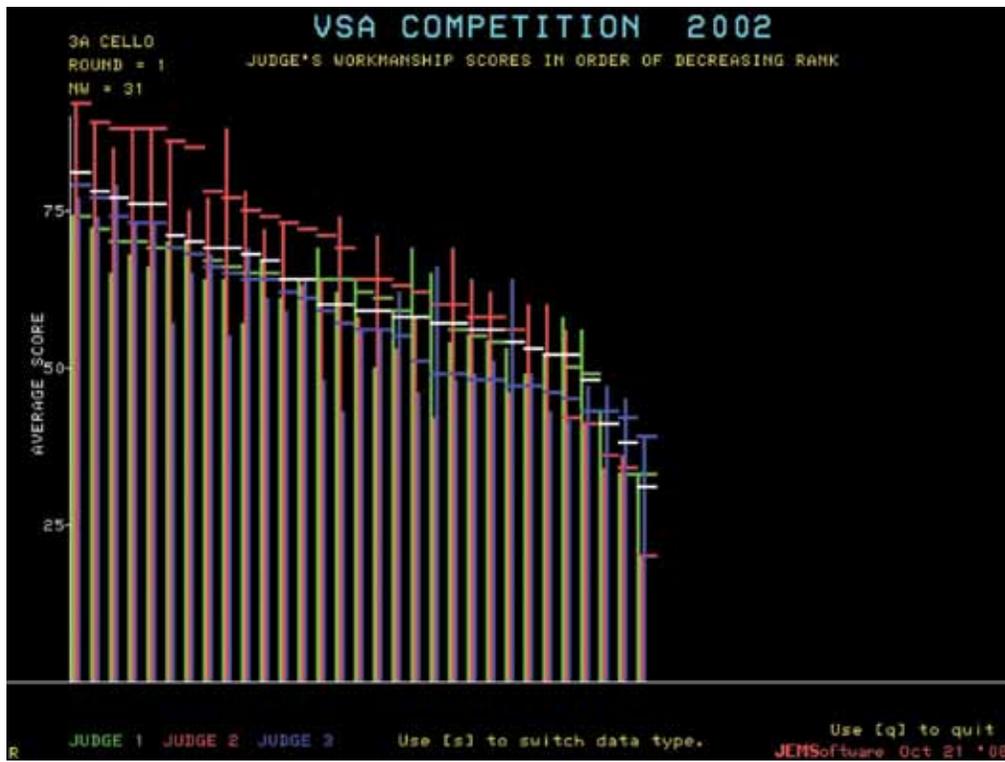


Figure 24. 2002 VSA Intl. Competition, Cellos, Round 1. Judges' Workmanship scores, plotted in order of decreasing rank, for 31 cellos. (The cascading white horizontal bars are the average scores of the three judges.)

Dr. King: That one. So this is a Round 1 here?

Dr. Regh: Yes, it is. There are 184 data points and it was Round 1 with the ultimate conclusion superimposed. The winning instruments were not known in Round 1. You only learned which ones won awards after the competition was over.

Dr. King: But we can say these were the results from Round 1, and what we see there over in the far left is this yellow line from roughly 79 up, and that was a gold medal violin.

Dr. Regh: That was a gold medal.

Dr. King: And that was identified in that way in Round 1.

Dr. Regh: It was identified correctly, but at the time we didn't know that.

Dr. King: We didn't know it was gold, but it received good marks.

Dr. Regh: Yes. There was no communication between the workmanship judges and the tone judges until Round 3.

Dr. King: Despite the chaos in Round 1, the judges were able to identify good workmanship and good tone. Then the next round serves to confirm a winner and so forth.

Dr. Regh: Absolutely, and let me add an important point. You have seen the difference in spread in the predictions, especially by the tone judges in Round 1 and in Round 2. The distribution really tightened up in Round 2, and they were still working independently and not communicating with each other. But it points to a potential problem in the way in which we conduct Round 1. In that huge hall with all those instruments laid out on the first day of competition, every instrument was played by each of the three tone judges at the same time. That had to be done very quickly because they had to meet the time requirements. The award winners must be identified in time for the VSA committee to process and produce the certificates and engrave the medals for presentation at the awards banquet.

There must be an influence of one judge playing near a second judge who is evaluating another instrument. Also, the cellos were in another corner, and the basses down in the other corner. The envi-

ronment in which we have conducted Round 1 has not been ideal, but because of the large number of instruments, we couldn't afford to transport each instrument into a separate room for the players. The large numbers, including about 200 violins, make that logistically impossible.

Steve Cheek: I thought Round 1 was just an elimination round and not a ranking round, but we see a ranking here. And that is from Round 1, right?

Dr. Regh: Yes. Every instrument gets evaluated identically in Rounds 1 and 2.

Mr. Cheek: Is there a score for each instrument in Round 1?

Dr. Regh: Yes. In fact, for those instruments that aren't advanced into Round 2, that single score in Round 1 will be the only one it gets and will be given to the makers in their final letter when the competition is over. Every contestant gets a print-out of every score, both for tone and workmanship, for all the rounds, plus any possible awards.

Mr. Cheek: Did you tell them where the boundaries were?

Dr. Regh: We tell them where their percentile rank was, so they have an approximate idea of where they were. To feed back exact numbers is meaningless due to the uncertainties and the statistics.

Mr. Cheek: Do you tell the contestants if their instruments made it into Round 2?

Dr. Regh: Yes, they get data from Round 2. They get a printout that separately lists all their scores in Round 1, all their scores in Round 2, and all their winnings in Round 3.

Mr. Chapman: Originally, we started out with workmanship and tone as two distinct categories that were not combined. So you could get a gold medal or a certificate in workmanship or in tone. In the mid-1980s the two were combined when we set up the system that Joe described, where if the instrument got a certificate in one and a gold medal rating in the other, then it was awarded a gold medal and so on. The thing that still bothers me, that I would like to see changed in the competition scoring, is when instruments receive a silver medal in workmanship or tone. For example, if an instrument merits a gold

medal in workmanship and nothing in tone, then there should be an automatic reconsideration by the other judges to make sure that they hadn't overlooked something or made a mistake. If the judges then say no, that would be the end of it.

Dr. Regh: That's a very good point, Eric. I'll have to go into the guts of the program to answer the question. The software is very sophisticated and it will allow you to look at discrepancies between any sets of data. And you can set a limit. For example, you can set the limit to 2, which will flag any instruments where the difference between judges is 2 or greater. And if you look at the data, let's say in Round 3, you can ask the program what happened to these instruments? What were the individual judgments of the judges? If they are close enough, I don't see any reason for reconsidering because of the way we generate the data. If there is a discrepancy, then we do talk to the judges and get an opinion as to whether that was legitimate or not. There's one illustration I showed where one judge was very conservative and gave scores with a mean value of about three, as opposed to another judge having a mean of about six. If two of the judges gave an instrument an eight and one gave it a five, that is not necessarily a problem. That just means the five is in a different distribution. So we do look for things like that and we look again at the data once the winners have been identified.

Mr. Chapman: The thing that still worries me is that the judges are in such a difficult situation, in this case looking at 200 violins and 100 violas. I think we should eliminate as many by-chance issues as we possibly can. The object is to come up with the right answer and the right instrument, rather than letting the data totally dictate the results.

Dr. Regh: The data do not dictate the results. I said that in Round 3 we take no data. The only thing we hope for is that Round 2 flushes up all of those instruments that are of a very high level. Then we have the judges play these instruments and do whatever they want to do to identify those instruments that, without any doubt, qualify for a gold medal. That is not based on data. That is based on exercising their expertise and competence on a limited set of instruments to identify the ones that are really outstanding. And we require that their decision be unanimous.

Mr. Chapman: Yes, but that still doesn't join the

workmanship and the tone evaluations. If you take the Round 3 tone instruments, they may not reflect the silver or gold medal in workmanship at all.

Dr. Regh: No, it does not. And the system that we designed is not supposed to.

Mr. Chapman: But that's my point. We need that safety valve to look at that balance, because it's not too often that there is a really gorgeous instrument that doesn't sound at all, or generally vice versa. We owe it to the competitors who have invested a great deal in the instruments, the travel, and everything else, to have that extra look just to make sure.

Dr. Regh: I agree that we need to do this, and I can also assure you that that is being done. From what I showed you after the fact, data from three competitions on all these instruments, I am confident that we are identifying the correct instruments. The only process flaw that I have been able to detect is the working environment for the tone judges in Round 1. Given that shortcoming, I think it's remarkable that we have had such a good identification of the instruments. Is a mistake possible? I am sure that we can come up with some scenario. The worst thing that we ever came into is a name being missed or someone that was supposed to win a gold medal and the name got dropped in some way. We did find the omission and made the correction, so the end result was still okay. That was a very rare occasion.

Mr. Chapman: I have two more suggestions. One is that the tonal judging be done in a concert hall. This goes back to my own prejudice from the years I was running the competitions, starting with 1978 at the University of California at La Jolla, 1980 at Hofstra University, and 1982 in Salt Lake City. All of the judging in the final round always took place in a concert hall. At the University of California at San Diego we used their main concert hall (with capacity of 1500 to 1800), which was right across from where the competition instruments were kept. At Hofstra University we had to carry the instruments across campus to their concert hall. In Salt Lake City we had a concert hall in the Hotel Utah itself, which seated perhaps 150 or 200 people. It concerns me still that in determining what a really first-class concert instrument is that deserves a gold medal in tone, that we don't do that in a concert hall. That is necessary even if we have to haul them across town somewhere to a facility

where it is possible to discriminate between instruments' tonal quality, projection, and articulation, which is difficult to do in small spaces. You can do your best guess in a small space, but I don't know very many professional musicians who don't get fooled occasionally by what they hear close up or under their ear or whatever.

Dr. Regh: The closest we've come to doing something like that is the session that Marilyn Wallin organizes on Saturday morning, where many of the winning instruments are played in a performance environment. So while you don't get to judge them, you get to hear them after the fact.

Mr. Chapman: But that's the problem. It's after the fact. Starting with the 1984 VSA competition in Ottawa, the instruments were closed off for the judging, which is understandable because there was just so much going on. I still believe that for the final round that playing—just as an educational opportunity if nothing else—should be open to the participants. They wouldn't have to be anywhere near enough to hear discussions among the judges. I think a lot of people come to the competitions because they want to hear the instruments, and hearing is believing. For them it's absolutely vital that they get that kind of exposure. What do these 10 or 15 instruments really sound like, and do I agree or not agree with the judges' assessments? Very often they don't. It's just like trying instruments in the shop. You can't get everybody to agree. But I think it's important for the participants to have that opportunity of listening to them and knowing their sound. The listeners wouldn't know who made the instrument, but they would hear something that they came a long way to hear.

Mr. Cheek: If the gold medal instruments jump out at you in the first round, do the silver medal winners also jump out at you after the first round?

Dr. Regh: No. Neither the gold nor the silver medal-winning instruments are obvious in Round 1.

Mr. Cheek: Well, you've been saying that from the data we have from Round 1, you can already predict the gold medal.

Dr. Regh: I can show you the statistics that would suggest that any one of these judges, if left alone in the room with the instruments, would come out with the correct winning instruments, sort of.

Any one of the judges showing competence can

come out with maybe 10 instruments out of 200 and say, these are the winners. So why do we need three judges and why do we need three rounds? Part of it is to get the results with great certainty. And part of it is to let the participating public and the contestants understand the integrity of the process that we are trying to run.

The VSA competition is considered to be one of the top events of its kind in the world. This is acknowledged by many of the other competitions and the competitors that I have talked to. The reason for that is precisely what we do. We have accountability. It is not likely that an individual judge could influence in any major way the results of the competition. We would see it. We have the data. We can go to the computer and figure out exactly what happens. It is good for you as competitors to know the kind of care we take to make that happen. When you finally do win a gold or silver medal or any kind of recognition in the VSA competition, it really means something. All around the world there are violinmakers who have VSA medals and certificates in their shops. They're proud to have received them. You can see from the large number of participants in every one of our competitions that it means something to participate and win here.

Mr. Cheek: With the information that you have now, namely that you can predict a gold medal from the first round, if indeed you allowed the two groups of judges to share scores, can you imagine if you have 500 violins to evaluate, you might not go past the first round because your mathematical models have shown that you don't need to go past the first round?

Dr. Regh: First of all, by looking at the data in Round 1, we cannot predict the gold medal-winning instruments because the data that were generated up here are going to be down there. But we don't know which are the winners yet. This is a superposition of the final results on our first data set.

Mr. Cheek: But that's only because you haven't compared the two sets of data from the tone and workmanship judges. Isn't that the case?

Dr. Regh: No. The winners aren't identified until Round 3. The data are not available. You only have two disconnected numbers here and disconnected numbers here. You have nothing else.

Mr. Cheek: Why are they disconnected? Because you choose not to share those numbers?

Dr. Regh: Because they're different ladders and we want them to be independent. That is part of the unique design of this competition.

Mr. Cheek: My point is that you could justify mathematically making an award based on the first round.

Dr. Regh: I don't want to do that.

Mr. Cheek: Okay. But your not wanting to do that is not a matter of statistical inadequacy or some kind of mathematical inadequacy. It's a public relations problem.

Dr. Regh: I will emphatically deny and refute that, because the intention of the competition, the overriding intention, is an education process. If you don't go to the trouble of looking at the instruments and collecting the data and feeding them back to the contestant, you do not have an education process and the entire thing is worthless. The only thing you get is something you can hang in your shop and people will like it or not like it.

Mr. Cheek: All I'm saying is that you have shown that you have some justification for identifying the winner, at least the gold medal winner, I don't know about the others, in the first round if you were to change your procedure to allow the two groups to talk. That you don't want to do that, I can understand.

Dr. Regh: No, you don't understand. The tonal judges only get briefly to look at the instrument for their tonal characteristics.

Walter Derlacki: I think the point of confusion is that these charts are not known until the end of the process, they do not get created at the end of Round 1.

Dr. Regh: The only influence it would have is on the process, not on the number of winners that come out in the end.

Mr. Derlacki: I think that may be part of the confusion that still exists. I must congratulate you on the effort that has gone into the statistical analysis because it helps resolve the question of whether

there are any influences that aren't correct. I'm not a maker; I only come to these events because I enjoy modern instruments. Great show!

Chris Germain: I've spoken with a number of the judges over the past competitions and I also hear consistency among them. It seems like the workmanship judges always say that even during the first day it's relatively easy for them to pick out the contenders from the non-contenders. The tonal judges, conversely, always seem at the end of the first day a little bleary and cross-eyed and unsure of exactly what took place.

I have two suggestions that I think would also have some statistical consistency, as well as result in a more favorable result that a lot of the makers would feel more comfortable with. First is if somehow we could relieve the tonal judges of this great burden of having to play 211 instruments in one day and feel like they've really picked out the winners. One way to do this would be on the first day to have the workmanship judges quickly determine which instruments exhibit a professional level of workmanship. After that, you would bring the tonal judges in and then you're assured that every instrument they're trying has also achieved this workmanship threshold. That's one way that everybody would feel that there'd be less chance of an inconsistent result.

My second suggestion is prompted by the fact that we don't really know too much about the likes and dislikes of the tonal judges going into the competition. On the Saturday after the end of the competition, we do hear the instruments being played and learn what each judge likes. But it would be very helpful in advance if the competitors knew, for instance, what instruments the tonal judges like, what kind of sound they like, and what kind of strings they play on. These things, since everybody has likes and dislikes, would help the competitors to prepare a little bit better in the tonal aspect.

Dr. Regh: By doing the screening beforehand, I think we would eliminate a very valuable input, and that is the evaluations of the judges in every category for each instrument, even of those makers who have not achieved a level that would allow their instruments to go beyond Round 1. I would not want to give up the opportunity to educate a maker to find out what he or she needs to do to make a better instrument next time. While I understand your comments, you don't understand the

tremendous workload these judges are subjected to, especially on the first day. They are professional musicians of the highest stature. They have to play and evaluate 211 instruments by the end of the day and come up with judgments that are representative of the quality of these instruments for the entire population. It is a tremendous task. We are at the limit of how much we can ask of the judges. If we have many more instruments, I think we'd have difficulty getting objective evaluations from those judges.

A suggestion had been made to have six judges instead of three for each category, which would halve the workload. It is statistically very difficult to stitch two sets of distributions together in a meaningful way to have a continuum. Every judge has a different taste. If you take three judges on one end with one set of tastes, and three judges on the other side with another set of tastes, and you join them together, they're not going to fit like this, they're going to fit like that. Also, what happens to the instrument when the workmanship judges recommend a certificate and the tone judges don't even select it for Round 2? The difference, in my opinion, is not the quality of the instruments, it is in the approach of the judges to make an evaluation of the instruments.

I have struggled with that dilemma for a while and I have not found a satisfactory way to divide the workload among more sets of judges. It is a difficult problem. Before I ran through the statistical evaluations of all our data, I was worried about how to go to the next step. However, the consistency of these data has shown me that the process, with the exception of the playing conditions in Round 1, is one that accurately predicts the winners and good instruments. Is it possible that one could make a mistake that would influence the outcome for one maker? I assume it could be. But that effect due to the competition is much less than what the maker or the environment would contribute. Changes in humidity, movement of the soundpost, or whatever, would have a greater influence on the evaluated ranking of an instrument than anything that we do in the competition process.

Mr. Chapman: We did, by the way, debate that issue of eliminating instruments after the workmanship evaluations and not passing them on to the tone judges. That was discussed years ago, but with the advice of people like Charles Beare, it was shot down. One of the reasons that idea was eliminated was because we could think of violins by superb makers like Giuseppe Guarneri *del Gesù* and a few

others that might not make it out of the first round, either. And that was the end of the discussion.

Dr. Regh: The VSA runs almost graduate-level courses in various aspects of violin and bow making at Oberlin College, and some of the directors of those courses are here. In that environment, there is a complete focus on all aspects of an instrument, and bows for that matter, and people are focused on how to get an instrument to sound better. There's a lot of testing going on. Many of the makers who have attended the Oberlin workshops have done very well in the VSA competitions. There is a correlation. I think the data clearly indicate that there is a known way of making very high-quality instruments. Ask those makers who are winning now and more than once in past competitions. They can do that consistently. There must be something that they know that should be important to every maker. Look how many instruments were entered in this year's competition. If their makers would have paid more attention to the tone, every one of them might receive a gold medal. And we would be delighted to make them and distribute them. But the fault is not on the part of the competition process. The improvements must come from the maker being more concerned about all aspects of the product that he or she makes. It is not a question of getting better judges.

Thomas Chinuchen: I'm a performer and a violin teacher, I make instruments, and I also have a collection of a few Italian antique instruments. I frequently search for good instruments, both for myself and for my college students. Often, when I take my students to a violin shop, they tend to like the commercial instruments. Under the ear they have a loud sound, which feels good to them, but when I stand some distance away from them, I don't hear much sound. Also, when they try them in a string quartet, the sound is small without projection. So that could be a big problem for the judges because the performer's technique always will be influenced by the instrument. If the judges are each in separate rooms when they evaluate the tone of the violins, they could hear only the sound they play, they cannot hear the sound from others playing.

Dr. Regh: In the final Round 3 the judges play for each other. Two judges sit apart, listening while one plays. And then they rotate.

Mr. Chinuchen: My point is that after 20 minutes, as performers our ears get tired and we don't pay attention to the detailed performance.

Dr. Regh: Yes. I will close by telling you a story that relates to this. Last year we were privileged to have two Stradivari violins and one by Giuseppe Guarneri *del Gesù*, known as the *Plowden*, to play and to analyze and to examine for physical and acoustical properties that we could learn from. Somewhere at the end of the session, we had a professional concert violinist play the instruments in a very comfortable living room environment. They sounded gorgeous. After about half an hour playing on these famous Cremonese instruments, even with your eyes closed, you could tell exactly which instrument was being played. They are so different, so unique, that your ears learned how to identify them.

Then we brought in two new violins made by two well-known American violinmakers, one made that year and the second made the year before. One of them was indistinguishable from the *Plowden*, even though it was only one year old. It had an incredible tone. I thought it was like having a champagne bubble bath—smooth, beautiful, and soothing. For me the output of the old and new violins was close enough to be identical. But then we asked the player how he felt about them, he said, “When I play the *Plowden*, there is a large window of opportunity—it's not very picky. When I play the other instrument, it has a smaller window. I have to focus harder and do some things slightly different, put a different kind of energy into it, but the output is the same.”

Mr. Chapman: Let's thank Dr. Regh for a very interesting and informative presentation.