Subjective Training and Evaluation Program (STEP)

Version 2.01

A computer-controlled system for audio presentation and subjective evaluation

Audio Research Labs
info@audioresearchlabs.com
June, 2016

Contents

Audio Research Labs

1   Minimum Requirements

The minimum system requirements to support this product are:
- Microsoft Windows 7 64-bit operating system
- Pentium 64-bit processor (i.e. i5 or i7)
- 4GB of memory (8 GB is recommended for 16 to 32 channel signals).
- 1 MB of hard disk space
- Stereo or multi-channel audio output device with ASIO drivers.

2   Quick Start

Assuming that your computer platform meets the minimum requirements, and that you have your sound card installed, get started by doing the following steps:

2.1   Install STEP

There is only one ARL STEP software program – the demo *is* the fully functional product, but it functions in demo mode when it finds no license file. When you purchase the software, you will receive your license file via email.

You should have already downloaded the self-extracting zip archive `Step64.exe`. Note as of this writing it is Version 2.00. We recommend that you unzip the files to the following folder (which is the default location), although it can be put in any directory.

    C:\Programs Files\ARL

Do this by double-clicking on `Step64.exe` and clicking on the `Unzip` button. The extraction process will create the folder

    C:\ Programs Files\ARL\Step64

that will contains the 64-bit STEP executable and STEP Version 2.00 user's manual

When you purchase the software, you will receive a license file (`license.txt`). Please copy that to the folder containing the Step64 executable. Note that a Version 2 license is required for Step Version 2 (i.e. Step64.exe).

2.2   Download STEP demo files

We strongly recommend that you download the STEP demonstration signal files since this manual will use examples from the demo files. The files are available in a zip archive from the STEP download page. Unpack the files at some location in your documents file structure, for example "C:\ARL"

It will create the folder STEP_Signal_Demo containing the following directory structure:

analysis	score files and data analysis spreadsheets
setup	channel identification signals (5.1 channels and 30 channels)
interleaved\demo1	signal and Session files for interleaved file format mono signals
interleaved\demo2	signal and Session files for interleaved file format stereo signals
interleaved\demo5	signal and Session files for interleaved file format 5.1 channel signals
mono\demo2	signal and Session files for mono file format stereo signals
mono\demo5	signal and Session files for mono file format 5.1 channel signals

2.3   Associate Session files with the STEP executable

STEP Session files control the presentation. The Session file extension is `*.asi`, and it is convenient (but not necessary) to associate this extension with the STEP program. To do this:
- double click on an example Session file (e.g. in the `demo2` directory).

- When you are presented with the "Windows cannot open this file" dialog box, select "Select program from a list" and click "OK."
- When you are presented with the "Open with" dialog box, select the "Always use the selected program to open…" box.
- Click on the "browse" button, browse to the folder into which you loaded the STEP executable (see Section 2.1), click on `step.exe` and finally click the "Open" button.

2.4 Run a test

The STEP session file can accommodate signals in two formats: interleaved WAV files or multiple mono WAV files. The STEP demonstration signal set has both formats, identified by directories `inter` (interleaved format) and `mono` (multiple mono file format). Under each of `inter` and `mono` are directories `demo1, demo2 and demo5`, that contain several example Session files and signal files.

Session files can make use of the following test methodologies:

Methodology	Description
`bs1116.asi`	BS.1116 test methodology [1].
`mushra.asi`	BS.1534 MUSHRA test methodology [2].
`ab5.asi`	P.800 Comparison Category Rating methodology with 5 rating grades [4].
`ab7.asi`	P.800 Comparison Category Rating methodology with 7 rating grades [4].
`Mos.asi`	P.800 Absolute Category Rating Mean Opinion Score [4].
`abx.asi`	ABX forced-choice methodology.
`training.asi`	Demonstrates the STEP training mode.
`playlist.asi`	Demonstrates the STEP playlist mode.

The interleaved file format signal files are named as:

`sigN_orig.wav`	Reference or unprocessed test stimulus *N*, where *N* is 1, 2 or 3.
`sigN_sysM.wav`	Stimulus *N* processed by system-under-test *M*, where *M* is 1, 2 or 3.

The mono file format signal files are named as:

`sigN_orig_<lab>.wav`	Reference or unprocessed test stimulus *N*, where *N* is 1, 2 or 3 and <lab> is a unique label that associated a mono signal file with an output loudspeaker (or ASIO output channel).
`sigN_sysM_<lab>.wav`	Stimulus N processed by system-under-test M, where M is 1, 2 or 3 and N and <lab> are as above.

Note that if you have only a demonstration version of STEP, the following behavior will be slightly different, as outlined in ANNEX 1.

Double-click on `mushra.asi` and the program will start. First it presents the *Welcome* panel;

Welcome panel:
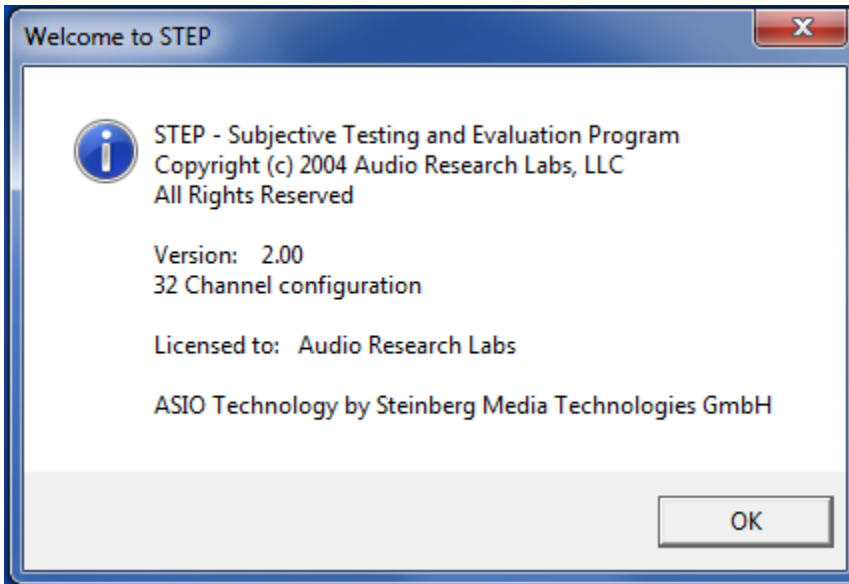
After a slight delay, it presents the *Administration* panel.

Administration panel:
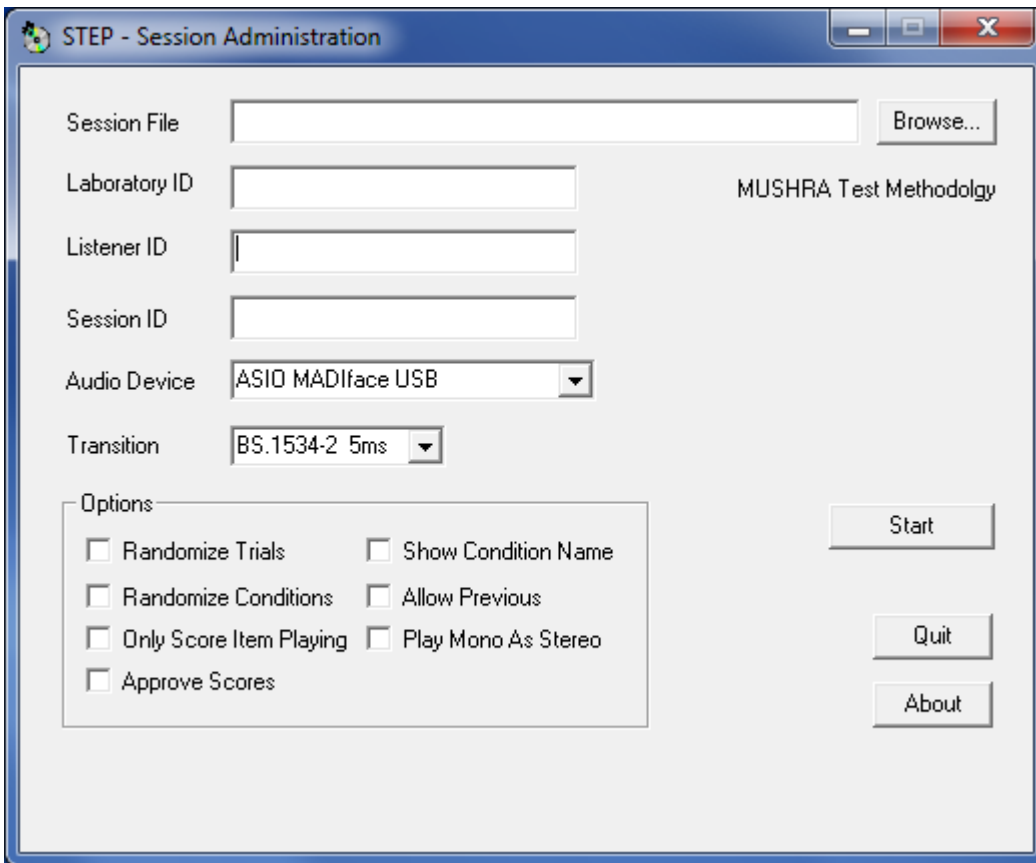
Enter a Session File, typically by using the Browse button. At this point the Session file is opened and parsed. In the case of long Session files of multi-channel signals, this may take several seconds, in which

Audio Research Labs

case a "progress bar" will appear at the bottom of the window. Next, enter values for Laboratory ID, Listener ID and Session ID. Select the desired Audio Device. You can leave Transition at the default value.

Click the "start" button. STEP will present the MUSHRA trial panel. As with the Administrator Panel, in the case of long Session files with multi-channel signals, reading the signals from disk into memory may take several seconds, in which case a "progress bar" will appear at the bottom of the window.

Trial panel:
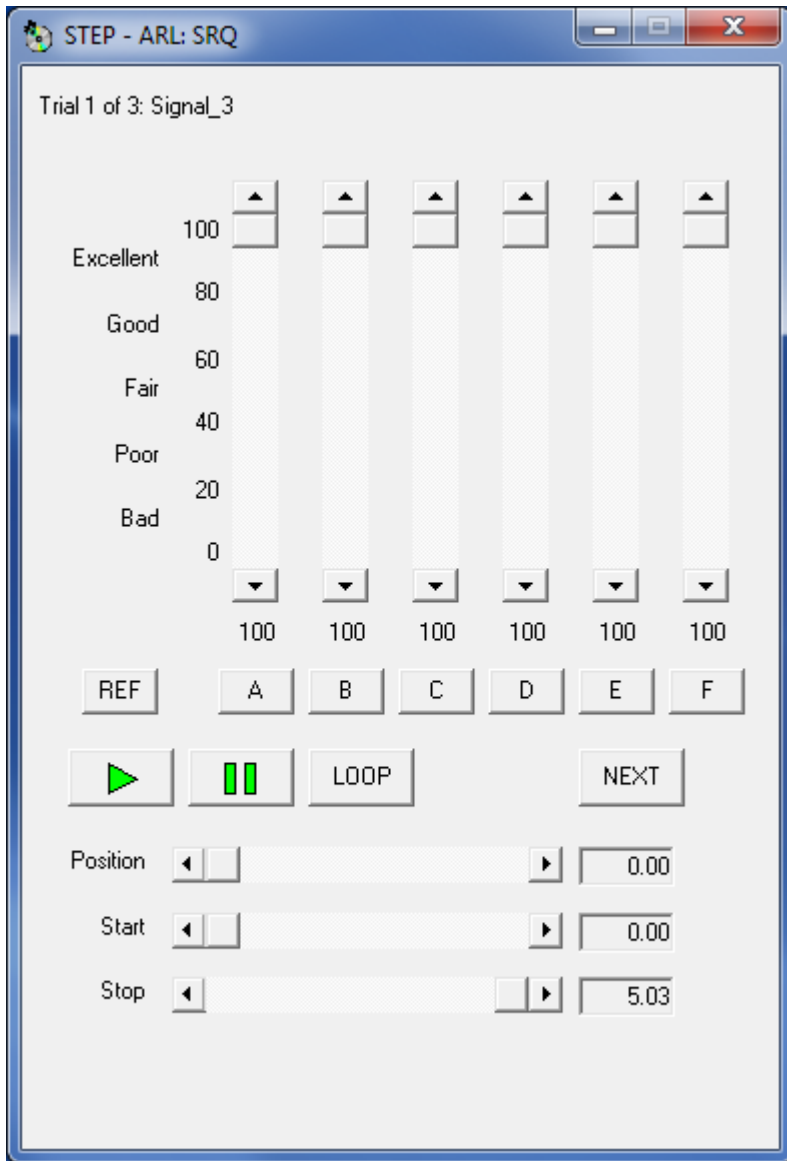
Click on the Ref button to play the reference signal, or on any of the A, B or C buttons to play the processed signals. Click on the "Loop" button to play the signal continuously. The horizontal Position slider indicates the instantaneous position in the signal waveform. Grabbing and moving the Start slider alters the start point for waveform looping, and similarly moving the Stop slider alters the end point, thus permitting a "loop and zoom" function that is particularly powerful for subjective evaluation.

Rate the processed signals by grabbing and moving the vertical sliders above the letter condition buttons. When you are satisfied with the ratings, click on the "Next" button to go on to the next trial. Note that if

Audio Research Labs

"Approve Scores" was selected, a confirmation window appears. When the last trial is scored, the MUSHA window is replaced by the Administrator window, and the STEP is ready for the next listener. Exit STEP by clicking on the Quit button in the Administrator window.

3 Operation

3.1 Typical system setup

3.1.1 Workstation

The STEP software is invaluable for use by a codec designer at his or her workstation. For example, it can be used in "Evaluation" mode to get a quick assessment of the quality impact of varying one or more coder parameters (each parameter set is a different System under Test). The workstation mode does not require any special configuration of equipment.

3.1.2 Listening Room

One typically uses STEP to conduct listening tests in an acoustically conditioned listening room or sound booth. In this case, Figure 1 shows a simple and effective configuration for use with a listening room. Here the PC external to the room runs STEP, and has a sound card that produces a multi-channel digital output (using e.g. the MADI format). This connects to MADI-compatible multi-channel D/A and associated audio amplifiers, which in turn connect to the speakers located inside the listening room. A Keyboard, Video, Mouse (KVM) switch permits the PC console to be either the Test Administrator Console (external to listening room) or the listener console, which would consist of only a presentation screen (e.g. LCD) and mouse, both located inside the listening room. After the test administrator sets up the test using the Administrator panel, KVM control is switched to the listener for the duration of the test session.
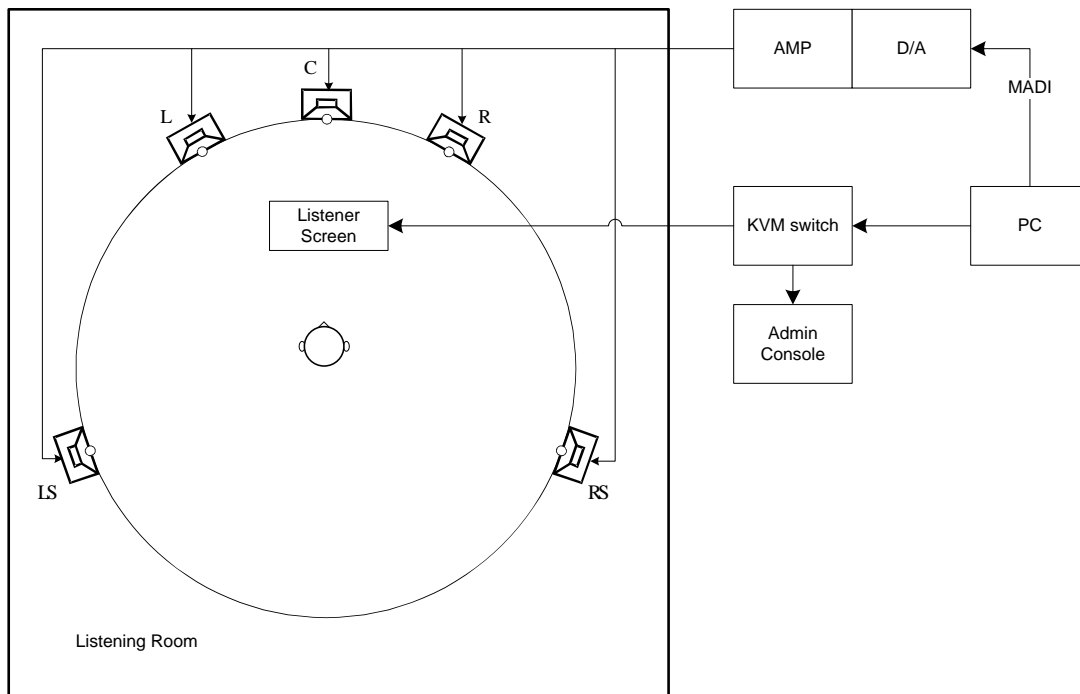

Figure 1. Suggested Configuration for Listening Room

3.2 Audio Output Device

STEP only supports audio output hardware that has an ASIO [5] driver. See www.audioresearchlabs.com/step for a list of audio interface hardware that is known to work with STEP.

3.3 Signal File Formats

3.3.1 Overview

STEP Version 2.0 supports both interleaved file signal formats and multiple mono-file signal formats, as detailed in this section. In addition, STEP Version 2.0 supports up to 32 channels of signal playout. Because this can entail a significant amount of data transfer, at the start of each Trial all stimuli are read from disk and buffered into memory. During the Trial, real-time playout only occurs from memory buffers, which minimizes the likelihood of a real-time play dropout.

STEP only supports reading WAV or Extensible WAV file formats. The WAV file contents can be any sampling frequency and can be signed integer values of 16, 20 or 24 bit word length.

3.3.2 Interleaved File Format

In this case, a stimulus is a single WAV file in which the channels are interleaved. More precisely, each sample of each audio channel is assembled in an ordered manner to form a sample frame, and the WAV file contents is a series of sample frames. The association the series of integer values in the sample frame with the loudspeaker signals is specified in the channel map file (see Section 3.4.2). While an interleaved channel map file may specify the association or mapping for 32 channels, a test can have fewer channels in the interleaved stimuli. Specifically, if the test stimuli have N interleaved channels, those must be the first N channels of the interleaved channel map file. STEP automatically feeds the remaining ASIO output channels with the zero signal.

3.3.3 Mono File Format

In this case, a stimulus is a collection of single-channel WAV files, where each file is a single audio channel. The association of audio signal channels to loudspeaker signals is specified in the mono channel label file (see Section 3.4.2). Unlike the interleaved file format, the mono file format can use any subset of ASIO output channels (i.e. loudspeaker signal feeds). For example, if the mono channel label file specifies a mapping for 32 channels, a stimuli may consist of an arbitrary and non-contiguous subset of those 32 channels. STEP automatically feeds the remaining ASIO output channels with the zero signal.

3.4 Session and Support Files

The format of the MUSHRA Session and Score files will be discussed in this section. The difference between MUSHA Session and Score files of that of the other test methodologies is discussed in Section 5.

3.4.1 Session File

By default, STEP Session files have extension `asi`, although any extension (or no extension) will work. The session file is a plain ASCII text file that can be created with any text editor. It can have "PC" line terminations (CR/LF) or "Mac/Linux" line terminations (LF).

A session file is composed of a "header" section followed by a "body" section.

3.4.1.1 Header

The simplest header section of a session file is just the line that specifies the test methodology:
`Session=MUSHRA`
STEP is not sensitive to case when parsing the Session header, so any mix of upper and lower case will work.

The possible session header lines are:
```
Session=BS.1116
Session=MUSHRA
Session=AB5
Session=AB7
Session=ABX
Session=MOS
Session=Train
```

The header section can specify a number of options, but STEP assumes defaults for all but the test methodology such that the header can be as short as one line (as shown above). Options are specified using "attribute/value" pairs, i.e. using a header line:
```
attribute=value
```

The following tables list optional header lines, their attribute text strings, the possible values text strings, and their default values. The attributes and values are show as literal strings. The exception is `interleaved_chn_map_file` and `mono_chn_label_file` that take a filename as value, and are discussed in detail in Section 3.4.2. The attribute `lock_admin_options_panel` has the value "true" then all of the check-box selections in the "options" panel in the Session Administration window are de-activated. Their values are the defaults for the test methodology, or are over-ridden by other lines in the session header.

This table indicates options that can only be set in the Session file header.

Attribute	Possible Values	Default
file_format	interleaved, mono	interleaved
interleaved_chn_map_file		Ascending order
mono_chn_label_file		No default
speaker_configuration	same, different	same
has_open_reference	true, false	true
lock_admin_options_panel	true, false	false

This table indicates options that can override the options specified in the Administration Panel Options pane. These options have default values that depend on the test methodology, as indicated in Section 3.5.

Attribute	Possible Values
randomize_trials	true, false
randomize_conditions	true, false
only_score_item_playing	true, false
play_mono_as_stereo	true, false

The options **Show Condition Name** and **Allow Previous** are only accessible from the Administration Panel Options pane.

3.4.1.2 Body

The body section of the Session file consists of a simple structure associated with each test trial, which is repeated for every trial in the test. The structure begins with a line having "#" as the first character followed a space or tab followed by the test item label. Subsequent lines in this structure have the pathname, relative to the directory containing the Session file, of each stimuli or condition (i.e. test item processed by system

9

under test) in the trial followed by space or tab followed by its associated system label (which is written to the Score file).

Depending on the test methodology and options (e.g. `has_open_reference`) the original signal is the first of these files (and will be associated with the REF button, and subsequent files are associated with condition buttons A, B, C, …. Note that STEP has a limit of 15 processed files plus the reference file, for a total of 16 files, although MUSHRA [2] recommends that there be no more than 12 processed files.

Here is an example structure associated with one trial in a MUSHRA test using interleaved signal files:

```
# Sig1
./test1/sig1_orig.wav    Ref
./test1/sig1_orig.wav    HR
./test1/sig1_sys1.wav    Sys1
./test1/sig1_sys2.wav    Sys2
./test1/sig1_sys3.wav    Sys3
./test1/sig1_lp70.wav    LP70
./test1/sig1_lp35.wav    LP35
```

The simplest MUSHA Session file (using interleaved signal files) is shown here:

```
session=MUSHRA
# Sig1
./test1/sig1_orig.wav    Ref
./test1/sig1_orig.wav    HR
./test1/sig1_sys1.wav    Sys1
./test1/sig1_sys2.wav    Sys2
./test1/sig1_sys3.wav    Sys3
./test1/sig1_lp70.wav    LP70
./test1/sig1_lp35.wav    LP35
# Sig2
./test1/sig2_orig.wav    Ref
./test1/sig2_orig.wav    HR
./test1/sig2_sys1.wav    Sys1
./test1/sig2_sys2.wav    Sys2
./test1/sig2_sys3.wav    Sys3
./test1/sig2_lp70.wav    LP70
./test1/sig2_lp35.wav    LP35
```

3.4.2 Channel Map Files

Interleaved Files

The optional interleaved channel map file for interleaved files is signaled with the following Session header line:

```
interleaved_chn_map_file=ITUR5-1.txt
```

In this case, the map file is for a 5.1 channel signal (ITU-R 5.1 channel speaker configuration) and the map file contains:

```
1       1
2       2
3       3
4       4
5       5
6       6
```

Each line in the file specifies two values: the first is the ASIO output channel number, the second is the channel number (or order) in the interleaved sample frame. Note that channels are counted from 1 (not 0). This is an optional line in the header, and if unspecified, the default mapping is just ascending order:

```
1       1
2       2
```

```
3       3
4       4
```
And so on.

In the case of interleaved files, the file pathname shown in the Session file is the full pathname.

Mono Files
The mandatory mono channel label file for multiple mono files is signaled with the following Session header line:

```
mono_chn_label_file=ITUR5-1.txt
```

In this case, the map file is also for a 5.1 channel signal (ITU-R 5.1 channel speaker configuration) and the file map file contains:

```
1       L
2       R
3       C
4       LFE
5       Ls
6       Rs
```

In the case of mono files, the file pathname shown in the Session file is the "base" pathname. The set of mono files that make up the signal are constructed by concatenating the base pathname, "_", the channel label and ".wav". For example, if one line in the body of the Session file is:

```
./demo5/sig1/sig1/sig1_orig    REF
```
then this one line indicates that STEP should read the following set of files:

```
./demo5/sig1/sig1/sig1_orig_L.wav
./demo5/sig1/sig1/sig1_orig_R.wav
./demo5/sig1/sig1/sig1_orig_C.wav
./demo5/sig1/sig1/sig1_orig_LFE.wav
./demo5/sig1/sig1/sig1_orig_Ls.wav
./demo5/sig1/sig1/sig1_orig_Rs.wav
```

Channel label files can be extended to as many channels as are needed to support a test (up to the STEP limit of 32). For example, a test with 22.2 channel signals that is run in a listening room with 30 loudspeakers might use the following channel label file:

```
1       A+000_E+00
2       A+030_E+00
3       A-030_E+00
4       A+060_E+00
5       A-060_E+00
6       A+090_E+00
7       A-090_E+00
8       A+110_E+00
9       A-110_E+00
10      A+135_E+00
11      A-135_E+00
12      A+180_E+00
13      A+000_E+35
14      A+045_E+35
15      A-045_E+35
16      A+030_E+35
17      A-030_E+35
18      A+090_E+35
19      A-090_E+35
20      A+110_E+35
21      A-110_E+35
```

11

```
22    A+135_E+35
23    A-135_E+35
24    A+180_E+35
25    A+000_E+90
26    A+000_E-15
27    A+045_E-15
28    A-045_E-15
29    LFE1
30    LFE2
```

IMPORTANT Note about "missing" files from mono file signal set

In the example above for 5.1 channel content, STEP read a mono file for every label in the channel label file. However, this is not required, and in the case of the 30-line channel label file above, this is typically not the case. STEP will attempt to open every file pathname constructed from the base filename and the appended channel label. The files that are found are reported in the Check file (check.txt, see the next section), and if not found then that ASIO output channel is fed with the zero signal.

3.4.3 Check File

The file `check.txt` is written as a result of parsing the Session file, and is written to the same directory as the Session file. The file reports

- the header options specified in the Session file
- the key for diagnostics for each stimulus file
- The format of each signal file and whether each signal file agrees in format with the other signal files. This is reported as one line per signal file with a blank line indicating a new trial in the session.

An example check file for signals in the interleaved file format is shown below. If everything is correct then it indicates the number of channels in the interleaved files and the status "OK."

```
Session Check

Session configuration
Session: MUSHRA
File format: interleaved
Speaker configuration: same
Open reference: yes
3 Trials
6 Conditions per Trial

Diagnostic key
Unexpected signal components: USC
Number of channels: NC
Sampling rate: SR
Audio word size: WS
Number of samples: NS
Channel Mask: CM


    Channels/Diagnostic   File
         6-channel OK : sig1_orig.wav
         6-channel OK : sig1_orig.wav
         6-channel OK : sig1_sys1.wav
         6-channel OK : sig1_sys2.wav
         6-channel OK : sig1_sys3.wav
         6-channel OK : sig1_lp70.wav
         6-channel OK : sig1_lp35.wav

         6-channel OK : sig2_orig.wav
```

```
6-channel OK : sig2_orig.wav
6-channel OK : sig2_sys1.wav
6-channel OK : sig2_sys2.wav
6-channel OK : sig2_sys3.wav
6-channel OK : sig2_lp70.wav
6-channel OK : sig2_lp35.wav


6-channel OK : sig3_orig.wav
6-channel OK : sig3_orig.wav
6-channel OK : sig3_sys1.wav
6-channel OK : sig3_sys2.wav
6-channel OK : sig3_sys3.wav
6-channel OK : sig3_lp70.wav
6-channel OK : sig3_lp35.wav
```

An example check file for signals in the mono file format is shown below. Under the heading "Channels/Diagnostic" are two lines that, read vertically, indicate the ASIO output channel number. If everything is correct then it indicates the mono channels in the stimulus signal set by means of an "x" in that column and the status "OK."

```
Session Check

Session configuration
Session: MUSHRA
File format: mono
Speaker configuration: same
Open reference: yes
Channel Label File (mono files): ITUR5-1.txt
Randomize Trials: yes
Randomize Conditions: yes
Only score item playing: yes
Play mono as stereo: no
3 Trials
6 Conditions per Trial

Diagnostic key
Unexpected signal components: USC
Number of channels: NC
Sampling rate: SR
Audio word size: WS
Number of samples: NS
Channel Mask: CM

                   Channels/Diagnostic   File
            11111111112222222222333
            12345678901234567890123456789012
            xxxxxx                      OK : ./orig/sig1/sig1_orig
            xxxxxx                      OK : ./orig/sig1/sig1_orig
            xxxxxx                      OK : ./lp35/sig1/sig1_lp35
            xxxxxx                      OK : ./lp70/sig1/sig1_lp70
            xxxxxx                      OK : ./sys1/sig1/sig1_sys1
            xxxxxx                      OK : ./sys2/sig1/sig1_sys2
            xxxxxx                      OK : ./sys3/sig1/sig1_sys3

            xxxxxx                      OK : ./orig/sig2/sig2_orig
            xxxxxx                      OK : ./orig/sig2/sig2_orig
            xxxxxx                      OK : ./lp35/sig2/sig2_lp35
            xxxxxx                      OK : ./lp70/sig2/sig2_lp70
            xxxxxx                      OK : ./sys1/sig2/sig2_sys1
            xxxxxx                      OK : ./sys2/sig2/sig2_sys2
            xxxxxx                      OK : ./sys3/sig2/sig2_sys3

            xxxxxx                      OK : ./orig/sig3/sig3_orig
            xxxxxx                      OK : ./orig/sig3/sig3_orig
            xxxxxx                      OK : ./lp35/sig3/sig3_lp35
            xxxxxx                      OK : ./lp70/sig3/sig3_lp70
```

```
            xxxxxx                              OK : ./sys1/sig3/sig3_sys1
            xxxxxx                              OK : ./sys2/sig3/sig3_sys2
            xxxxxx                              OK : ./sys3/sig3/sig3_sys3
```

A check file for a more complicated test is shown here (slightly re-formatted to fit to the page). Notice that in the first portion has the line

> Speaker configuration: different

which indicates that in this test different stimuli will play to different loudspeakers. This can be confirmed by inspecting the matric of x's under the channel number indication.

```
Session Check

Session configuration
Session: MUSHRA
File format: mono
Speaker configuration: different
Open reference: yes
Channel Label File (mono files): CfP_loudspeaker_config.txt
Randomize Trials: yes
Randomize Conditions: yes
Only score item playing: yes
Play mono as stereo: no
12 Trials
5 Conditions per Trial

Diagnostic key
Unexpected signal components: USC
Number of channels: NC
Sampling rate: SR
Audio word size: WS
Number of samples: NS
Channel Mask: CM

              Channels/Diagnostic   File
          1111111111222222222333
 1234567890123456789012345678901
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Orig/CO_01_Church/CO_01_Church
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Orig/CO_01_Church/CO_01_Church
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./LP35/CO_01_Church/LP35_CO_01_Church
 xxx    xx     xx xx        x     OK : ./Test2/Sys1/CO_01_Church/Sys1_128_CO_01_Church
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Test2/Sys2/CO_01_Church/Sys2_128_CO_01_Church
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Test2/Sys3/CO_01_Church/Sys3_128_CO_01_Church


 xxx      xx xxx  xx  xxxx        OK : ./Orig/CO_04_Fountain/CO_04_Fountain
 xxx      xx xxx  xx  xxxx        OK : ./Orig/CO_04_Fountain/CO_04_Fountain
 xxx      xx xxx  xx  xxxx        OK : ./LP35/CO_04_Fountain/LP35_CO_04_Fountain
 xxx    xx   x xx        x        OK : ./Test2/Sys1/CO_04_Fountain/Sys1_128_CO_04_Fountain
 xxx      xx xxx  xx  xxxx        OK : ./Test2/Sys2/CO_04_Fountain/Sys2_128_CO_04_Fountain
 xxx      xx xxx  xx  xxxx        OK : ./Test2/Sys3/CO_04_Fountain/Sys3_128_CO_04_Fountain


 xxx    xx   x  xx  xx   x   x    OK : ./Orig/CO_05_BarcoTR/CO_05_BarcoTR
 xxx    xx   x  xx  xx   x   x    OK : ./Orig/CO_05_BarcoTR/CO_05_BarcoTR
 xxx    xx   x  xx  xx   x   x    OK : ./LP35/CO_05_BarcoTR/LP35_CO_05_BarcoTR
 xxx    xx      xx  xx       x    OK : ./Test2/Sys1/CO_05_BarcoTR/Sys1_128_CO_05_BarcoTR
 xxx    xx   x  xx  xx   x   x    OK : ./Test2/Sys2/CO_05_BarcoTR/Sys2_128_CO_05_BarcoTR
 xxx    xx   x  xx  xx   x   x    OK : ./Test2/Sys3/CO_05_BarcoTR/Sys3_128_CO_05_BarcoTR


 xxx    xx      xx  xx            OK : ./Orig/CO_08_Musik2/CO_08_Musik2
 xxx    xx      xx  xx            OK : ./Orig/CO_08_Musik2/CO_08_Musik2
 xxx    xx      xx  xx            OK : ./LP35/CO_08_Musik2/LP35_CO_08_Musik2
 xxx    xx      xx  xx            OK : ./Test2/Sys1/CO_08_Musik2/Sys1_128_CO_08_Musik2
 xxx    xx      xx  xx            OK : ./Test2/Sys2/CO_08_Musik2/Sys2_128_CO_08_Musik2
 xxx    xx      xx  xx            OK : ./Test2/Sys3/CO_08_Musik2/Sys3_128_CO_08_Musik2


 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Orig/CO_11_Betty3/CO_11_Betty3
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Orig/CO_11_Betty3/CO_11_Betty3
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./LP35/CO_11_Betty3/LP35_CO_11_Betty3
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Test2/Sys1/CO_11_Betty3/Sys1_128_CO_11_Betty3
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Test2/Sys2/CO_11_Betty3/Sys2_128_CO_11_Betty3
 xxxxxxx  xxxxxx  xx  xxxxxxxxx   OK : ./Test2/Sys3/CO_11_Betty3/Sys3_128_CO_11_Betty3
```

3.4.4 Score File

STEP Score files have extension `txt`. The format of the MUSHRA score files is shown below. The file is tab-separated ASCII text with a header line followed by one line for each signal line in the associated

Session file. Note that the "Sig" column has the signal or test item label from the session file and the "Sys" column has the system label from the session file.

```
Lab    Listener  Session   File                     Sig    Sys    Score
lab1   lis1      1         ./test1/sig1_orig.wav    Sig1   HR     97
lab1   lis1      1         ./test1/sig1_sys1.wav    Sig1   Sys1   96
lab1   lis1      1         ./test1/sig1_sys2.wav    Sig1   Sys2   66
lab1   lis1      1         ./test1/sig1_sys3.wav    Sig1   Sys3   98
lab1   lis1      1         ./test1/sig1_lp70.wav    Sig1   LP70   28
lab1   lis1      1         ./test1/sig1_lp35.wav    Sig1   LP35   22
lab1   lis1      1         ./test1/sig2_orig.wav    Sig2   HR     100
lab1   lis1      1         ./test1/sig2_sys1.wav    Sig2   Sys1   86
lab1   lis1      1         ./test1/sig2_sys2.wav    Sig2   Sys2   74
lab1   lis1      1         ./test1/sig2_sys3.wav    Sig2   Sys3   53
lab1   lis1      1         ./test1/sig2_lp70.wav    Sig2   LP70   34
lab1   lis1      1         ./test1/sig2_lp35.wav    Sig2   LP35   20
```

Any time the subject exits STEP prior to completing a test (actually, after each completed trial), STEP creates an interim score file in the same directory as the session file. This file "remembers" all the subject's scores through the last completed trial. In this way a subject can exit at test at any and re-start exactly where he or she left off.

3.4.5 Initialization File

Each time you quit STEP while in a testing mode, it creates a "step.ini" file in the same directory as the Session file. This file "remembers" all the options that the administrator may have selected. The next time STEP is started in that directory, default option settings are over-ridden by the initialization file settings. This insures consistency from one listener session to the next. If one does not desire this behavior, simply delete the "step.ini" file prior to starting STEP.

3.5 Administrator Panel

The Administrator panel controls all parameters of the test presentation. It is shown here, and each of the settings made using this panel is described.
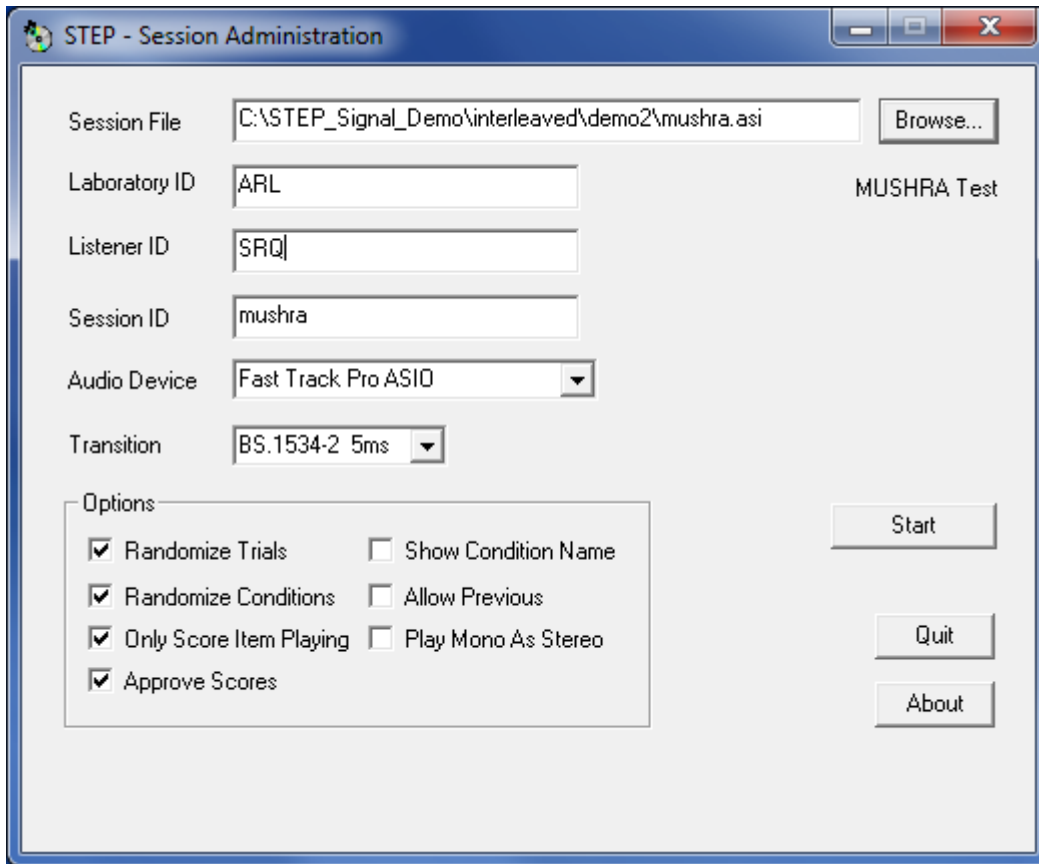
Session File – This text box indicates the Session file. The Browse button permits the administrator to easily select the desired Session file. At this point the Session file is opened and parsed. In the case of long Session files containing multi-channel signals, this may take several seconds, in which case a "progress bar" will appear at the bottom of the Session Administration window.

The parsing of the Session file will produce a report contained in the file "`check.txt`" located in the same directory as the Session file. The format of the check file is discussed in Section 3.4.3.

Laboratory ID, Listener ID and Session ID – The administrator must enter values for Laboratory ID, Listener ID in these text boxes. The Laboratory ID is typically a short string indicating the test site. Similarly, the Listener ID is a short string indicating the listener. STEP fills in the Session ID based on the Session filename and the channel mapping filename, but this can be over-written by the user by typing into the textbox. The literal string "Scores", the Session ID, the Laboratory ID and the Listener ID are concatenated (separated by "_") to provide a unique Score filename.

Audio Device – This drop-down box permits selection of the audio output device.

Transition – This drop-down box permits selection of the transition format when STEP switches from playing one stimulus to another. Choices are:
    Cross-Fade   Fade out the first stimulus while fading in the next. Cross-fade waveforms are a raised cosine. The transition occurs in one ASIO buffer interval.
    BS.1544-2-5ms Fade out of the first stimulus in 5 ms, followed by 5 ms of silence followed by fade-in of the next stimulus. This is what is recommended in BS.1534-2 [2.].
    BS.1544-2-10ms Same as above, but all intervals are 10 ms.

Options
The Options panel provides the following options, which are selected via check-box:

16

Randomize Trials – If selected then STEP will automatically randomize the order of presentation of trials with respect to the order indicated in the Session file. This is the default for all test modes. Regardless of whether or not this is selected, the Score file still corresponds exactly to the Session file.

Randomize Conditions – If selected then STEP will automatically randomize the order of conditions on the trial or test panel (i.e. randomize the correspondence between the processed files in the Session file and the lettered buttons on the test panel). This is the default for all test modes. Regardless of whether or not this is selected, the Score file still corresponds exactly to the Session file.

Only Score Item Playing – If selected then only the slider associated with the condition currently being played out can be adjusted for scoring.

Approve Scores – If selected then STEP will raise a dialog box with a confirmation checkbox that the listener must select in order to proceed to the next trial.

Show Signal Name – If selected then the signal filename will appear at the bottom of the trial or test panel. This is the default for training. By default it is *not* selected for any test mode. Note that, if the Session file is not "double-blind" selecting this may reveal the identity of the System under Test. However, if the purpose of the session is "engineering evaluation" and it is desired to know what condition a letter button corresponds to, this option should be selected.

Allow previous – If selected then the Prev button will appear on the test panel. This is the default for training. By default it is *not* selected for any test mode.

Play Mono as Stereo – If selected then mono test items are played to both left and right output channels. If not selected, mono items are played to left channel only.

The following table indicates the options selected by default for each test methodology:

Option	BS-1116	MUSHRA	AB5, AB7, ABX Comparison	MOS	Training
Has Open Reference	X	X			X
Randomize Trials	X	X	X	X	
Randomize Conditions	X	X	X	X	
Only Score Item Playing					
Approve Scores	X	X	X	X	
Show Condition Name					
Allow previous					X
Play Mono as Stereo	X	X	X	X	X

Some test methodologies have additional checks on the scoring:

BS-1116	One of A or B must be scored at 5.0
MUSHRA	One of the sliders (score for A, B, etc.) must be scored at 100
ABX	X must be sored as one of A or B

Start – This starts the test, which causes the Administrator panel to be replaced by the appropriate trial or test panel.

Quit – This exits STEP.

About – This shows information about this version of STEP.

3.6 Test Panels

3.6.1 Common Controls and Behavior

For all test methodologies each invocation of the test panel represents a *trial*, and the trial number and signal name is displayed in the upper left corner of the panel. In each trial the listener must make one or more subjective judgments by setting the values on a set of vertical sliders. The subjective rating scale is shown to the left of the set of vertical sliders.

Play is initiated by clicking on the Condition (or systems under test) buttons beneath each vertical slider, and the presentation software supports seamless switching from one condition to another. At any time the listener can click the "pause" icon button to pause or the "play" icon button to play the selected signal. Selecting the alternate action "Loop" button causes the signal to play continuously. Each of these has a keyboard shortcut, as listed here:

Key	Function
R	Ref button click
A, B, C, …	A, B, C, … button click
Space Bar	Alternately Play or Pause
L	Loop On/Off

The horizontal Position slider indicates the instantaneous position in the signal waveform. Grabbing and moving the Start slider alters the start point for waveform looping, and similarly moving the Stop slider alters the end point, thus permitting a "loop and zoom" function that is particularly powerful for subjective evaluation.

Note that a subject can exit the test at any time (by selecting the "X" button in the top right corner of the MUSHRA panel), and the subject's interim results are saved in the listener's Score file. Re-entering the test will start exactly where the listener previously left off. When all trials have been scored, the Administrator window appears again.

3.6.2 MUSHRA

In each trial in the MUSHRA test methodology the listener must rate several conditions, which are denoted by letters (A, B, C, etc.) underneath the vertical scroll bars. The reference is indicated by the button "REF" to the left of the letter buttons. The listener selects which of reference or conditions A, B, C, etc. he or she wishes to hear at a given instant by clicking on the REF or letter buttons. When the listener is satisfied with the scoring, he or she clicks on the NEXT button to proceed to the next trial. The MUSHRA panel is shown here:
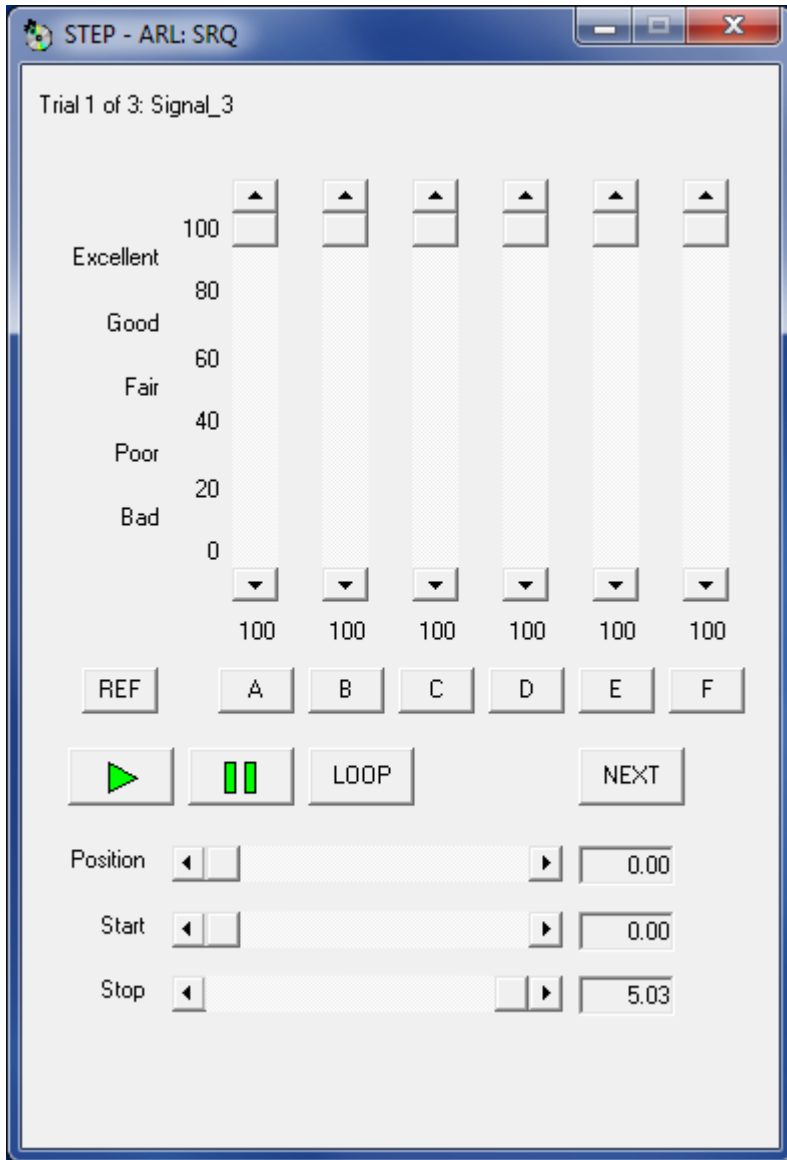
3.6.3 BS.1116

In each trial in the BS.1116 test methodology the listener must rate two conditions which are denoted by letters A and B underneath the vertical scroll bars. One of them is a hidden reference. The listener also has available the "open reference" which is indicated by the button "REF" to the left of the letter buttons. The listener selects which of reference or conditions A or B he or she wishes to hear at a given instant by clicking on the REF or letter buttons. Note that, because the listener knows that one of A or B is a hidden reference, the "Check Scores" option checks that one of the A or B scores has the value 5.0 (implicitly indicating that it is the hidden reference). When the listener is satisfied with the scoring, he or she clicks on the NEXT button to proceed to the next trial. The BS.1116 panel is shown here:
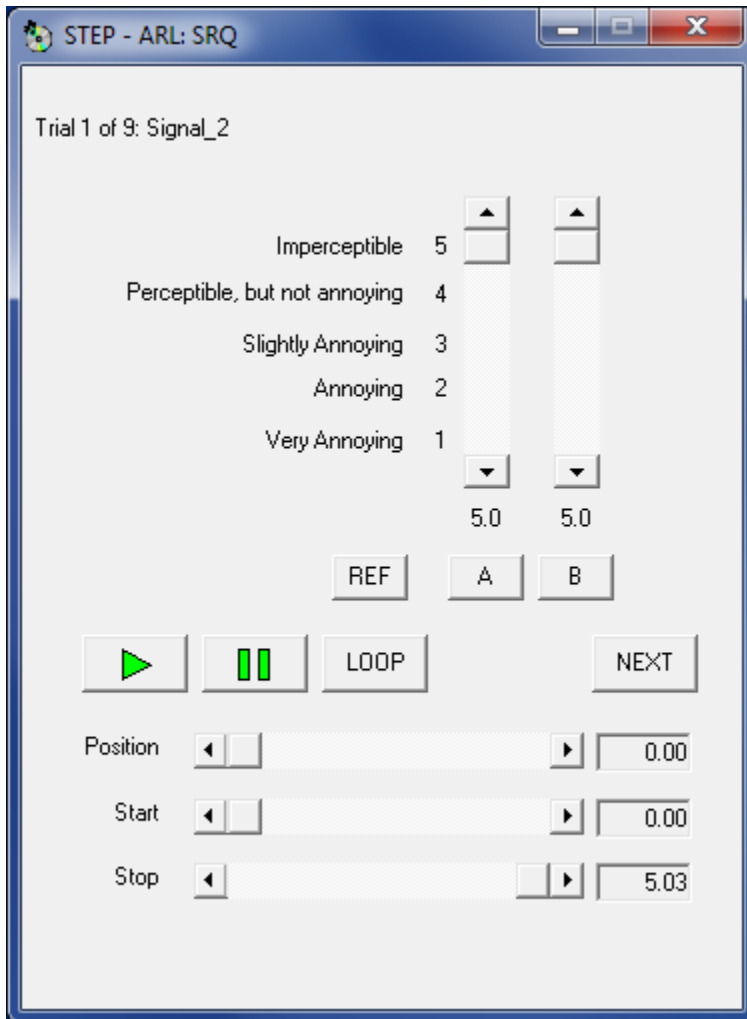
3.6.4 A-B Comparison

The A-B Comparison mode invokes a two-comparison forced-choice test methodology using either a 5-point or a 7-point response scale. The listener must make a single judgment by adjusting the vertical slider, which indicates the extent that condition A is (or is not) better than condition B. The optional reference (specified by has_open_reference=true in the session header) is selected by the button "REF" to the left of the letter buttons. When the listener is satisfied with the scoring, he or she clicks on the NEXT button to proceed to the next trial.

The A-B Comparison Session file supports quite a bit of flexibility, as described here:
5- or 7-point scale – You can invoke either a 5-point scale (+/- 2) or a 7-point scale (+/-3) via the session value as follows:

```
session=AB5 (5-pont scale)
session=AB7 (7-pont scale)
```

The A-B Comparison panel is shown here:

Audio Research Labs

3.6.5 ABX Comparison

The ABX Comparison mode invokes a two-comparison forced-choice test methodology. The listener must make a single judgment by clicking on either the A or B radio-buttons to indicate that X is the signal A or the signal B. When the listener is satisfied with the scoring, he or she clicks on the NEXT button to proceed to the next trial.
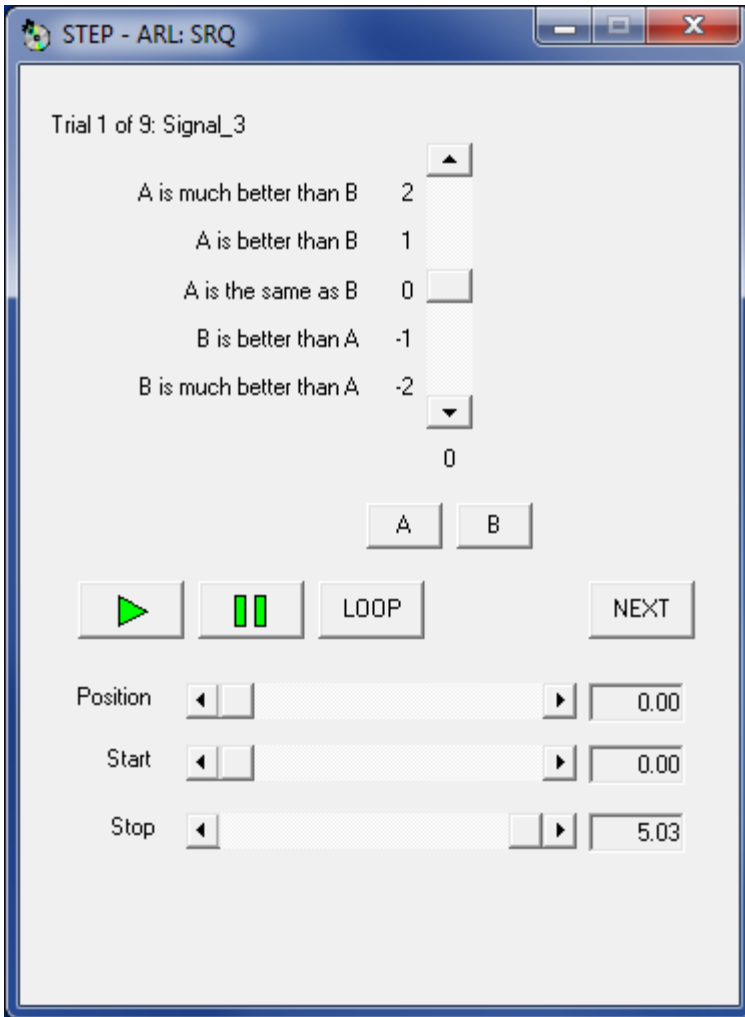
The A-B Comparison panel is shown here:
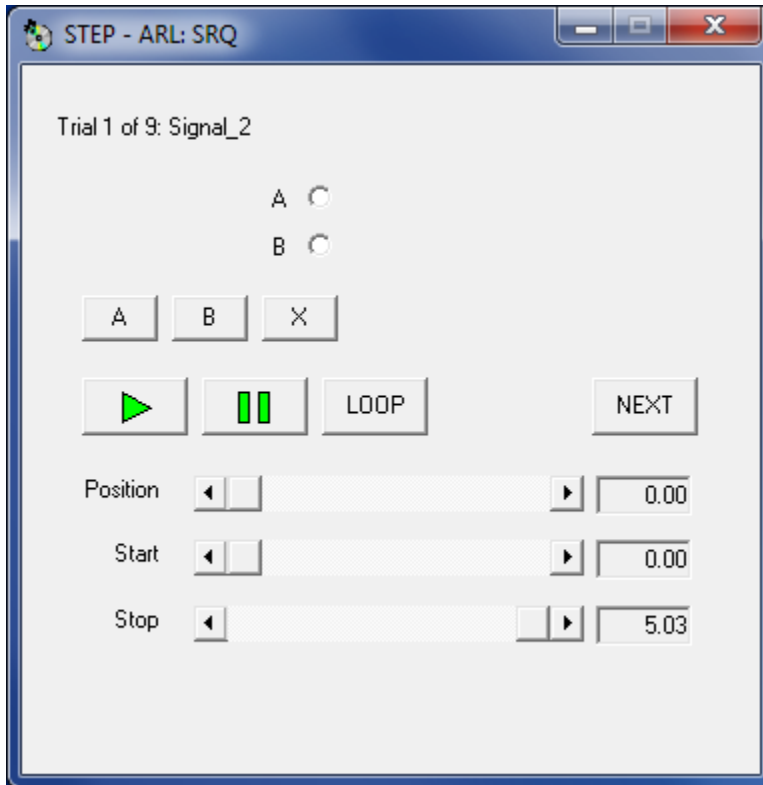
3.6.6 MOS

The MOS test is a single-stimulus Absolute Category Rating (ACR) test methodology. The listener must make a single judgment of the subjective quality of signal A. When the listener is satisfied with the scoring, he or she clicks on the NEXT button to proceed to the next trial.

The MOS panel is shown here:
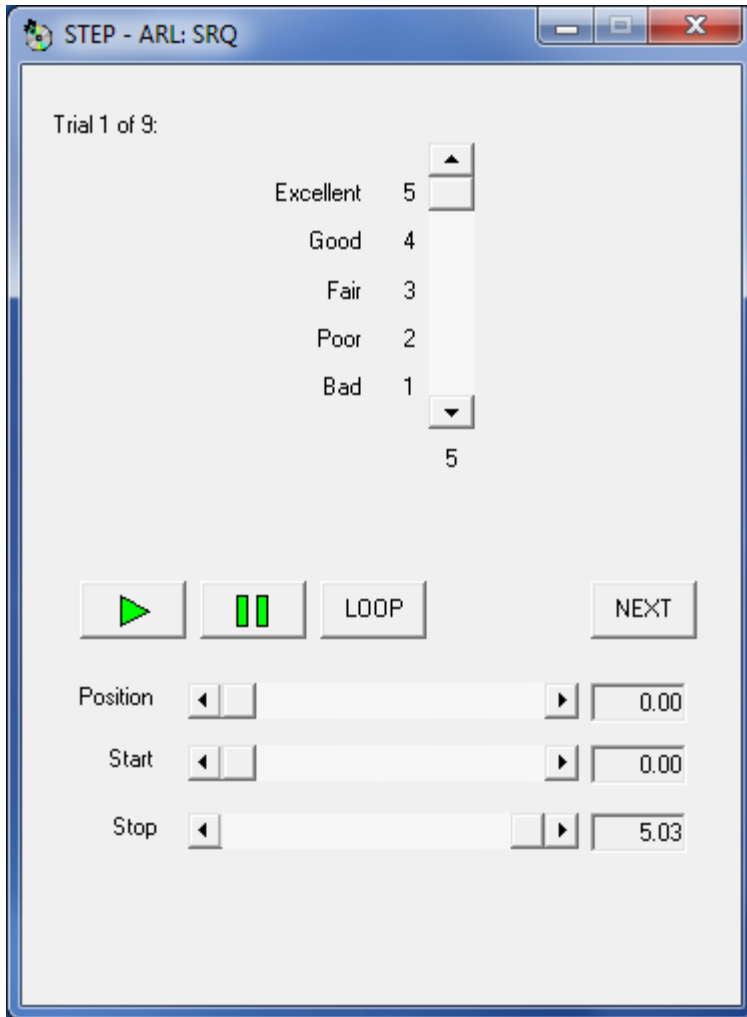
3.6.7 Training

This mode can be used either for listener training or "Engineering Evaluation." It does not require or support scoring, so the vertical sliders are not present. Since there is no scoring, it is not required to enter either a Laboratory or Listener ID. Optionally, the processed filename is shown at the bottom of the panel. Also, it allows browsing forward or back in the Session file, so that both PREV and NEXT buttons are available. The Training panel is shown here:

3.6.8 Playlist

This mode provides a simple means of playing a set of files. The "session" file is just a list of files to play, such as:

```
sig1_orig.wav
sig1_sys1.wav
sig1_sys2.wav
sig1_sys3.wav
```

It only supports sequential playing of files, so there are no "A," "B" or "C" buttons, only NEXT and PREV.

Audio Research Labs

4 Training

Every listener should receive training prior to their first test session. This has two functions: first, to familiarize the listener with the range and types of impairments that he or she will hear in the test and second, to familiarize the listener with the trial panel to be used in the test.
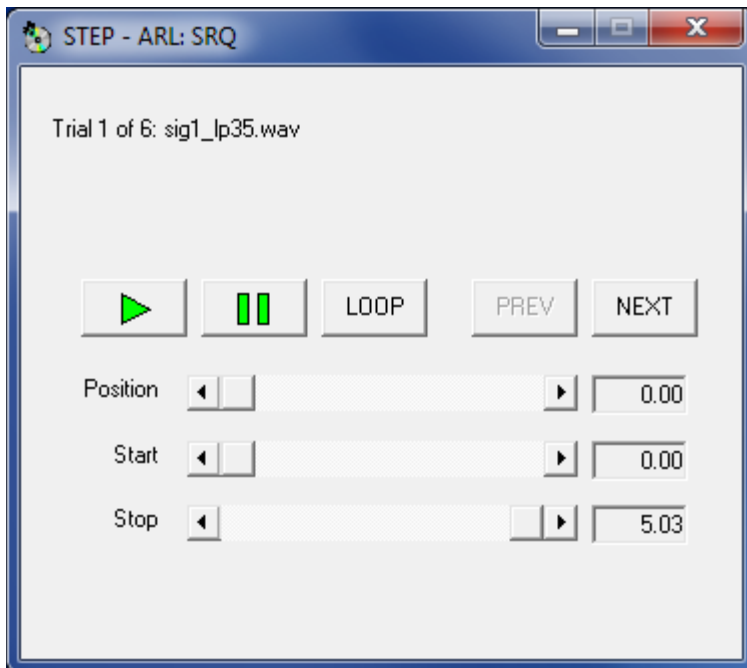
For the first objective, STEP can be used in *training* mode. In training mode the signal name and system name are both visible on the trial panel (see Section 3.2.1) and the correspondence between filenames in the session file and buttons on the user panel are not randomized (i.e. REF is the first file, A is the second file, etc.). Using the training panel, the test administrator and the listeners can step through representative test files and discuss what impairments should be apparent. Remember – the administrator can discuss what *impairment* might be heard but not what *score* it might receive! A portion of the training Session file is shown here:

Line in Session file		Description
session=training		Session type
# Sig1		Displayed signal name
sig1_sys1.wav	Sys1	System under test Sys1
sig1_sys2.wav	Sys2	System under test Sys2
sig1_sys3.wav	Sys3	System under test Sys3
# Sig2		Displayed signal name
sig2_sys1.wav	Sys1	System under test Sys1
sig2_sys2.wav	Sys2	System under test Sys2
sig2_sys3.wav	Sys3	System under test Sys3

For the second objective, STEP can be used in the target test mode, but with a set of training files. The test administrator can walk the listeners through using the trial panel to listen to and score the test items. Remember that the training items should be distinct from the test items!

5 Score analysis using Excel

5.1 Overview

It is straightforward to analyze the results produced by STEP using a Microsoft Excel Pivot Table. An example of Session files, concatenated Score files and Excel Pivot Table analysis for each methodology that STEP supports is included in the demo directory. A step by step walk though of the analysis is given in this section. MUSHRA is presented first, as it is most straightforward. Each other methodology is presented as differences relative to the MUSHRA analysis.

5.2 Score file concatenation

There is an additional executable in the STEP directory: step_sc.exe. This supports easy assembly of individual score files into a single score file that is ready to open in Excel. It is easiest to use if you create a shortcut to step_sc.exe, for example in your current working directory or on your desktop. Double-clicking on the shortcut brings up this panel:
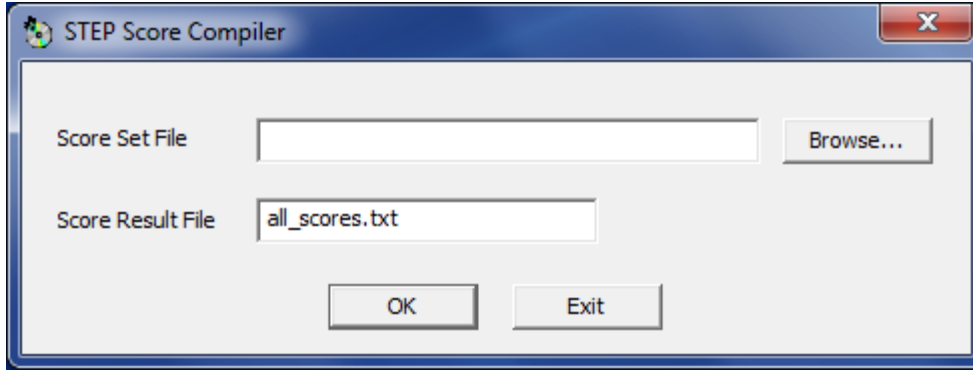
The "Browse" button permits you to select a file that contains the pathname of all individual score files that you wish to include in the test. Note that these pathnames are relative to the directory containing the Score Set File. For example:

```
scores\Lis_1.txt
scores\Lis_2.txt
scores\Lis_3.txt
scores\Lis_4.txt
scores\Lis_5.txt
scores\Lis_6.txt
```

After selection, the pathname of this file is shown in the Score Set File box.

The Score Result File box contains the default filename that will contain the concatenated scores. You can edit this as appropriate. Clicking on the OK button creates the concatenated score file.

5.3 MUSHRA

5.3.1 Session file

Assume that the STEP Session file uses "in the clear" naming. In this example, test items are named `sig1`, `sig2`, etc. and systems under test are named `sys1`, `sys2`, etc. but they could just as easily be named `pop1`, `symphony1` (where different signals are different music genres) and `AAC_Enc1`, `AAC_Enc2`, (where we are testing different MPEG Advanced Audio Coding encoders). Note that the Session file contains signals and Systems under Test in a systematic order: Sig1, Sig2, Sig3, etc. for each trial and HR, Sys1, Sys2, Sys3, LP70 and LP35 for each system within a trial. Creating a Session file in this manner is straightforward, and since STEP by default randomizes the presentation order of trials and the assignment of systems to Condition letters A, B, C, etc., there is no reason not to create a Session file in this manner.

In this example, original (i.e. unprocessed) items are in subdirectory `originals`, anchor files (e.g. low pass filtered originals) are in subdirectory `anchors`, and processed files are located in a separate subdirectory for each System under Test (e.g. `coder1, coder2`, etc.). A portion of the Session file is shown here:

Line in Session file		Description
`session=mushra`		Session type
`# Sig1`		Displayed signal name
`originals/sig1_orig.wav Ref`		Open reference
`originals/sig1_orig.wav HR`		Hidden reference
`coder1/sig1.wav`	`Sys1`	System under test Sys1
`coder2/sig1.wav`	`Sys2`	System under test Sys2
`coder3/sig1.wav`	`Sys3`	System under test Sys3
`anchors/sig1_lp70.wav`	`LP70`	7.0 kHz lowpass anchor
`anchors/sig1_lp35.wav`	`LP35`	3.5 kHz lowpass anchor
`# Sig2`		Displayed signal name

```
originals/sig2_orig.wav Ref              Open reference
originals/sig2_orig.wav HR               Hidden reference
coder1/sig2.wav         Sys1             System under test Sys1
coder2/sig2.wav         Sys2             System under test Sys2
coder3/sig2.wav         Sys3             System under test Sys3
anchors/sig2_lp70.wav   LP70             7.0 kHz lowpass anchor
anchors/sig2_lp35.wav   LP35             3.5 kHz lowpass anchor
```

A portion of an equivalent Session file that captures signal and system in the filename rather than using directory structures is shown below. If we had wanted to make a truly double-blind experimental setup, we could have named every signal "sigN_sysM.wav, where N is signal number and N is system number, which would include the Hidden Reference, all Systems under Test and all low-pass anchors.

Line in Session file		Description
session=mushra		Session type
# Sig1		Displayed signal name
sig1_orig.wav	REF	Open reference
sig1_orig.wav	HR	Hidden reference
sig1_sys1.wav	Sys1	System under test Sys1
sig1_sys2.wav	Sys2	System under test Sys2
sig1_sys3.wav	Sys3	System under test Sys3
sig1_lp70.wav	LP70	7.0 kHz lowpass anchor
sig1_lp35.wav	LP35	3.5 kHz lowpass anchor
# Sig2		Displayed signal name
sig2_orig.wav	REF	Open reference
sig2_orig.wav	HR	Hidden reference
sig2_sys1.wav	Sys1	System under test Sys1
sig2_sys2.wav	Sys2	System under test Sys2
sig2_sys3.wav	Sys3	System under test Sys3
sig2_lp70.wav	LP70	7.0 kHz lowpass anchor
sig2_lp35.wav	LP35	3.5 kHz lowpass anchor

5.3.2 Score file

This Session file shown above results in a Score file for each listener, where the filename is whatever text string was entered into the Session ID field field concatenated with the string in the Listener ID and with the suffix `txt`. The first part of the Score file for listener "SRQ" is shown below.

The rows of the Score file consist of tab-separated fields with one score per row, such that the Score files can be directly opened by Microsoft Excel. The signal name string (after the "#" in the Session file) appears in the `Sig` column of the Score file. Similarly, system name string (after the file pathname in the Session file) appears in the `Sys` column.

```
Score File                                                Line Description
Lab   Listener  Session   File            Sig   Sys   Score   Header line
ARL   SRQ       mushra    sig1_orig.wav   Sig1  HR    100     Hidden reference
ARL   SRQ       mushra    sig1_sys1.wav   Sig1  Sys1  96      Sys1
ARL   SRQ       mushra    sig1_sys2.wav   Sig1  Sys2  66      Sys2
ARL   SRQ       mushra    sig1_sys3.wav   Sig1  Sys3  98      Sys3
ARL   SRQ       mushra    sig1_lp70.wav   Sig1  LP70  28      7.0 kHz anchor
ARL   SRQ       mushra    sig1_lp35.wav   Sig1  LP35  22      3.5 kHz anchor
ARL   SRQ       mushra    sig2_orig.wav   Sig2  HR    100     Hidden reference
ARL   SRQ       mushra    sig2_sys1.wav   Sig2  Sys1  86      Sys1
ARL   SRQ       mushra    sig2_sys2.wav   Sig2  Sys2  74      Sys2
ARL   SRQ       mushra    sig2_sys3.wav   Sig2  Sys3  53      Sys3
ARL   SRQ       mushra    sig2_lp70.wav   Sig2  LP70  34      7.0 kHz anchor
ARL   SRQ       mushra    sig2_lp35.wav   Sig2  LP35  10      3.5 kHz anchor
```

5.3.3 Excel data worksheet

The set of Score files must be concatenated, but with only one header line as the first line in the file. The resulting file (which should retain the extension txt) should be opened by Excel and loaded into a worksheet named, for example "Scores." The first portion of such a worksheet is shown below:

Lab	Listener	Session	File	Sig	Sys	Score
ARL	SRQ	mushra	sig1_orig.wav	Sig1	HR	100
ARL	SRQ	mushra	sig1_sys1.wav	Sig1	Sys1	96
ARL	SRQ	mushra	sig1_sys2.wav	Sig1	Sys2	66
ARL	SRQ	mushra	sig1_sys3.wav	Sig1	Sys3	98
ARL	SRQ	mushra	sig1_lp70.wav	Sig1	LP70	28
ARL	SRQ	mushra	sig1_lp35.wav	Sig1	LP35	22
ARL	SRQ	mushra	sig2_orig.wav	Sig2	HR	100
ARL	SRQ	mushra	sig2_sys1.wav	Sig2	Sys1	86
ARL	SRQ	mushra	sig2_sys2.wav	Sig2	Sys2	74
ARL	SRQ	mushra	sig2_sys3.wav	Sig2	Sys3	53
ARL	SRQ	mushra	sig2_lp70.wav	Sig2	LP70	34
ARL	SRQ	mushra	sig2_lp35.wav	Sig2	LP35	10

5.3.4 Excel PivotTable

Use the score worksheet data to create a Pivot Table in Excel as follows:

Recent versions of Excel

Begin by doing the following:
- Concatenate all Score files, retaining only the first header line. Note that step_sc.exe can be used to for this. See Section 5.2.
- Open the resulting file using Excel. Rename worksheet "Sheet 1" (which should contain the scores) to "Scores."

To create the Pivot Table, do the following:
- Under the Insert tab, click on PivotTable. This pops up a window permitting you to select the PivotTable data.
  - You should select all the score data including the header row.
  - For "Where do you want PivotTable to be placed," select "New Worksheet"
  - Click "OK"
- Build the Report:

- o Drag Lab, Listener, Session and Sig down to the "Filters" or "Report Filter" box. Ignore the File label, it will not be used.
- o Drag Sys down to the "Rows" or "Row Labels" box.
- o Drag Score down to the "Values" box *three times*. For each one, use the drop down menu to select "Value Field Settings."
  - ▪ Set the first to "Average"
  - ▪ Set the second to "StdDev"
  - ▪ Set the third to "Count"

Excel 2000

These are the steps unique to Excel 2000.

To create the Pivot Table, do the following:
- Under the Data tab, select PivotTable (which will pop up the PivotTable Wizard). In the Wizard do the following:
- Step 1 of 3:
  - o For question "Where is the data that you want to analyze?" select "Microsoft Excel list"
  - o For question "What kind of report do you want to create?" select "PivotTable"
  - o Select Next
- Step 2 of 3
  - o For question "Where is the data that you want to use?" click on the *selection* icon and select all the listener data (in "Scores" worksheet) and then click on the *exit* icon.
  - o Select Next
- Step 3 of 3
  - o For question "Where do you want to put the PivotTable report?" select "New worksheet"
  - o Click on the Layout button
- Layout
  - o Drag Lab to the Page drop zone
  - o Drag Listener to the Page drop zone
  - o Drag Session to the Page drop zone
  - o Ignore the File item – this will not be used.
  - o Drag Sig to the Page drop zone
  - o Drag Sys (i.e. Condition or System under Test) to the Row drop zone
  - o Drag Score to the Data drop zone *three times* (where it will appear as "Sum of Score")
    - ▪ Double-click on the first "Sum of Score" and change to "Average"
    - ▪ Double-click on the first "Sum of Score" and change to "StdDev"
    - ▪ Double-click on the second "Sum of Score" and change to "Count"
  - o Click on OK to exit Layout panel
  - o Click on Finish to exit PivotTable wizard.
- Reformat
  - o In the PivotTable sheet, right-click anywhere in the PivotTable and select "Show PivotTable Toolbar."
  - o On the PivotTable Toolbar, left-click the "Report Format" icon
    - ▪ Select the upper left format choice (selection indicated by bold outline)
    - ▪ Click OK to exit Format Panel

5.3.5 Construct Plot

The final steps are, for the most part, the same in all versions of Excel.

- Construct a table as shown below, where for each table row
  - o Construct the 95% confidence (CI) interval for each of the Conditions using the following Excel formula. Note that the Student t-distribution is used. This is always correct, and converges to the Gaussian distribution as count become greater than 50.
    - ▪ =TINV(5%,Count-1)*StdDev/SQRT(Count), where StdDev and Count are the PivotTable values for that condition.
  - o Condition is the label for the condition or System under Test.

- o   Mean is the PivotTable Average value for that condition.
- o   Upper is the Mean plus the 95% confidence interval value.
- o   Lower is the Mean minus the 95% confidence interval value.

CI	Sys	High	Low	Mean
4.03	Sys1	85.90	77.85	81.88
4.33	Sys2	80.13	71.47	75.80
4.68	Sys3	68.75	59.40	64.08
2.20	HR	98.35	93.95	96.15
3.87	LP70	33.24	25.51	29.38
4.24	LP35	22.21	13.74	17.98

- • Plot the result using the "Stock" chart type
  - o Select "Layout" tab under "Chart Tools" and fill in Title (e.g. "Test 1"), X axis as "Systems" and Y axis as "Score".
  - o Delete the "legend" as it has no meaning in this context
- • Click on the Y-axis of the chart
  - o Right-click "Format axis"
  - o Select "Scale" tab and designate Minimum to be 0, Maximum to be 100
  - o Select "Number" tab and set Decimal places to be 0

The resulting Pivot Table and 95% CI plot are shown here:
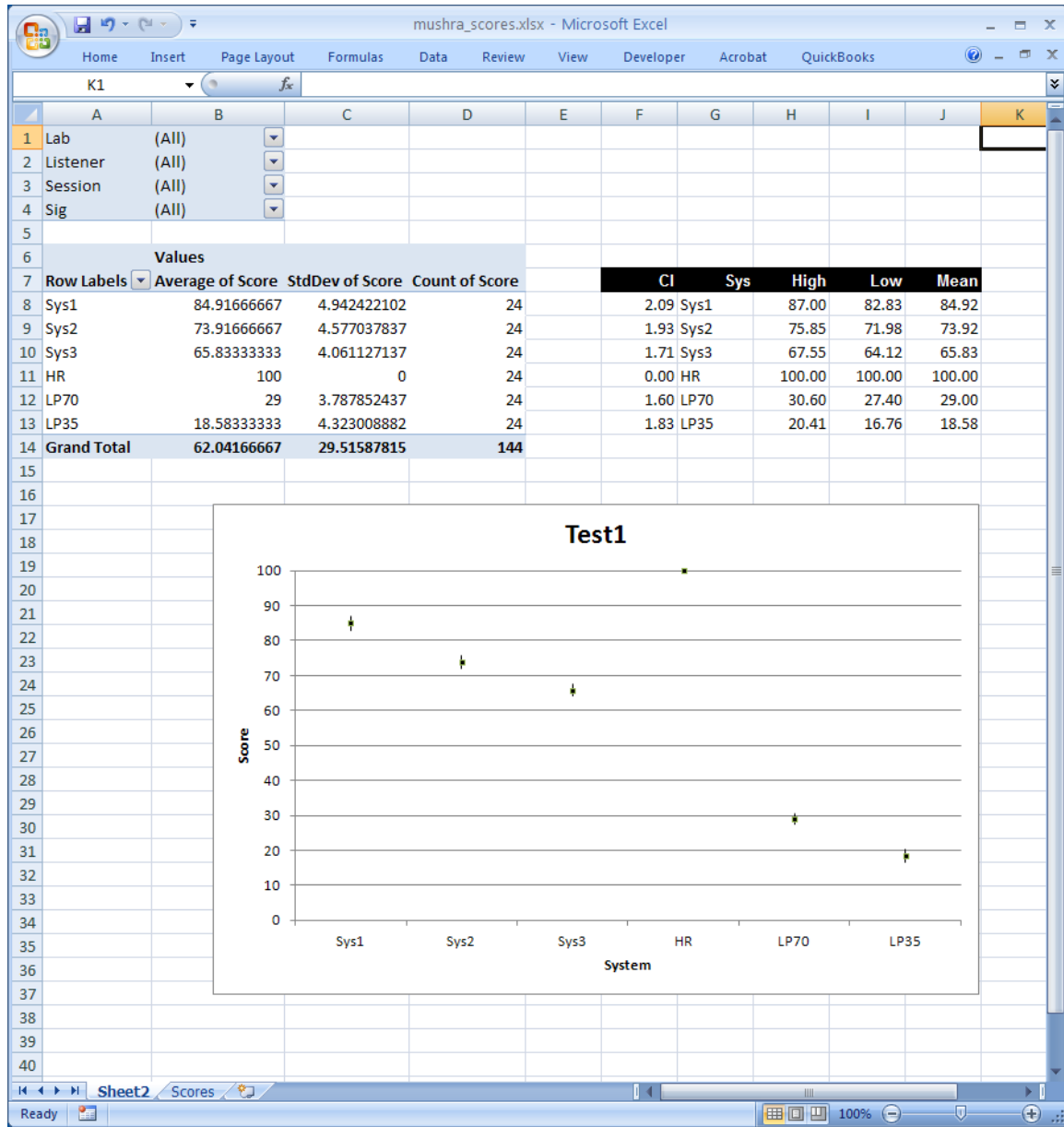
5.4 BS-1116

The analysis of data from BS-1116 test methodology will be presented as differences relative to the MUSHRA analysis found in the previous section.

5.4.1 Session file

We use a Session file that captures signal and system in the filename, and a portion of the Session file is shown below. As with the MUSHRA Session file, the BS.1116 Session file is constructed in a systematic way, since by default STEP automatically randomizes the order of signal presentation and the assignment of Hidden Reference and System under Test to Condition buttons A and B.

Audio Research Labs

Line in Session file		Description
session=BS.1116		Session type
# Sig1		Displayed signal name
sig1_orig.wav	REF	Open reference
sig1_orig.wav	HR	Hidden reference
sig1_sys1.wav	Sys1	System under test Sys1
# Sig1		Displayed signal name
sig1_orig.wav	REF	Open reference
sig1_orig.wav	HR	Hidden reference
sig1_sys2.wav	Sys2	System under test Sys2
# Sig1		Displayed signal name
sig1_orig.wav	REF	Open reference
sig1_orig.wav	HR	Hidden reference
sig1_sys3.wav	Sys3	System under test Sys3
# Sig2		Displayed signal name
Sig2_orig.wav	REF	Open reference
Sig2_orig.wav	HR	Hidden reference
sig2_sys1.wav	Sys1	System under test Sys1
# Sig2		Displayed signal name
Sig2_orig.wav	REF	Open reference
Sig2_orig.wav	HR	Hidden reference
sig2_sys2.wav	Sys2	System under test Sys2
# Sig2		Displayed signal name
Sig2_orig.wav	REF	Open reference
Sig2_orig.wav	HR	Hidden reference
sig2_sys3.wav	Sys3	System under test Sys3

5.4.2 Score file

A portion of the corresponding Score file for one listener is shown in the following table. The Score file presents the signals and Systems under Test in the same order as the Session file. Realize, however, that when Randomize Trials and Randomize Conditions are selected in the Administrator panel (which they are by default), each listener experiences a randomized presentation. In addition, the BS-1116 test methodology the Score file lists the filenames and scores associated with each of buttons A and B for each trial. We will see that this format makes for easier data analysis.

Lab	Listener	Session	FileA	FileB	Sig	SysA	SysB	ScoreA	ScoreB
ARL	SRQ	bs1116	sig1_orig.wav	sig1_sys1.wav	Sig1	HR	Sys1	5	4.4
ARL	SRQ	bs1116	sig1_orig.wav	sig1_sys2.wav	Sig1	HR	Sys2	5	4.3
ARL	SRQ	bs1116	sig1_orig.wav	sig1_sys3.wav	Sig1	HR	Sys3	5	3.2
ARL	SRQ	bs1116	sig2_orig.wav	sig2_sys1.wav	Sig2	HR	Sys1	5	5
ARL	SRQ	bs1116	sig2_orig.wav	sig2_sys2.wav	Sig2	HR	Sys2	5	3.9
ARL	SRQ	bs1116	sig2_orig.wav	sig2_sys3.wav	Sig2	HR	Sys3	5	3.5

5.4.3 Excel data worksheet

Concatenate all the listener Score files, retaining only the first header line and open the resulting file using Excel. The first portion of the resulting worksheet is shown here:

Lab	Listener	Session	FileA	FileB	Sig	SysA	SysB	ScoreA	ScoreB
ARL	SRQ	bs1116	sig1_orig.wav	sig1_sys1.wav	Sig1	HR	Sys1	5	4.4
ARL	SRQ	bs1116	sig1_orig.wav	sig1_sys2.wav	Sig1	HR	Sys2	5	4.3
ARL	SRQ	bs1116	sig1_orig.wav	sig1_sys3.wav	Sig1	HR	Sys3	5	3.2
ARL	SRQ	bs1116	sig2_orig.wav	sig2_sys1.wav	Sig2	HR	Sys1	5	5

ARL	SRQ	bs1116	sig2_orig.wav	sig2_sys2.wav	Sig2	HR	Sys2	5	3.9
ARL	SRQ	bs1116	sig2_orig.wav	sig2_sys3.wav	Sig2	HR	Sys3	5	3.5

5.4.4 Excel Pivot Table

Systematic Session file
The example Session file shown above is in systematic order. If you are using a Session file in systematic order, then you only need to do the following:

Immediately to the right of the score data, create one new column (K) with heading Diff Grade. This will calculate the difference between the System under Test score and the Hidden Reference score. Enter the following formula in that column:

    =J2-I2

and copy this formula through to the last row of score data.

Randomized Session file
If you wish to use a Session file that is *not* in systematic order (for example, due to double-bind test requirements), this section presents a procedure to un-scramble the assignment of Hidden Reference and System under Test to Condition buttons A and B.

Immediately to the right of the score data, create five new columns (columns K, L, M, N and O) with headings HR, Sys, HR Score, Sys Score and Diff Grade. Enter the following formulas:

    under the HR column (K)

        =IF(G2="HR",G2,H2)

    under the Sys column (L)

        =IF(H2="HR", G2,H2)

    under the HR Score column (M)

        =IF(G2="HR",I2,J2)

    under the Sys Score column (N)

        =IF(H2="HR",I2,J2)

    under the Diff Grade column (O)

        =N2-M2

and copy these formulas through to the last row of score data. In this way for each trial column HR always has the hidden reference name, column Sys always has the System under Test name, column HR Score always has the hidden reference score, column Sys Score always has the System under Test score. This makes calculation of the difference grade straightforward. If you wish to do listener post-screening, adjust the column references in the next section accordingly.

Post-Screening
The BS-1116 test methodology assesses two items of information at each trial: the first is whether the listener can identify the hidden reference, which is done by giving it the score of 5.00, and the second is the listeners rating of the System under Test in a given trial. Failure to correctly identify the hidden reference may indicate an unacceptably low level of listener reliability and can be a reason to exclude a listener's scores from the data analysis. If and how this is determined typically depend on the systems under test. For example, a miss-identification of the hidden reference might not be counted against the listener if the System under Test for that particular trial receives a very high score, e.g. greater than 4.90.

The following is one way to use hidden reference identification to perform post-screening on listener data. (Note that it assumes a Session file having systematic order.) Create enter headings PostScr and HR ID in columns L and M. Note that rows 2:10 corresponds to the scores for the first listener. Enter the following formulas:

    In the first cell (L2) under the PostScr column (L)

        =IF(SUM(M2:M10)>0.95*COUNT(M2:M10),1,0)

    in cell L3

        =L2

and copy this formula through row 10.

under the HR ID column (M)
> =IF(I2=5,1,0)

and copy this formula through row 10.

Note that the above assumes that a listener score block consists of 9 scores. You will have to change the "10" to whatever is appropriate for your test. Now copy the cell block for the first listener (L2:M10) to all subsequent listeners. In this way, the PostScr column for each trial will indicate whether that listener correctly identified the hidden reference 95 percent of the time.

Construct Pivot Table

To construct the Pivot Table for the systems under test, proceed as in Section 6.2.4, but
- Select all data for the Pivot Table (i.e. columns A through M or Q)
- In the Layout part of the Pivot Table
    - drop Lab, Listener, Session, Sig and PostScr (if present) to the Report Filter drop zone.
    - drop SysB in the Row drop zone (the System under Test)
    - drop Diff Grade in the Data drop zone (three times, setting to Average, StdDev and Count)
    - ignore the FileA and FileB items

Construct the chart of system mean scores and 95% confidence intervals, but
- Click on the Y-axis of the chart, right-click and select "Format axis"
    - Select "Scale" tab and designate Minimum to be -4, Maximum to be 0
    - Select "Number" tab and set Decimal places to be 1
- Click on the X-axis at the top of the chart, right-click and select "Format axis"
    - Select "Scale" tab and designate Minimum to be -4, Maximum to be 0
    - Select "Patterns" tab and set Tick Mark Labels to be "high"

This should result in the following Pivot Table and chart:
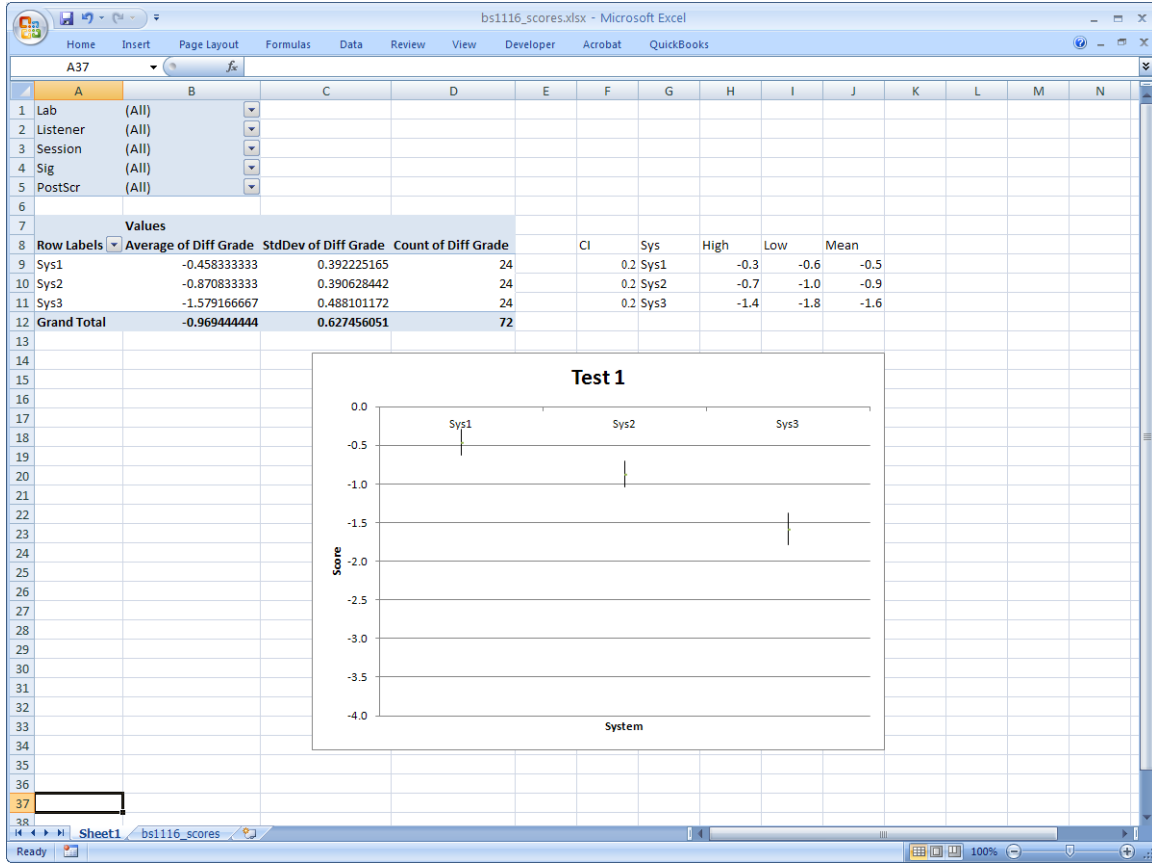
5.5 Comparative: AB5 and AB7

5.5.1 Session file

A portion of an example Session file for the 7-category A-B Comparison forced choice test methodology *with* Reference is shown below. There are two options in creating the A-B Comparison Session file:

- Use 5 or 7 rating categories
- Include open reference and show REF signal button

Open Reference Option
If the option of having an open reference and associated REF button is desired, then include the
`has_open_reference=true`
line in the session file (as shown below).

5 or 7 rating categories
If the session value is AB5 then a 5-category rating scale is displayed. Alternatively, if the session value is AB7 then a 7-category rating scale is displayed.

Line in Session file		Description
`session=AB7`		Session type
`has_open_reference=true`		Include Open Reference
`# Sig1`		Displayed signal name
`sig1_orig.wav`	`REF`	Open Reference
`sig1_sys1.wav`	`Sys1`	System under test Sys1
`sig1_sys2.wav`	`Sys2`	System under test Sys2
`# Sig1`		Displayed signal name
`sig1_orig.wav`	`REF`	Open Reference

35

```
sig1_sys1.wav     Sys1           System under test Sys1
sig1_sys3.wav     Sys3           System under test Sys3
# Sig1                           Displayed signal name
sig1_orig.wav     REF            Open Reference
sig1_sys2.wav     Sys2           System under test Sys2
sig1_sys3.wav     Sys3           System under test Sys3
# Sig2                           Displayed signal name
sig1_orig.wav     REF            Open Reference
sig2_sys1.wav     Sys1           System under test Sys1
sig2_sys2.wav     Sys2           System under test Sys2
# Sig2                           Displayed signal name
sig1_orig.wav     REF            Open Reference
sig2_sys1.wav     Sys1           System under Test Sys1
sig2_sys3.wav     Sys3           System under Test Sys3
# Sig2                           Displayed signal name
sig1_orig.wav     REF            Open Reference
sig2_sys2.wav     Sys2           System under Test Sys2
sig2_sys3.wav     Sys3           System under Test Sys3
```

5.5.2 Score file

The set of Score files must be concatenated, retaining only the first header line in the file. However, no further editing is required. The first portion of the Score file is shown here:

```
Lab        Listener   Session   FileA          FileB          Sig     SysA    SysA    Score
lab1       lis1       ab7       sig1_sys1.wav  sig1_sys2.wav  Sig1    Sys1    Sys2    2
lab1       lis1       ab7       sig1_sys1.wav  sig1_sys3.wav  Sig1    Sys1    Sys3    3
lab1       lis1       ab7       sig1_sys2.wav  sig1_sys3.wav  Sig1    Sys2    Sys3    1
lab1       lis1       ab7       sig2_sys1.wav  sig2_sys2.wav  Sig2    Sys1    Sys2    3
lab1       lis1       ab7       sig2_sys1.wav  sig2_sys3.wav  Sig2    Sys1    Sys3    3
lab1       lis1       ab7       sig2_sys2.wav  sig2_sys3.wav  Sig2    Sys2    Sys3    2
```

5.5.3 Excel data worksheet

The scores.txt file is opened with Excel and put in the "Scores" worksheet. A portion of the worksheet is shown here:

Lab	Listener	Session	FileA	FileB	Sig	SysA	SysA	Score
lab1	lis1	ab7	sig1_sys1.wav	sig1_sys2.wav	Sig1	Sys1	Sys2	2
lab1	lis1	ab7	sig1_sys1.wav	sig1_sys3.wav	Sig1	Sys1	Sys3	3
lab1	lis1	ab7	sig1_sys2.wav	sig1_sys3.wav	Sig1	Sys2	Sys3	1
lab1	lis1	ab7	sig2_sys1.wav	sig2_sys2.wav	Sig2	Sys1	Sys2	3
lab1	lis1	ab7	sig2_sys1.wav	sig2_sys3.wav	Sig2	Sys1	Sys3	3
lab1	lis1	ab7	sig2_sys2.wav	sig2_sys3.wav	Sig2	Sys2	Sys3	2

5.5.4 Excel Pivot Table

Since a score is associated with a comparison, an additional column must be added to the Scores worksheet. Create a new column (J) with the label Sys. In cell J2 insert the formula

=CONCATENATE(G2,"-",H2)

and copy this formulas through to the last row of score data. Now the new column shows both systems associated with Condition buttons A and B, which will be used as a label in the analysis chart.

To construct the Pivot Table for the systems under test, proceed as in Section 6.2.4, but
 • In the Layout part of the Pivot Table

- o   drop Lab, Listener, Session and Sig in the Report Filter drop zone
- o   drop Sys in the Row drop zone
- o   drop Score in the Data drop zone (three times, setting to Average, StdDev and Count)
- o   Ignore the FileA, FileB, SysA and SysB items

Construct the chart of system mean scores and 95% confidence intervals, but
- Click on the Y-axis of the chart, right-click and select "Format axis"
  - o   Set the Minimum to be -3 and Maximum to be 3 (or Minimum to be -2 and Maximum to be 2 if this is a 5-level category comparison)
  - o   Set the Decimal places to be 0
- Set chart title and axis titles

This should result in the following Pivot Table and chart:
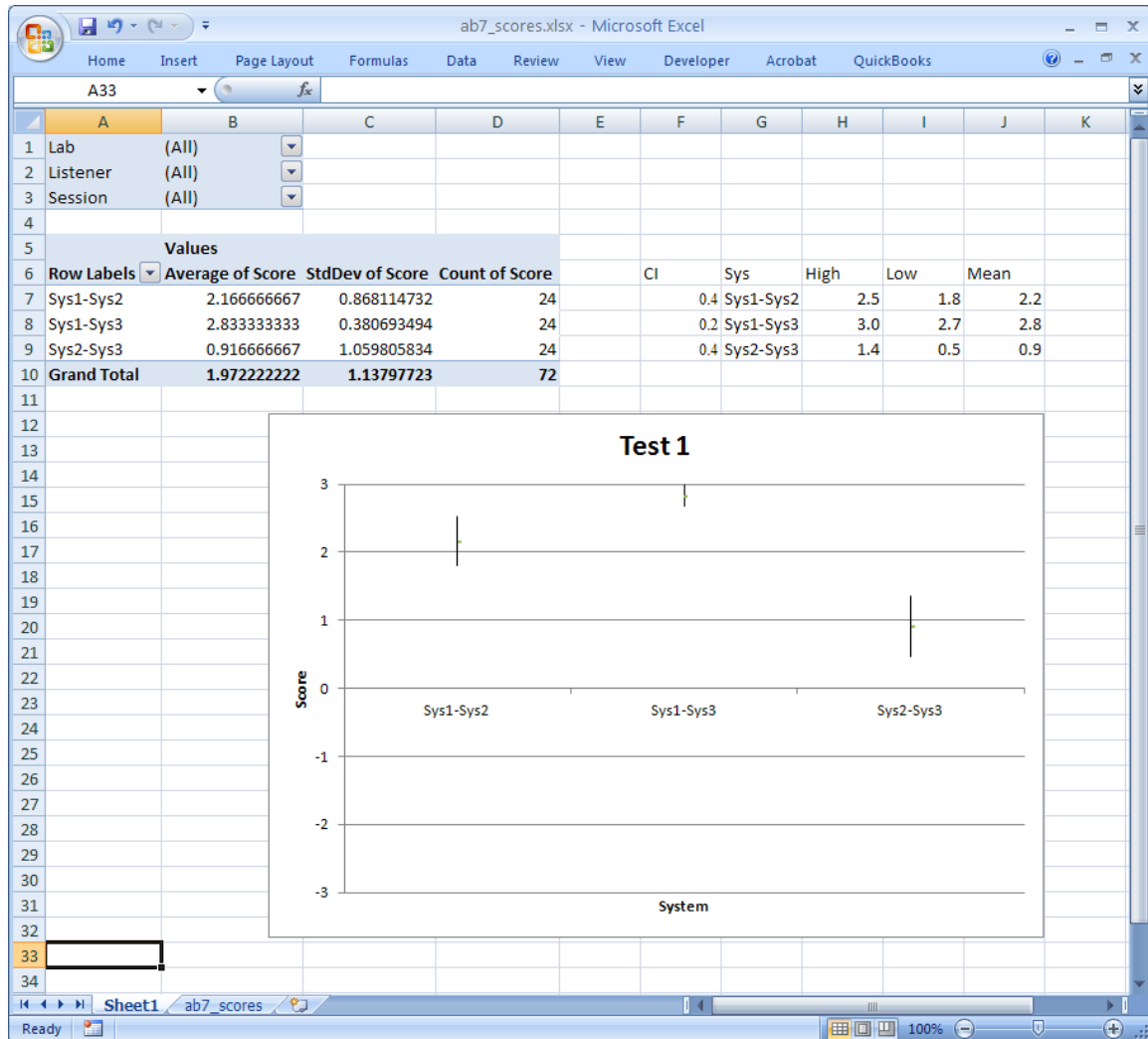
5.6   Comparative: ABX

The ABX forced choice comparison is unique amongst the test methodologies in that the outcome is 1 (correct identification) or 0 (incorrect identification). Any statistic whose outcome is binary has the Binomial distribution which, as we will see, has implications when computing the 95% confidence interval on the mean scores. An example of a statistic with a Binomial distribution is the outcome of a coin toss where $P(Heads) = \theta$ and $P(Tails) = 1 - \theta$. Note that if $\theta \neq 0.5$ the coin is not "fair."

37

5.6.1 Session file

A portion of an example Session file for the ABX Comparison forced choice test methodology is shown below. Note that there are only two systems lines per signal block. STEP will randomly assign them to Condition buttons A and B and then randomly assign one of them to Condition button X.

Line in Session file		Description
session=ABX		Session type
# Sig1		Displayed signal name
sig1_sys1.wav	Sys1	System under test Sys1
sig1_sys2.wav	Sys2	System under test Sys2
# Sig1		Displayed signal name
sig1_sys1.wav	Sys1	System under test Sys1
sig1_sys3.wav	Sys3	System under test Sys3
# Sig1		Displayed signal name
sig1_sys2.wav	Sys2	System under test Sys2
sig1_sys3.wav	Sys3	System under test Sys3
# Sig2		Displayed signal name
sig2_sys1.wav	Sys1	System under test Sys1
sig2_sys2.wav	Sys2	System under test Sys2
# Sig2		Displayed signal name
sig2_sys1.wav	Sys1	System under test Sys1
sig2_sys3.wav	Sys3	System under test Sys3

5.6.2 Score file

The set of Score files must be concatenated, retaining only the first header line in the file. However, no further editing is required. Note that the ABX score file shows the identity of the systems assigned to the "A" and "B" buttons (i.e. in the columns SysA and SysB). The first portion of such a Score file is shown here:

Lab	Listener	Session	FileA	FileB	Sig	SysA	SysB	X	Score
lab1	lis1	abx	sig1_sys1.wav	sig1_sys2.wav	Sig1	Sys1	Sys2	A	0
lab1	lis1	abx	sig1_sys1.wav	sig1_sys3.wav	Sig1	Sys1	Sys3	B	1
lab1	lis1	abx	sig1_sys2.wav	sig1_sys3.wav	Sig1	Sys2	Sys3	A	1
lab1	lis1	abx	sig2_sys1.wav	sig2_sys2.wav	Sig2	Sys1	Sys2	B	1
lab1	lis1	abx	sig2_sys1.wav	sig2_sys3.wav	Sig2	Sys1	Sys3	B	0
lab1	lis1	abx	sig2_sys2.wav	sig2_sys3.wav	Sig2	Sys2	Sys3	A	1

5.6.3 Excel data worksheet

The scores.txt file is opened with Excel and put in the "Scores" worksheet. A portion of the worksheet is shown here:

Lab	Listener	Session	FileA	FileB	Sig	SysA	SysB	X	Score
lab1	lis1	abx	sig1_sys1.wav	sig1_sys2.wav	Sig1	Sys1	Sys2	B	0
lab1	lis1	abx	sig1_sys1.wav	sig1_sys3.wav	Sig1	Sys1	Sys3	A	1
lab1	lis1	abx	sig1_sys2.wav	sig1_sys3.wav	Sig1	Sys2	Sys3	B	1
lab1	lis1	abx	sig2_sys1.wav	sig2_sys2.wav	Sig2	Sys1	Sys2	B	1
lab1	lis1	abx	sig2_sys1.wav	sig2_sys3.wav	Sig2	Sys1	Sys3	A	0
lab1	lis1	abx	sig2_sys2.wav	sig2_sys3.wav	Sig2	Sys2	Sys3	B	1

5.6.4 Excel Pivot Table

Since a score is associated with a comparison, an additional column must be added to the Scores worksheet. Create a new column (K) with the label Sys. In cell K2 insert the formula

Audio Research Labs

=IF(G2<H2,CONCATENATE(G2,"-",H2),CONCATENATE(H2,"-",G2))

and copy this formulas through to the last row of score data. Since column data SysA and SysB show the identity of the A and B buttons (rather than systematic order as with other tests), the Excel formula must sort the identity to obtain the systematic ordering. Now the new column shows both systems associated with Condition buttons A and B in sorted order, which will be used as a label in the analysis chart.

To construct the Pivot Table for the systems under test, proceed as in Section 6.2.4, but
- In the Layout part of the Pivot Table
  - drop Lab, Listener, Session and Sig in the Report Filter drop zone
  - drop Sys in the Row drop zone
  - drop Score in the Data drop zone (three times, setting to Average, StdDev and Count)
  - Ignore the FileA, FileB, SysA, SysB and X items

As stated above, scores of the ABX forced choice comparison have the Binomial distribution. Let θ be the fraction of correct identifications of X, which is also equal to the average of the score. Then the mean score is

$$\mu = \theta$$

We define the following variables

$$a = n + z_{\alpha/2}^2$$
$$b = -(2n\theta + z_{\alpha/2}^2)$$
$$c = n\theta^2$$

where n is the total number of responses (*Count of Score* in the Pivot Table) and $z_{0.05/2} = 1.96$.
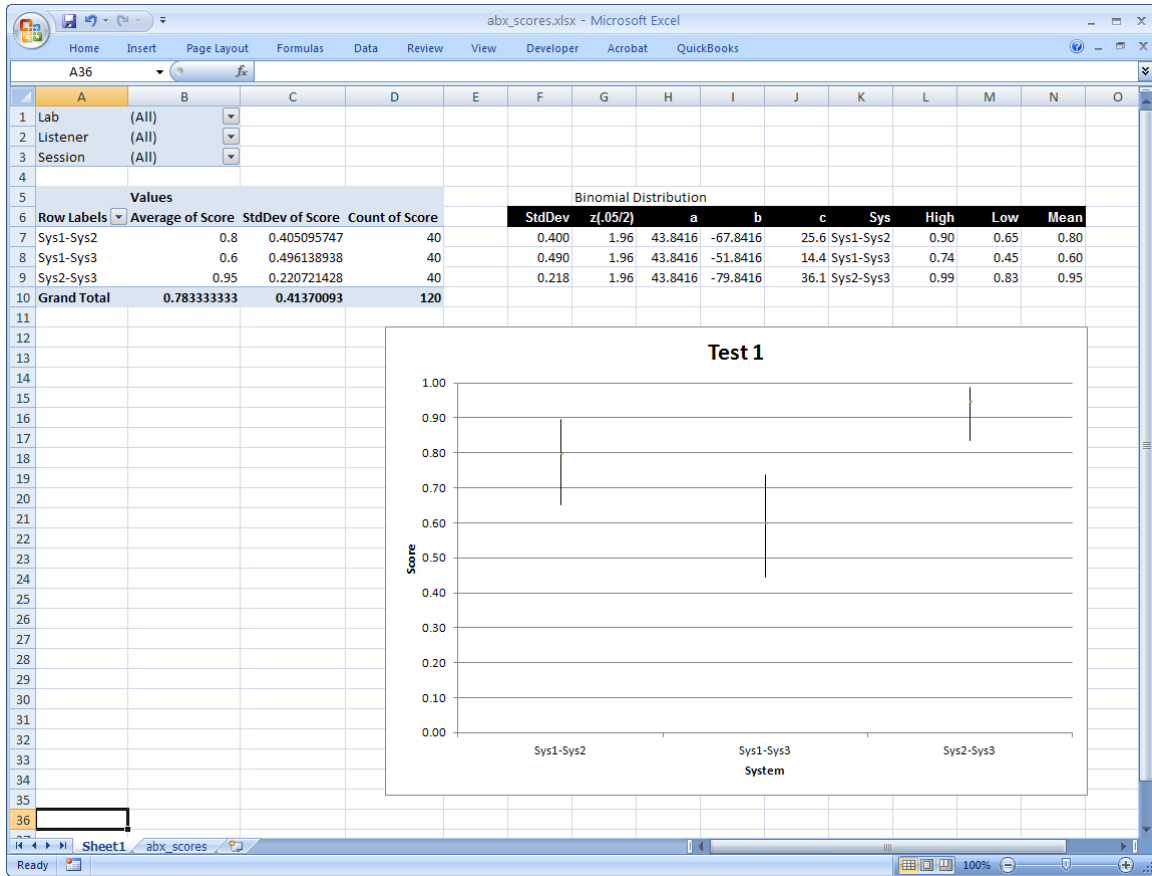
Then the 95% confidence interval on the mean is given by the two roots of the quadratic equation:

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

This is implemented in the example spreadsheet, which is shown below. Note that the standard deviation (StdDev) computed in the Pivot Table is *not* used in calculation of the 95% confidence interval, although it would yield quite close results. For comparison, the standard deviation (StdDev) for the Binomial distribution statistic is computed.

According to the analysis, Sys2 has a mean score that is different from Sys1 at the 95% level of significance (i.e. the 95% CI does not encompass the 50% point), while the mean scores of Sys3 and Sys1 are not different at the 95% level of significance (i.e. the 95% CI does not encompass the 50% point). Hence Sys3 achieves transparency, that is subjects are not able to identify X when Sys3 is one of A or B with a probability greater than random chance.

abx_scores.xlsx - Microsoft Excel

Row Labels	Average of Score	StdDev of Score	Count of Score		StdDev	z(.05/2)	a	b	c	Sys	High	Low	Mean
	Values						Binomial Distribution						
Sys1-Sys2	0.8	0.405095747	40		0.400	1.96	43.8416	-67.8416	25.6	Sys1-Sys2	0.90	0.65	0.80
Sys1-Sys3	0.6	0.496138938	40		0.490	1.96	43.8416	-51.8416	14.4	Sys1-Sys3	0.74	0.45	0.60
Sys2-Sys3	0.95	0.220721428	40		0.218	1.96	43.8416	-79.8416	36.1	Sys2-Sys3	0.99	0.83	0.95
Grand Total	0.783333333	0.41370093	120										

Lab (All), Listener (All), Session (All)

Test 1 — chart of Score vs System (Sys1-Sys2, Sys1-Sys3, Sys2-Sys3)

5.7 MOS

5.7.1 Session file

A portion of an example Session file for the Absolute Category Rating (ACR) or Mean Opinion Score (MOS) test methodology *with* Reference is shown below. The MOS test is possible with and without an open reference.

Open Reference Option
If the option of having an open reference and associated REF button is desired, then include the
`has_open_reference=true`
line in the session file (as shown below).

Line in Session file		Description
`session=MOS`		`Session type`
`has_open_reference=true`		`Include Open Reference`
`# Sig1`		`Displayed signal name`
`sig1_orig.wav`	`REF`	`Reference`
`sig1_sys1.wav`	`Sys1`	`"A" System`
`# Sig1`		`Displayed signal name`
`sig1_orig.wav`	`REF`	`Reference`
`sig1_sys2.wav`	`Sys2`	`"A" System`
`# Sig1`		`Displayed signal name`
`sig1_orig.wav`	`REF`	`Reference`
`sig1_sys3.wav`	`Sys3`	`"A" System`
`# Sig2`		`Displayed signal name`

Audio Research Labs

```
sig2_orig.wav     REF     Reference
sig2_sys1.wav     Sys1    "A" System
# Sig2                    Displayed signal name
sig2_orig.wav     REF     Reference
sig2_sys2.wav     Sys2    "A" System
# Sig2                    Displayed signal name
sig2_orig.wav     REF     Reference
sig2_sys3.wav     Sys3    "A" System
```

5.7.2 Score file

The set of Score files must be concatenated, retaining only the first header line in the file. However, no further editing is required. The first portion of the Score file is shown here:

```
Lab       Listener Session File           Sig     Sys       Score
lab1      lis1     mos     sig1_sys1.wav  Sig1    Sys1          5
lab1      lis1     mos     sig1_sys2.wav  Sig1    Sys2          4
lab1      lis1     mos     sig1_sys3.wav  Sig1    Sys3          3
lab1      lis1     mos     sig2_sys1.wav  Sig2    Sys1          5
lab1      lis1     mos     sig2_sys2.wav  Sig2    Sys2          4
lab1      lis1     mos     sig2_sys3.wav  Sig2    Sys3          3
```

5.7.3 Excel data worksheet

The scores.txt file is opened with Excel and put in the "Scores" worksheet. A portion of the worksheet is shown here:

Lab	Listener	Session	File	Sig	Sys	Score
lab1	lis1	mos	sig1_sys1.wav	Sig1	Sys1	5
lab1	lis1	mos	sig1_sys2.wav	Sig1	Sys2	4
lab1	lis1	mos	sig1_sys3.wav	Sig1	Sys3	3
lab1	lis1	mos	sig2_sys1.wav	Sig2	Sys1	5
lab1	lis1	mos	sig2_sys2.wav	Sig2	Sys2	4
lab1	lis1	mos	sig2_sys3.wav	Sig2	Sys3	3

5.7.4 Excel Pivot Table

To construct the Pivot Table for the systems under test, proceed as in Section 6.2.4, but
- In the Layout part of the Pivot Table
  - drop Lab, Listener, Session and Sig in the Report Filter drop zone
  - drop Sys in the Row drop zone
  - drop Score in the Data drop zone (three times, setting to Average, StdDev and Count)

Construct the chart of system mean scores and 95% confidence intervals, but
- Click on the Y-axis of the chart, right-click and select "Format axis"
  - Set the Minimum to be 1 and Maximum to be 5
  - Set the Decimal places to be 0
- Set chart title and axis titles

This should result in the following Pivot Table and chart:
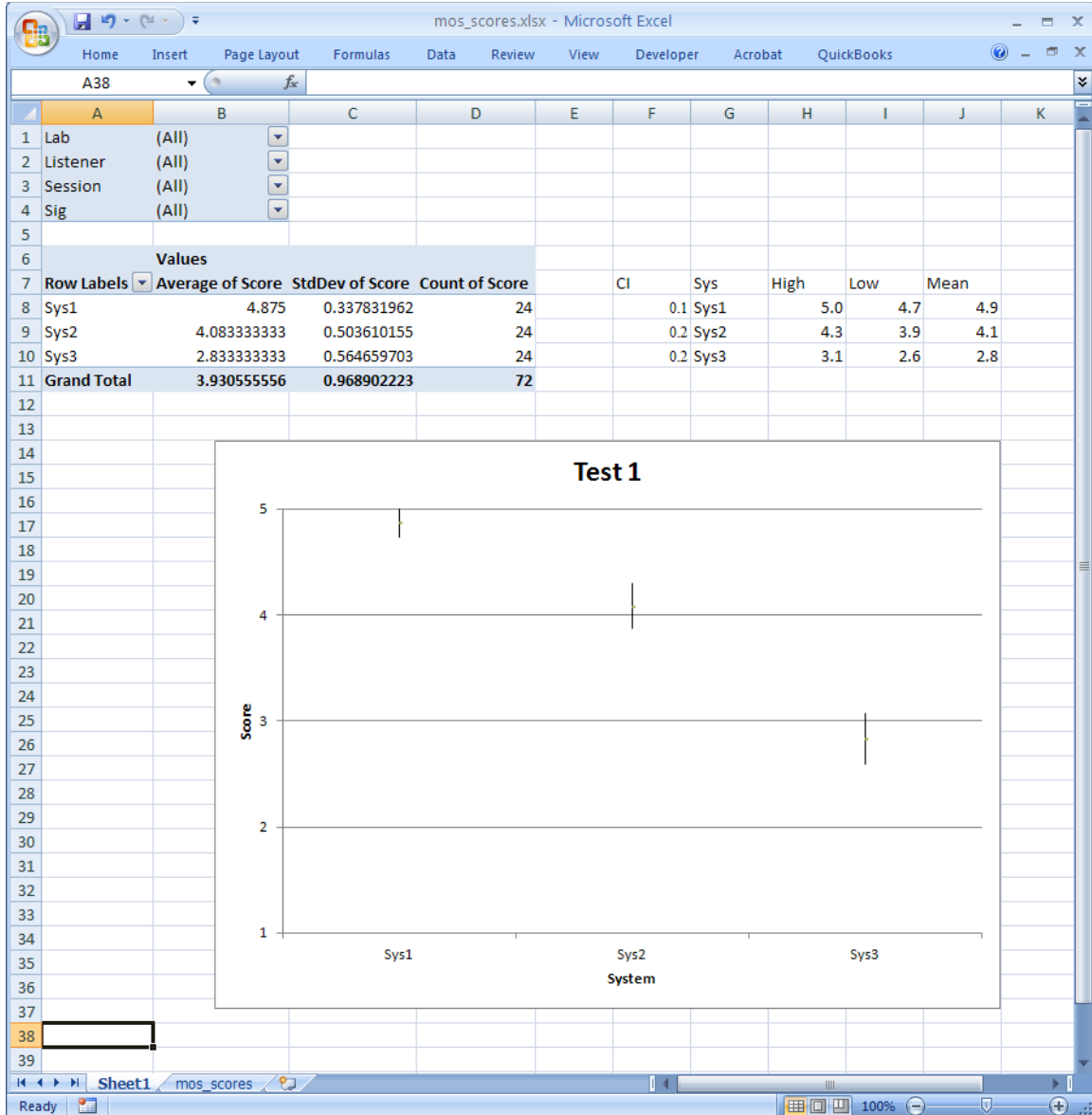
6 Bibliography

1. ITU-R Recommendation BS. 1116-3, "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems," Geneva, Switzerland, 2015. Available at http://www.itu.int/rec/R-REC-BS.1116/en.
2. ITU-R Recommendation BS.1534-2, "Method for the subjective assessment of intermediate quality level of coding systems: MUlti-Stimulus test with Hidden Reference and Anchor (MUSHRA)," Geneva, Switzerland, (06/20014). Available at http://www.itu.int/rec/R-REC-BS.1116/en.
3. ITU-R Recommendation BS.1284-1, "General Methods for the Subjective Assessment of Sound Quality." Geneva, Switzerland, 2003. Available at http://www.itu.int/rec/R-REC-BS.1116/en.
4. ITU-T Recommendation P.800, (08/96) "Methods for objective and subjective assessment of quality," Geneva, Switzerland, 1996. Available at http://www.itu.int/rec/T-REC-P.800-199608-I/en.

5. WAVE or RIFF WAVE sound file format: http://www-mmsp.ece.mcgill.ca/documents/AudioFormats/WAVE/WAVE.html
6. Steinberg ASIO developer's page: http://www.steinberg.net/en/company/developers.html

ANNEX I - Behavior of demo version of STEP

6.1 Start-up Behavior

Double-click on Session file `mushra.asi` and the program will start. First it presents the *Welcome* panel. The user must click the "OK" button to close this window so that STEP can present the *Administration* panel.

Welcome panel:
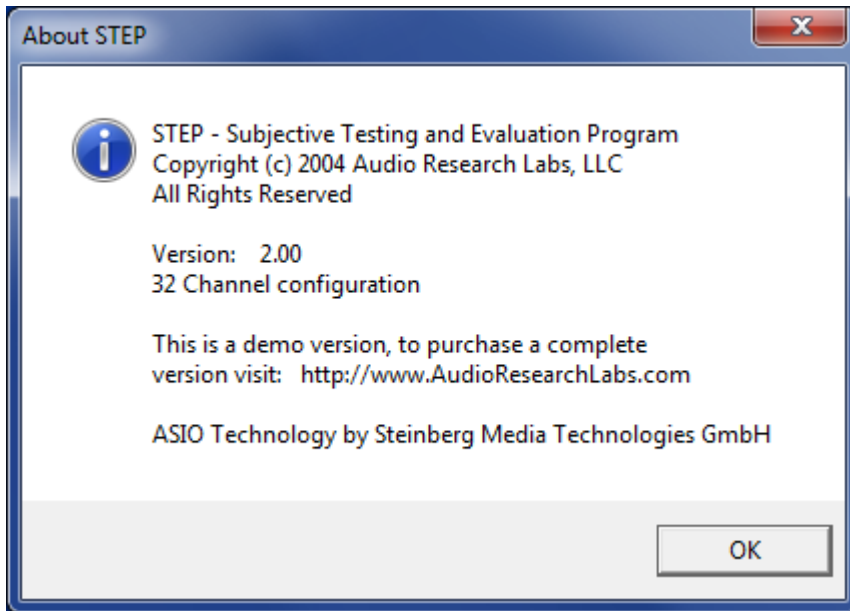
6.2 Limitations of Demonstration Version of STEP

The demonstration version of STEP shows all the capabilities of the program, but has the following limitations:
- It will present only the first 2 seconds of the signal files
- It will only write out scores of "0" in the score file