

# Blind Listening Evaluation of Classical Guitar Sound Ports

*R.M. Mottola*

Copyright (c) 2008 by R.M. Mottola

[This article was originally published in *American Lutherie #96*, Winter 2008. It is republished here at <http://LiutaioMottola.com>.]

**Abstract** - A blind listening evaluation of classical guitar soundports was performed to ascertain whether players could hear the difference between an instrument with an open port and the same instrument with a closed port, all in a context that is representative of a player's instrument selection decision. The study was conducted to comply with the specifications detailed in standard ASTM E 2139-05, Standard Test Method for Same-Different Test from ASTM International (formerly called the American Society for Testing and Materials). Twenty four guitar player subjects were recruited to play a single guitar twice while blindfolded. In each of these two trials the open/closed port configuration of the guitar was selected at random. After playing the instrument in each port configuration the subject was asked to indicate whether there was or was not a difference heard between the two configurations. Results of this experiment indicate that the port open and port closed states are not perceivably different.

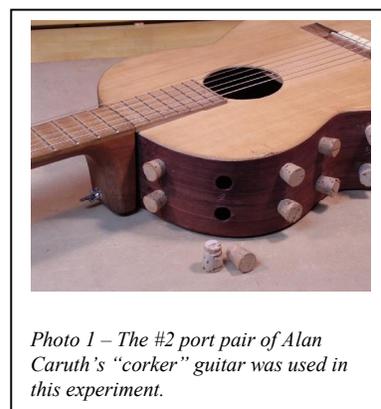
*American Lutherie #91* featured a survey article by Cyndy Burton on the use of sound ports in guitars and other stringed instruments. (1) These ports are small holes in the ribs of the instrument, intended to direct more of the sound of the instrument towards the player and thus serve as monitors. Research performed by Alan Caruth and published in *American Lutherie #94* showed that sound does indeed emanate from an open port. During the first half of 2008 I conducted research to determine if players could actually hear a difference in an instrument equipped with ports which could be opened or closed. This article describes that research.

**Introduction.** The question this research intended to address was whether players could hear the difference between an instrument with an open port and the same instrument with a closed port, all in a context that is representative of a player's instrument selection decision. Although at first blush it may seem overkill, a controlled, blind, multi-person study is necessary to address this question. Human sensory evaluation is quite a mature area of study, and years of research have demonstrated a number of things that tend to confound less formal attempts at evaluation. Humans vary greatly in sensory sensitivity, and even the sensitivity of an individual can vary greatly as a result of exposure, training and bias. The latter is a major issue for research, as substantial bias can result from the context in which a sensory assessment is made, and can also result from simultaneous input from another sense. We understand some of this intuitively – it is difficult to differentiate small volume changes in a quiet musical passage immediately after hearing loud volume material for a few minutes. In the realm of taste (which is not unrelated – general principles of human response to sensory input pretty much apply to all senses), chefs know that food is perceived to taste better if it is also visually appealing. And sensory assessment is highly subject to expectations, something which must be controlled for in any experiment intending to produce meaningful results. I am just touching on some of the issues here. Anyone interested in a more complete background in sensory evaluation may want to take a look at a textbook on the subject. The introductory chapters of the book *Sensory Evaluation Techniques* by Meilgaard et al (2) provide a good succinct overview.

The work by Alan Caruth mentioned has shown that sound emanates from the sound port(s) of a guitar so equipped. And a simple test can be performed which indicates that an open sound port has an audible effect on the sound of an instrument, although in a context which is not particularly representative of actual instrument selection or performance. If an open string is struck and let ring, and then a hand is used to alternately cover and uncover the sound port at a rate of about 1 Hz, pretty much everyone can hear the resulting wah wah style pitch filtering. But human perception of sonic difference is highly influenced by both the way in which sound samples are presented and by the amount of time separating them. This informal test is optimized for detection of difference, but as mentioned is not representative of the kind of comparison that an individual interested in comparing two different instruments would perform.

Another shortcoming of the informal test described above is that it is not blind, a shortcoming shared by most informal evaluations. An interesting non-lutherie experiment was performed recently by Antonio Rangel and others on the perceived taste of wine, and published in Proceedings of the National Academy of Sciences. (3) In this study, subjects were asked to sample wines and then to report on how good they were. They were told nothing about the wine other than its cost, and the cost ranged from cheap to very expensive. The general trend of the results showed that the subjects found that the more expensive the wine the better is tasted. But the experimenters used exactly the same wine for all trials and the cost figures they gave were all made up. These results (and the results contained in a recent book on the same subject, *The Wine Trials* by Robin Goldstein (4)) are consistent with many previous experiments in the area of human sensory perception – people’s perception of sensory input is highly influenced by expectations as well as by ancillary sensory input.

Experimental Design. The sound port listening evaluation was designed to eliminate expectations and ancillary sensory input as influential factors in the experiment and to emulate the type of comparison a player would perform in selecting an instrument. A single test instrument (Alan Caruth’s “corker” classical guitar, pictured in **photo 1**) with ports that could be opened or closed was used. A single port pair located approximately 4” from the neck heel on the side of the instrument was opened or closed for this experiment. This port pair was used because it showed substantial volume in Caruth’s experiment and also “points at” and is located close to the player’s face when the instrument is held in typical playing position. Port placement and area are probably critical factors in this and similar investigations, as frequency spectrum, volume and proximity to the player’s ears will all vary with position and/or size. Assessors were evaluated one at a time in a quiet room following an explanation of the experiment and instructions. Assessors performed their evaluation blindfolded so they could not see which port configuration was in use.



*Photo 1 – The #2 port pair of Alan Caruth’s “corker” guitar was used in this experiment.*

As mentioned, human sensory evaluation is a mature area of study, and as such the experimental methodology for this study did not have to be designed completely from scratch. The study was conducted to comply with the specifications detailed in standard ASTM E 2139-05, *Standard Test Method for Same-Different Test* from ASTM International (formerly called the American Society for Testing and Materials). The standard includes specifications for the design and implementation of the experiment as well as the methods used to analyze the results.

Each assessor played the instrument twice, the first time with one port configuration (the port pair opened or closed) and the second time with another port configuration (the port pair opened or closed). All port permutations (open/open; closed/open; open/closed; closed/closed) were used in the experiment in equal number but each assessor evaluated just one permutation. During instruction the assessors were made aware that there was a 50% chance that the port configuration of the instrument would be identical both times they played it. The assessor played the instrument for 30 seconds in the first port configuration, then the instrument was taken from the assessor and configured for the second port configuration and then returned to the assessor 10 seconds later. The assessor then played the instrument again and was asked to state whether the instrument sounded the same or different between the two trials. Since removing or replacing the corks in the ports makes some squeaking noise, the corks were made to squeak even in those cases where there was no change in port configuration between trials. To minimize sensory fatigue and for statistical accuracy each assessor participated in only one two-trial session. A preliminary experiment indicated that potential assessors could not determine port configuration by detecting weight differences of the instrument with the corks in or out. Assessors were free to choose the material played in the evaluation. Although the choice of material could affect one’s ability to differentiate between an open and a closed port, preparatory investigation indicated that as a practical matter assessors could not be expected to both play a recently learned piece and pay careful attention to listening to the instrument at the same time. Some of the assessors were luthiers, and as an anecdotal observation, the luthiers tended to play in a more probing fashion, working at the extremes of volume, pitch and timbre.

One critical area not specified in the ASTM test standard was the manner in which the samples are presented to the assessor. Simultaneous presentation is preferred but that is not possible for a player playing an instrument. Sequential presentation does present some potential problems in an evaluation of this type. There are certainly well established presentation methods which provide for better detection of sonic differences, such as the ABX method (described succinctly and accessibly at <http://www.hydrogenaudio.org/forums/index.php?showtopic=16295>), but none of these were deemed practical for evaluation of an instrument by a player. The method chosen also emulates

well a typical real world evaluation of two instruments by a player, and the corresponding level of sensitivity was considered to be appropriate for this type of evaluation.

**Data Collection.** The test was run on 24 assessors at 4 sites. The number of assessors was determined by practical considerations, primarily the limited resources available to conduct this experiment. This number was suboptimal in terms of the sensitivity of the test, as will be described in detail in a later section. Assessors were all competent professional musicians or conservatory students and/or luthiers. Some assessors were known to be primarily classical guitar players as described below. Six (6) assessors were recruited from among professional players picking up repair instruments at the repair shop of a busy music store (cm). The test was conducted at this site in a small, quiet lesson room. Eleven (11) assessors were recruited from among player/luthiers at a monthly meeting of a regional luthiers group (nel). The test was conducted at this site in a quiet, moderately sized wood shop. Four (4) assessors were recruited from among professional classical guitar players and/or conservatory students attending a classical guitar master class workshop (nec). The test was conducted at this site in a quiet, small recital hall. Three (3) assessors were recruited among professional classical guitar players and/or conservatory students (lr). The test was run at this site in a quiet suburban living room. Assessor responses were recorded by the author in all cases, and are listed in **table 1**.

Subject #	Trial 1 port configuration	Trial 2 port configuration	Subject's Response	Location	Classical player
1	open	open	different	cm	?
2	closed	open	same	cm	?
3	open	closed	different	cm	?
4	closed	closed	different	cm	?
5	open	open	different	cm	?
6	closed	open	different	cm	?
7	open	closed	different	nel	?
8	closed	closed	same	nel	?
9	open	open	different	nel	?
10	closed	open	different	nel	?
11	open	closed	different	nel	?
12	closed	closed	different	nel	?
13	open	open	different	nel	?
14	closed	open	different	nel	?
15	open	closed	different	nel	?
16	closed	closed	different	nel	?
17	open	open	different	nel	?
18	closed	open	different	nec	yes
19	open	closed	different	nec	yes
20	closed	closed	same	nec	yes
21	open	open	same	nec	yes
22	closed	open	same	lr	yes
23	open	closed	different	lr	yes
24	closed	closed	different	lr	yes

Table 1 – Port configurations and assessor responses for all 24 assessors.

**Analysis and Interpretation of results.** The data from the test are summarized in **table 2**. Following instructions for analysis specified in the ASTM test standard, the initial analysis is to determine whether the number of *different* responses from those assessors receiving different port configurations is less than or equal to the number of *different* responses from those that received the same port configuration in both trials. In this case the hypothesis of no assessed difference can not be rejected, which is to say that, given the limitations of the experiment, assessors could not tell the difference between open and closed ports. Although close, the results of this test do not meet these criteria, as the number of *different* responses from those assessors receiving different port configurations (10) is

greater than the number of *different* responses from those that received the same port configuration in both trials (9). In this case the test standard specifies that statistical analysis of the data is required to determine whether the samples are perceptibly different or not. The analysis specified is called Fisher's Exact Test, a statistical test of the equality of two independent binomial proportions. This test is described in detail in the textbook *Principals and Procedures of Statistics* by R.G.D. Steel and J.H. Torre and also in other statistics texts. Simply put, the goal of the statistical analysis for this experiment is to determine the likelihood that the test results indicate the ability of assessors to detect the difference between an open and a closed port. Just looking at the summary data in **table 2** indicates that this is not likely, and in fact the analysis yielded a resulting p-value of 0.5, which indicates that the port open and port closed states are not perceptibly different.

Discussion. The results indicate that the population of assessors could not perceive a difference whether or not a port was open in this experiment. In any experiment there are always factors which would limit the extent to which it is appropriate to extrapolate from the results to real world application. Two I would like to discuss here are the composition of the assessor population and the sensitivity of the test as a consequence of the limited number of assessors.

Response:	Assessor Received:		TOTAL
	Matched Pair	Unmatched Pair	
Same	3	2	5
Different	9	10	19
TOTAL	12	12	24

Table 2 – Summary data for the test is represented in a 2x2 table.

That the assessor population is not homogeneous by a number of criteria is a given. One criterion that could affect the applicability of these results is the type of instrument usually played by each assessor. I did not track this reliably during data collection, but it is reasonable to assume that those assessors whose primary instrument is not classical guitar typically play steel string acoustic guitar or electric guitar. It is possible that close familiarity with the type of instrument used in this experiment (classical guitar) could influence sensitivity to subtle differences in tone or volume from that type of instrument. It is anecdotally interesting to note (but not statistically significant) that the known classical guitar players used as assessors later in the experiment scored better at determining a port difference than did the rest of the assessors. Again, the numbers are too small to consider in the context of this study, but this is something that subsequent similar studies may want to consider.

As mentioned earlier, the number of assessors used in this study was sub optimal. The ASTM test standard provides data in tabular form to help determine the number of assessors needed for desired

$\alpha$ -risk	$\beta$ -risk	p1	$\Delta$	Number of assessors
0.05	0.2	0.3	0.3	84
0.3	0.2	0.3	0.4	24

Table 3 – The number of assessors required is a function of the desired sensitivity of the test. Nominal values for sensitivity parameters shown in the first row indicate a test using 84 assessors is required. The test outlined in this article used only 24 assessors, resulting in the sensitivity parameter values shown in the second row. Descriptions of these parameters appear in the text.

test sensitivity, given some information about the likely profile of the data. With no previous assessment studies of this type I had no data profile information available. Using the default values for that and for the parameters specifying the desired test sensitivity, a minimum of 84 assessors would have been required to obtain results with reasonable confidence in their accuracy. As studies of this type are often limited by practical considerations in the number of assessors (as this study certainly was), the tables in the ASTM test standard can be read to indicate test sensitivity as a function of the number of actual assessors, in this case 24. **Table 3** shows both the nominal test sensitivity values and associated number of assessors required (84), and the test sensitivity values I selected based on the number of assessors used in this experiment. The sensitivity parameters include:

$\alpha$  (alpha) risk – The probability of concluding that a perceptible difference exists when in reality one does not. This is also known as Type I error or significance level;

$\beta$  (beta) risk – The probability of concluding that no perceptible difference exists when in reality one does. This is also known as Type II error;

$\Delta$  (delta) – The minimum difference in proportions that the test should detect;

$p_1$  – The proportion of assessors in the population that would respond *different* to the matched (i.e., not different) sample pair;

For the case of the sensitivity parameter values resulting from the use of 24 assessors in this experiment, the following can be stated about the sensitivity of this test:

There is a 30% probability that the test will indicate perceivable differences where there are none ( $\alpha = 0.3$ ); The test provides an 80% certainty ( $\beta = 0.2$ ) of detecting a 40% difference ( $\Delta = 0.4$ ) between the proportion of assessors who would correctly identify the unmatched port configuration and the proportion who would say the matched port configuration was different in the population. The logic may be a bit hard to follow, but what this means in simple terms is that the probability of the test being acceptably accurate with this number of assessors is good but not great.

Of possible interest to future research in this area is the proportion of *different* to *same* responses, which is heavily biased toward *different*. Various acoustic, psychoacoustic and psychosocial possibilities could explain this bias and this could itself be an interesting area of research.

**Conclusions.** The results of this experiment indicate that the port open and port closed states are not perceptibly different. These results may be of practical value for those considering adding sound ports to guitars. Future efforts may also want to attempt to identify narrower populations that may be more able to detect the presence of a sound port. If this can be done, such ports may prove advantageous to that population. Additional research using larger assessor populations and a more consistent test environment may also provide better confidence in study results.

**Acknowledgements.** Research is both expensive and time consuming. I was very fortunate in this effort to be aided by a number of generous individuals and organizations. The guitar used in this experiment was provided by Alan Caruth. The New England Luthiers group (<http://newenglandluthiers.org>) provided a population of assessors (players) at one of their regularly scheduled monthly meetings. Dennis Keller and Jim and John Mouradian at Cambridge Music in Cambridge MA (<http://www.cambridgemusic.com>) provided space to perform some of the evaluations, and a steady stream of assessors as well. Robert Sullivan, Eliot Fisk and Steven Lin of the New England Conservatory (<http://www.newenglandconservatory.edu>) also provided space and access to assessors at Boston GuitarFest 2008, an annual classical guitar workshop and competition. This article is substantially better than its initial draft due to careful review and commentary by Mark French and David Cohen. Content of the commentary section of this article is due largely to discussions with Cyndy Burton and Jeffrey Elliot. The assessors were kind enough to take time to participate in this study.

## Bibliography

1. **Burton, Cyndy.** There's A Hole in the Bucket. *American Lutherie* #91. 2007, p. 6.
2. **Meilgaard, Morton.** *Sensory Evaluation Techniques*. s.l. : CRC Press, 2007.
3. *Marketing actions can modulate neural representations of experienced pleasantness.* **Rangel, Antonio.** s.l. : Proceedings of the National Academy of Sciences, 2008.
4. **Goldstein, Roblin.** *The Wine Trials*. s.l. : Fearless Critic Media.