Accuracy – Can’t live with it Can’t live without it.

But, what does it really mean?

James D. Johnston
Chief Scientist, Immersion Networks

Why am I giving this talk?

- There is a lot of confusion about what “accurate” means.
- Some examples of accuracy:
  - Measures (via least mean squares) very nearly error free.
  - “Sounds Like” the original
  - Sounds like the listener wants (strictly speaking, that’s preference, more on that later!
  - Sounds like a “real place” and how do you “measure” that?
  - Measures like a “real place”, and how do you propose to measure that

- Some things that are not accuracy

  - “I like the sound of that (lp/tape/cd/download/whatever)”.
  - ‘xxx is the best compressor for broadcast’
  - CD’s always sound artificial
  - LP’s are always too noisy
  - Tube amplifiers sound <better/worse> than solid state amplifiers

  - The list goes on and on. There is no arguing preference. Just don’t confuse it with accuracy!

Accuracy vs. Preference vs. Perception

- Accuracy requires measurement
  - Sometimes electronic measurements suffice
  - Sometimes listen tests are required
  - Sometimes, it's hard to even establish exactly what "accurate" can be defined as.

- Preference
  - Is a personal thing. It does not extend beyond the individual holding the preference. Yes, a person can prefer 'accurate', now how do they define that?
  - Can affect what aspects of measurable accuracy are important
    - And can show that accuracy is not always preferred.
  - Can change with time, experience, or for no reason at all!

- Perception
  - Can be affected by many “outside” influences (we’re talking sound here, not anything else)
    - That’s not an invalid thing. If you prefer the “red wire”, so be it.
    - Must be understood in many cases to be affected by outside influences
    - Can be easily changed by events, circumstances, other people, et c.
  - Can be measured
    - But this is very often hard to do.
    - Must be set up for “quick switching without clicks, etc” in order to be sensitive
    - Must be blinded to all influences except the one(s) under test.
    - Can be used to measure preferences
    - Can be influenced by preference, of course, and almost always is.
    - For some things, perception is pretty much uniform across people (masking, audibility, and other basic parameters)
    - For other things, perception can vary wildly based on subject’s preferences (pre-echo sensitivity, stereo soundstage, distortion preference)

So, what's easy to measure?

Some devices, such as amplifiers, pre-amps, ADC's, DAC's, and the like, are intended not to modify the signal, merely adjust gain, capture, or amplify it. These devices can be measured electronically by comparing input and output. There are a few concerns:

- Do the test in situ, i.e. in the way they will be actually used. Use realistic inputs and output loads (not just resistors for an amplifier measurement, perhaps?)
- Use meaningful input stimuli.
  - For instance, measuring "THD" at 1kHz, well, it's kind of indicative but far from conclusive
- Don't be afraid to use complex stimuli, in fact, you probably should use complex stimuli, along with some good mathematical methods to disentangle those stimuli at measurement time.
- Give up on '1 number' kinds of things except in extreme cases.
  - 120dB SNR 20-20K at full power is, probably a decent '1 number'
  - 90dB SNR at 1kHz, -10dB from full output level – try again! It could be just fine, and it could be really, really bad.
- Ergo: Not that easy, but easier than anything else!

But, now, you didn’t say “how much is enough”.

- No, I did not. Earlier I gave an example of 120dB SNR at maximum power.
  - This is probably safe for perfect transparency anywhere in the real world for several reasons.
  - It is necessary, of course, to make sure that SNR does not decrease faster than level (i.e. center clipping, etc.)
- How much is enough really is a question for a given application. There are some limits, like the noise floor of the atmosphere, background noise, and equipment capability.

- In terms, however, of “accuracy” in this case, you have your answer. You have one or more numbers or a graph that tells you the actual accuracy of your system.
- You can, as well, have error spectra and the like calculated. The question of “how accurate” can be answered to outstanding precision even if you don’t like the answer.

Ok, what about transducers?

- For now, we'll stick to the loudspeaker end.
- Yes, it is entirely possible to measure the direct response of a loudspeaker, its distortion, and so on. BUT:
  - Most of us don't listen to loudspeakers in an anechoic setting.
  - Loudspeakers have radiation patterns. Off axis radiation is non-zero.
  - In order to even consider a measurement, you'll have to decide on the room you're working in as well as the loudspeaker, thanks to those room reflections of the off-axis radiation
  - Of course, the back of the room will also reflect the on-axis sound, as well.

- There are some things to consider there:
  - If the room reflection from off-axis radiation has a different coloration, that will affect how the speaker “sounds” to the listener, because that will reach the ears, just like the original direct sound (and we’re off into preference right there)
  - If the room has modes (especially at bass frequencies), those modes store room energy and let it build up, affecting the sound reaching the ears.
  - As Toole and others have shown, a well-controlled radiation pattern as well as direct pattern most often lead to a positive preference. This is “a good start”.
- Can we measure “accuracy” there, like we can from an amplifier output?
  - Well, yes, you can, for a given point in a room, get a 4-valued response for the output of the speaker at that point.
  - Interpretation of that is well-defined as far as mathematics.
  - Such information can be more difficult to relate to perception and preference.

4-valued? Huh? What's that about?

- Well, going back to basic acoustics, at any one point in the air, there are 4 variables, the pressure and 3 volume velocity variables, showing which way, and how fast the instantaneous air movement (remember, you have to move air to create pressure!) is happening. It's not just pressure, even though the eardrum reacts to pressure.
- If we have many closely spaced such measurements, of course, there will be a relationship between those measurements.
- In general, in places where the velocity is highest, the pressure is lowest, and vice versa.
- So, the eardrum response to pressure right? So who cares about all that pesky other stuff?
  - Yeah. Nice try. Your head creates pressure from volume velocity. What pressure it creates depends on the direction and velocity. It has more effect at higher frequencies.
  - That's how HRTF's come about, plainly put.

So, how do we determine “accurate” in soundfield capture terms?

- It’s not that simple. A very, very simple method to capture one point in space has been around for quite some time in the form of the “B format”
- This captures only one point. Now, that’s pretty dandy in a free field, of course but most of us do not live, work, or listen in an anechoic chamber.
- In a real acoustic setting, large or small, live or reasonably dead, moving the microphone a little bit will give you quite a surprisingly different measurement of the “original”.

- Which one is right? Well, probably neither, you’ll have to spatially sample the space if you’re going to do full-on analytic capture inside a given bandwidth.
  - At least you can probably stick to pressure if there is no reflection or absorption inside a sphere of some appropriate size. (You can calculate the rest using the wave equation).
  - So, how many channels is that? Figure a microphone every 1.5 to 2 CM, on the facets of a geodesic, give or take, to qualify to actually capture the soundfield up to 20kHz.
  - Now, do you think that’s even a stationary measurement? That means no currents, heat eddies, wall movements, and so on.

So, the point?

If you want to MEASURE ACCURACY then you must know what the target is.

There has to be something to measure against.

Now then, “Sounds like the original”

- That’s harder. The only way to decide that is via a subjective test. Now, we’ve all seen various rhetoric about how this test or that test does this or that, **HOWEVER,** there is a gold standard for how to determine this, and it’s called the double-blind ABX test.
  - These are hard to set up.
  - They require controls, level observation, and many other things
  - They don’t tell you “how different”.
- For “how different” there’s the ABC/HR test.
  - It will also tell you “how different” when things are different.

- This not a talk on how to do subjective tests, in fact, that’s not a lecture, that’s a semester class for those already familiar with the general idea.

Of course, there’s the same question:

You need to be able define the “right answer”
Again, sometimes this is easy.

Sometimes, it’s well-nigh impossible.

Sounds like: When you can define the right answer.

- Examples would be things like codecs, MP3, AAC, AC3, and newer ones.
  - You have the original PCM input.
  - You have the PCM output.
  - So, you compare them in a DBT of some sort.

- That will get you an answer, for the individual taking the test, to some probability of correct conclusion.

- If you do enough tests, on enough people, for long enough, with enough different kinds of material, while avoiding listener fatigue and burnout, you might even be able to come to a tentative conclusion or two about the universality of such a device. Maybe.

Note: DBT means “double blind test.

So, then, we get to “Does this sound like the original performance in the original hall”

- Yes, you have an “original” in the actual performance.
- You can also present the thing you’re testing to the listener.

- There’s just a few problems:
  - Nobody, and I mean NOBODY can capture the entire soundfield around your head, and yes, heads move during a concert.
  - Even if you could capture it, and then render it properly, the detailed memory of the original is gone. The fine details of an acoustic memory are gone in 200 milliseconds, give or take.

  - Ok, how do you swap between original and playback?

  - And, of course, you can not “loop” the orchestra/band/performer. Each performance will be slightly different.

That’s not totally impossible, for instance:

- One can test, to some limits, localization. Is there directionality that’s “something like the original”.
- For some special situations, you can even test the accuracy of things like distance, direction, and elevation. This may even matter sometimes (say if you’re using auditory feedback to someone in a very real-time situation).
- Some have proposed testing “original” by using a “dummy head” in the original soundfield, and using that, in headphones, for the “original”.
  - Whose head do you model?
  - How do you predict the head motion of the listener?
  - So, yeah, you can do some things with this kind of test, but it’s still very limited.

Of course, there is more complexity

- In headphone virtualization, for instance, getting front externalization can be hard without customization for each listener.
  - On the other hand, capturing head motion can help a generic HRTF externalize much better.
    - Subjects can, and will, learn on the fly.
    - They'll even learn if they're just listening, and it's not a test.
  - Better have that latency very low.
    - Latency can create a drastic case of motion sickness.
    - Motion sickness generally causes the individual to reject the entire situation.
- In such settings, one can also learn the intended localization.
- Remember, when you change your shirt/blouse/hat/hair/glasses, you change your HRTF, and you adapt JUST FINE.

Where does this leave us?
Let’s reproduce something the audience
PREFERS.

- Like I said earlier, one can test preferences.
- Preferences can change, including during the test (been there seen that)
- Preferences may differ for different sources (say, precise imaging for a 4 piece quartet, vs. “wall of sound” for a metal band).

- I think that Ricky Nelson said it best:

  You can’t please everyone, so you’ve got to please yourself!

Just to remind everyone

- You can in fact test accuracy very, very well
  - If you can define “the right answer”
  - If you can use objective measurement, for example least-mean-squares
  - If you can relate that to the needs of the situation.
- In other words, yes, you can very well measure an ADC/DAC/amplifier/preamp, wire, or other such stuff that is intended to not make changes.

- Beyond that, testing changes that are intended to be made, but not be audible, can be tested, with some difficulty.
- Beyond that, in realistic situations, what you care about is preference.