Pulling all of this together is an important study [6] in which the opinions of 12 selected and trained listeners are compared to those of 256 listeners from various backgrounds. The relative ratings of the products were essentially the same for small groups of listeners extracted from each population. The consequential difference was in the statistical confidence one could place in the opinions. The selected and trained listeners were much more reliable in their ratings, meaning that trustworthy results could be obtained in much less time. The trained listeners also provided comments that were easily interpreted by design engineers to help them focus on aspects of performance that needed working on, while other listeners tended to use less technically descriptive terms.

BLIND vs. SIGHTED TESTS – SEEING IS BELIEVING

When you know what you are listening to, there is a chance that your opinions might not be completely unbiased. In scientific tests of many kinds, and even in wine tasting, considerable care is taken to ensure the anonymity of the devices or substances being subjectively evaluated. Many people in audio follow the same principle, but others persist in the belief that, in sighted tests, they can ignore such factors as price, size, brand, etc. and arrive at the unbiased truth. In some of the "great debate" issues, like amplifiers, wires, and the like, there are assertions that disguising the product identity *prevents* listeners from hearing small differences. "Proof" of this is the observation that perceived characteristics that seemed to be obvious when the product identities were known, are either less obvious or non-existent when the products are hidden from view. The truth is not always what we wish it to be.

In the category of loudspeakers and rooms, however, there is no doubt that differences exist and are clearly audible. To satisfy ourselves that the additional rigor was necessary, we tested the ability of some of our trusted listeners to maintain objectivity in the face of visible information about the products.

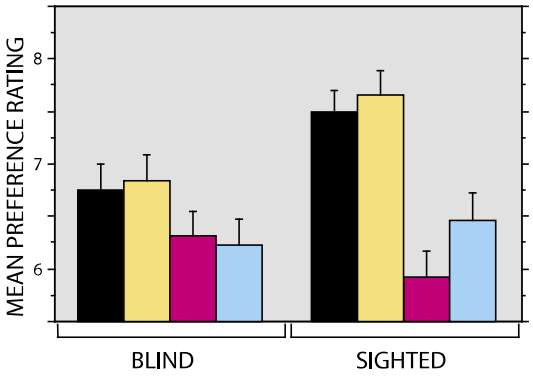

Figure 4. A comparison of blind and sighted evaluations of the same four loudspeakers by the same group of listeners. From Ref. 7.

The results are very clear. Figure 4 shows that, in subjective ratings of four loudspeakers, the differences in ratings caused by knowledge of the products is as large or larger than those attributable to the differences in sound alone. The two left-hand bars are scores for loudspeakers that were large, expensive and impressive looking, the third bar is the score for a well-designed, small, inexpensive, plastic sub/sat system. The right-hand bar represents a moderately expensive product from a competitor that had been highly rated by respected reviewers.

When listeners entered the room for the sighted tests, their positive verbal reactions to the big speakers and the jeers for the tiny sub/sat system foreshadowed dramatic ratings shifts – in opposite directions. The handsome competitor's system got a higher rating; so much for employee loyalty.

Other variables were also tested, and the results indicated that, in the sighted tests, listeners substantially ignored large differences in sound quality attributable to loudspeaker position in the listening room and to program material. In other words, knowledge of the product identity was at least as important a factor in the tests as the principal acoustical factors. Incidentally, many of these listeners were very experienced and, some of them thought, able to ignore the visually-stimulated biases [7].

At this point, it is correct to say that, with adequate experimental controls, we are no longer conducting "listening tests", we are performing "subjective measurements".