Perspective: real ear verification of hearing aid gain and output

Abstract

It is the role of the audiologist to ensure that hearing aids are programmed and fitted to optimize benefit. Research has shown that haphazard fittings lead to reduced performance for the hearing aid user. This paper reviews the evidence supporting the use of validated prescriptive methods such as the NAL-NL2. The use of prescriptive methods includes ensuring that the fitting targets are met relative to ear canal SPL. This verification only can be made using probe-microphone measures; current techniques and procedures for this verification are discussed.

Keywords: probe microphone measures, real ear measures, hearing aid verification, prescriptive hearing aid fittings, NAL-NL2

H. Gustav Mueller[1]

1 Vanderbilt University School of Medicine, Nashville, United States

Introductory remarks by the editor

We are using a somewhat different format for this special Review Paper. H. Gustav Mueller, PhD, has been a leading advocate for the real-ear verification of hearing aid performance for over 40 years, and has published numerous articles and book chapters on the topic. He authored the first book on real-ear verification back in 1992, and recently, a second textbook, Speech Mapping and Probe Microphone Measures. Dr. Mueller also is known as the editor of the popular column "20 Questions" that appears each month at AudiologyOnline. For this invited Review Paper, we're going to turn things around and ask Dr. Mueller the questions.

Interview

ZAUD: We hear that you might be one of the pioneers of the real-ear verification of hearing aid performance. True?
Mueller: I've been at it for a long time, so I guess that does make me a pioneer. We started looking at the practicability of these measures in 1979 when I was at Walter Reed Medical Center, Washington D.C. In those early days, we actually placed a small hearing aid microphone down in the ear canal. It was a procedural nightmare, but we were able to obtain some meaningful measures, and were excited about the potential of this procedure. At the time, we were using pure-tone aided testing in the sound booth to verify our prescriptive fittings. We were well aware of all the negative issues and pitfalls surrounding these measures, and were anxious to abandon them. We presented a paper on our early experiences with real-ear measures at the 1980 conference of the American Speech and Hearing Association [1] – that was 40 years ago, so this certainly isn't something new.

ZAUD: This was before we used probe tubes to assist in the ear canal measure?
Mueller: Yes, the "probe-tube" version wasn't introduced until 1983 or so (the Rastronics CCI-10), and was not really commercially available for another year (at least in the U.S.). The probe-tube approach was a life saver – throwing away a plugged tube was a minor thing compared to the gummed-up microphones we had in the past. By 1985 we had all our fitting rooms at Walter Reed equipped for probe-mic verification, and we were off and running. We of course expected this to soon become the best practice standard for fitting hearing aids, and commonly used by all audiologists. After all, why wouldn't you want to know the SPL at the ear drum?

ZAUD: You say "the standard." But that never happened?
Mueller: No – I guess that's partly why I'm here with you today. I can't speak for other countries, but in the U.S., my best guess is that no more than 30–40% of audiologists who fit hearing aids conduct probe-mic verification routinely, and that hasn't really changed since the 1990s.

ZAUD: Why do you think there is a reluctance to use this verification tool?
Mueller: It's a combination of several factors. Some say that they simply don't have the time, a weak excuse I believe. One issue is that I don't think the concept of verification is well understood. To verify something, we start with a set of standards to verify against – in the world of fitting hearing aids that would be an evidence-based validated prescriptive method. On social media, I often read long discussions among audiologists regarding whether or not to do REM (a popular term for probe-microphone measures). In these online discussions, audiologists talk about "REM" as if it were a way to fit hearing aids. It isn't. It's simply a verification of the "best known way" to fit hearing aids. I know clinics where the audiologists use the manufacturers' default fitting, conduct

probe-measures but do not change the programming, and then tell their colleagues that they "fit by probe" (whatever that means). That isn't the way it works. We have to buy into the fact that an evidence-based standard is the starting point, and go from there.

A second factor is that audiologists often are encouraged by manufacturers to use the manufacturer's proprietary fitting. They sometimes are told that certain hearing aid features do not work correctly unless the manufacturer's first-fit is selected in the fitting software. Audiologists tell me that they follow this guidance. There are no ear-canal targets for the manufacturer's fitting available on probe-microphone equipment, so it's impossible to do real-ear verification. Recently, some manufacturers, through the use of autoREMfit, have made it possible to use real-ear data to fit to their proprietary targets. The problem of course remains, that these targets have not been validated.

A third factor is that some audiologists believe that combining their clinical experience with comments from the wearer will provide a fitting more optimal than that of a prescriptive fitting approach. Denis Byrne, was the developer of the original National Acoustic Laboratories (NAL) prescriptive fitting approach, dating back to 1976 [2]. He passed away in 2000, and a year later Harvey Dillon gave the Denis Byrne Memorial Lecture at the annual meeting of the American Speech and Hearing Association. Harvey paraphrased Denis's thoughts on relying on clinical experience to fit hearing aids as follows [3]:

- If you can't write down the rules you use, you probably don't understand what you do.
- If it's not written down, no one else can do it, and no one can test whether it's better or worse than some alternative approach.
- If you can't evaluate your procedure you can't improve it.

Another important point is that when fitting a hearing aid, you have to start someplace. Why not with a validated approach? In his 2012 article, Earl Johnson [4] reviewed the problem of going rogue when selecting the best frequency response for a new hearing aid user. He suggests that an experienced clinician can rule out a large number of possible frequency responses, so we can assume that the optimal frequency response falls within a 20 dB range in each of the 16 channels of a typical modern hearing. We also know that there is a need for a somewhat "smooth" response across the side-by-side channels of adjacent frequencies – we wouldn't put 20 dB of gain in one channel, and 0 dB of gain in the adjacent channel, nor is it even possible due to overlapping channels. When we construct somewhat smooth frequency responses, we eliminate 99% of the available 16 channel, 20 dB range frequency response choices. We then apply further logic, only selecting frequency responses that in theory could simultaneously provide the best speech intelligibility, acceptable loudness and sound quality. After all of which, there are still *1,430* possible frequency responses from which to choose for any particular hearing aid user!

And that is only for one input level. Sounds like the fitting process is going to require more than one office visit.

ZAUD: You mention the need for verification of the "best known way." But do we know that there really is a best way?

Mueller: Well, we certainly know what *isn't best*, and that is what is commonly used, and I'll be happy to talk about that later. There probably are several "equally-best" ways. There are three or four prescriptive fitting methods that have been rigorously validated. Let's talk about the NAL approach, simply because it's been around for the longest, is the most researched, is used around the world, and for adults, the current version is very similar to the other methods available. It started with the original NAL [2], which then led to the NAL-R [5], followed by the NAL-NL1 [6] and we now have the NAL-NL2 [7]. I published an evidenced-based review of the earlier NAL methods in 2005 [8].

One method to evaluate the appropriateness of a given prescriptive fitting is to fit individuals accordingly, provide them with highly trainable hearing aids, and then allow them to adjust the products to what they prefer based on real-world use. We have these types of studies for the NAL prescriptive method. Ben Hornsby and I were curious if previous experience with a given hearing aid fitting would influence preferred gain with new instruments [9]. We often hear that hearing aid users want new hearing aids that sound like their old ones. We specifically selected participants (n=20; all bilateral wearers) who had used their current hearing aids for at least two years, and who we knew had been fitted to a specific manufacturer's proprietary fitting, which tended to provide gain substantially below that of the NAL-NL1. We fitted these individuals bilaterally with trainable hearing aids (e.g., input-specific gain training, and a treble adjustment) to the NAL-NL1 prescriptive method. The participants used the hearing aids in the real world for two weeks. They had a diary to complete, which included a variety of assigned listening situations that potentially would encourage gain adjustments (note: on follow-up, data logging showed that all participants had at least 130 gain adjustments during the trial period).

The results are shown in Figure 1. Displayed are the mean NAL-NL1 targets, and the mean values for the REARs for the hearing aid user's present instruments, the original programmed output, and the trained output, for both low and high frequency bands. As predicted, the participants had been fitted substantially below NAL-NL1 targets – nearly 10 dB for the 55 dB SPL input. Observe, however, that following training, they did not train down to what they had been using, but rather, used significantly more gain (p<.001), only 2–3 dB below NAL-NL1 targets. It may simply be coincidental, but these were NAL-NL1 targets, and the NAL-NL2 targets are roughly 3 dB lower.

The Mueller and Hornsby [9] study was with experienced hearing aid users. Perhaps even more compelling data is from a study using trainable hearing aids conducted by Catherine Palmer [10]. The participants in this study
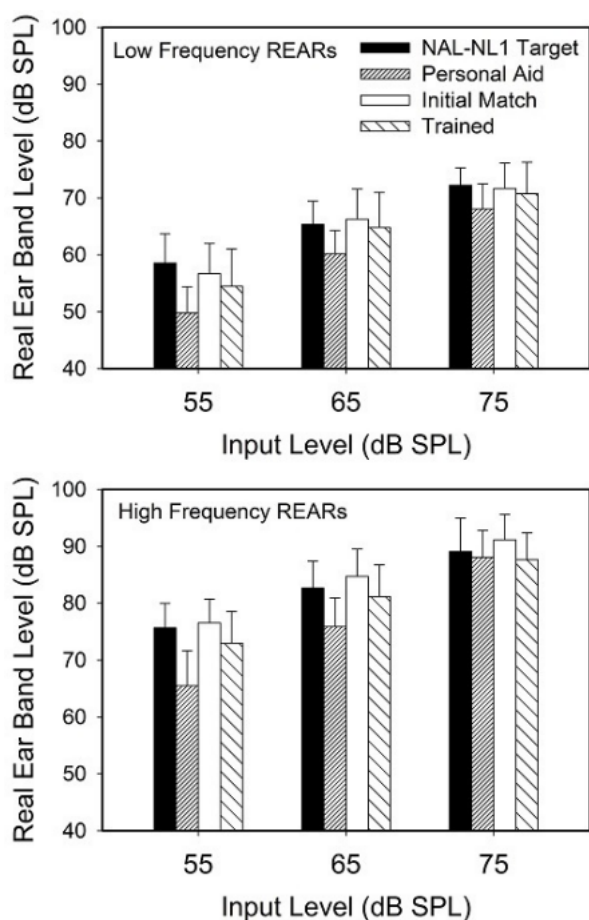
Figure 1: Mean real ear aided response (REAR; ear canal SPL) shown for the averaged low-band and high-band frequencies for the real- speech inputs for three levels. Data shown for four conditions: NAL-NL1 target, NAL-NL1 best match to target, trained gain, and average gain for the patient's personal hearing aids.

were 36 new users of hearing aids. One group of 18 was fitted to the NAL-NL1, used this gain prescription for a month, and then trained the hearing aids for the following month. The second group of 18 was also fitted to the NAL-NL1, but started training immediately, and trained for two months. Importantly, these individuals were using hearing aids that had input-specific training, and had the potential to be trained up or down by 16 dB – providing ample opportunity for them to zero in on their preferred loudness levels. In general, after two months of hearing aid use, both groups ended up very close (within 1–2 dB) to the NAL-NL1 targets for average inputs. Palmer reports that the Speech Intelligibility Index (SII) for soft speech was reduced 2% for the first group, and 4% for the group that started training at the initial fitting. Again, this was with NAL-NL1, not the current NAL-NL2.

The NAL-NL2 prescriptive method was evaluated in a trainable hearing study by Keidser and Alamudi [11]. In this research, 26 hearing-impaired individuals (experienced hearing aid users) were fitted with trainable hearing aids, which were initially programmed to NAL-NL2. Following three weeks of training, the authors examined the new trained settings for both low and high frequencies, for six different listening situations. That is, the training

was situation specific based on the hearing aid's classification system; a given participant could train increased gain for music, and decreased gain for speech-in-noise. The participants did tend to train down from the NAL-NL2 for all six situations, but only by a minimal amount. For example, for the speech in quiet condition for the high frequencies, the average value was a gain reduction of 1.5 dB (0.95 range = 0 to –4 dB), and for the speech in noise condition, there was an average gain reduction of only 2 dB (0.95 range =+0 .5 to –4.5 dB). The trained gain for the low-frequency sounds for these listening conditions was even closer to the original NAL-NL2 settings.

These studies all suggest that on average, the NAL prescription is a reasonable starting point. A skeptic, however, might point out that in all three studies, the starting point was the NAL prescription, which could have influenced the ending point [12]. Let's then look at a recent study from Sabin et al. [13]. These authors evaluated the outcomes of self-fitting hearing aids that were initially set to 0 dB REIG, so the starting point was not biased toward any fitting rationale. For later reference, the hearing aids were programmed to a real-ear verified NAL-NL2. The real-world performance of the self-fitting approach (n=38) was evaluated via a month-long field trial. There was a strong correlation between user-selected and audiologist-programmed gain (r=0.66, p<.0001). On average, the user-selected gains were only 1.8 dB lower than those selected by the audiologists based on the NAL-NL2 prescription.

These studies all have used the NAL prescription as the reference, but it seems unlikely that the findings for DSLv5 would be much different, simply because for adults, this prescription method is very similar to the NAL-NL2. Johnson and Dillon [14] compared these two methods for five different mild-to-moderate sensorineural hearing loss configurations. Rarely did prescribed gain for the key frequencies of 500–4,000 Hz differ by more than 3–4 dB, and when the SIIs for an average-level input were averaged for the five different configurations the difference between the two methods was 0.01 (DSL SII=0.70; NAL SII=0.69).

ZAUD: Given that many audiologists choose not to verify to the NAL-NL2 targets, what do you believe is their concern?

Mueller: In some cases, for some products, fitting to the NAL-NL2 will cause a feedback issue, but by far, what I hear the most is that NAL-NL2 prescribed targets provide more loudness than what the average wearer wants. This just doesn't match with the research evidence. I've already reviewed that when hearing aid users have the opportunity to train away from the NAL fitting, they don't. But we can go back to the research that led to the NL2 modification of the NL1 [15]. At the time, there were data that suggested that indeed NL1 called for slightly more gain than desired by the average user. For this reason, gain for average inputs for NL2 were lowered by about 3 dB. Based on the preferred loudness level data from

nearly 200 hearing aid users, it was shown that by lowering the gain by this amount, about 60% of individuals would fall within a ±3 dB window of the fitting target. So yes, we then would expect about 20% to say that an NAL-NL2 fitting was too loud, but also, 20% to say that the NAL fitting was too soft. From my experience, this seems about right.

Now, I can think of some reasons why clinicians might state that their NAL targets call for too much gain – two of them involve procedural issues:

- Prior to the testing of each individual, the probe-mic software must be set to correspond with the specifics of the fitting. The two most important factors are whether the fitting is bilateral or unilateral, and if the person being fitted is a new user, or experienced. I'll use an example from Johnson [4] to illustrate the importance of setting up the equipment correctly when using NAL-NL2. Our examples are a woman obtaining her first set of hearing aids, and an experienced male user, obtaining a new set of hearing aids. To make it easy, let's say that they have the same hearing loss; 20–30 dB HL in the low frequencies sloping down to 70 dB HL at 2 kHz and above. While their hearing loss is the same, the prescribed insertion gain at 2,000 Hz for the female for a 65 dB SPL speech input would be 16 dB, whereas for the male it would be 21 dB. If the equipment was not set up correctly, you could think that you are at the NAL-NL2 target (for the woman) when in fact you were 5 dB over target – possibly big enough difference to exceed preferred loudness and impact the success of the fitting. The greater the hearing loss, the bigger the differences will be.

- A second procedural issue concerns the equalization method used by the clinician. Most probe-mic systems default to concurrent equalization – that is, the reference microphone is active during the presentation of the test signal. This helps correct for minor head movement during the 10–12 seconds that the signal is presented. But, this equalization method cannot be used with open fittings. Consider that for nearly all probe-mic systems, the reference microphone is located at the ear lobe, just below the ear-canal opening. The amplified signal leaks out of the ear, is picked up by the reference microphone, and if it is louder than the input signal (which it usually is) this will prompt a reduction in the input signal. The audiologist might think that they are presenting a 65 dB SPL signal, when in fact it's only 60 dB SPL. This will likely generate an output that is below the 65 dB target, so the audiologist now increases gain by 5 dB, which causes 5 dB more to leak out of the ear, and the input signal goes down another 5 dB. It is very possible that the ear-canal output would appear to be at target, when in fact it's 10 dB or more over target. This usually is observed in the 2,000–3,500 Hz region, because of the residual resonance of the ear canal. This is why stored equalization, not concurrent needs to be used, even when the fitting is only partially open (see Mueller

et al. for review [16]). This then is a possible reason why an audiologist might report that his or her hearing aid users believe the NAL-NL2 targets are too loud – they are not fitting to the target. We talked about it back in 2006 [17], but it still seems to be a reoccurring problem.

- A third issue relates to the understanding of the prescriptive target. The target is not a "dot" on the fitting screen, but rather a range. At least one probe-mic manufacture has a ± vertical bar at the target for each frequency to remind us of this. How big is the range? As mentioned earlier, we would expect that about 60% of individuals would be okay falling within ±3 dB of the center of the target. At least two different fitting guidelines have used ±5 dB as acceptable (International Society of Audiology [18]; British Academy of Audiology [19]), and we know that 5 dB would probably not be more than two JNDs for a broadband signal [20], so a ±5 dB range would seem clinically acceptable. The point being, that if a given hearing aid user preferred 5 dB below the precise target, they are still *fitted to target*. It remains important, however, to maintain a smooth frequency response that more or less follows the precise prescriptive pattern – we would not want to be 5 dB over at 1,000 Hz, then 5 dB under at 2,000 Hz, and then back over at 3,000 Hz.

- The final issue is related to counseling. Yes, it is true that when we first program the hearing aids for a new user, the first thing we often hear is, "Wow, everything seems loud." But this does not mean that we immediately grab our mouse and start turning down gain. Rather, the follow-up comment from the audiologist would then be something like, "Yes, that is the expected perception, it should sound loud; you haven't been hearing these sounds for many years. You'll adjust to this after a few days of use." Of course, there are some cases where the new user simply will not accept an output level that is close to target, but most will experience at least some acclimatization to loudness after some listening exprience. For these individuals, therefore, it's usually possible to increase gain during post-fitting visits, or implement an automatic gain increase in the fitting software, so that in the end, audibility is acceptable for both the user and the practitioner.

ZAUD: You mentioned the common use of manufacturers' proprietary algorithms. How do they compare with the generic prescriptive methods?

Mueller: Let me first talk a little about why I think these fitting algorithms exist. I'll start with an example from the mid-1990s. I was serving as a consultant for a major hearing aid manufacturer, and the DSL[i/o] had just been introduced [21]. Our clinical audiology advisory team convinced the R & D folks that this should be the default fitting for WDRC instrument that was soon to be launched (WDRC was a big deal in those days). They bought off on it, and it was part of the fitting software. Within months,

the sales staff was inundated with complaints from the field that the new product was not well received, and sales were dismal. The report from the field was that the new product had too much gain, sounded "tinny" and was feedback prone. New software was soon introduced and the fitting screen now showed "DSL [i/o]*". The note for the asterisk simply stated that the DSL had been modified. In the background, overall gain was reduced, and gain above 2,000 Hz rolled off considerably (solving both the tinny and feedback issues). Sales increased immediately. The point of the story is that manufactures, to stay profitable, have to satisfy a wide range of fitting goals for dispensers around the world – some with PhDs and others with no college at all. Sometimes decisions are not based on science.
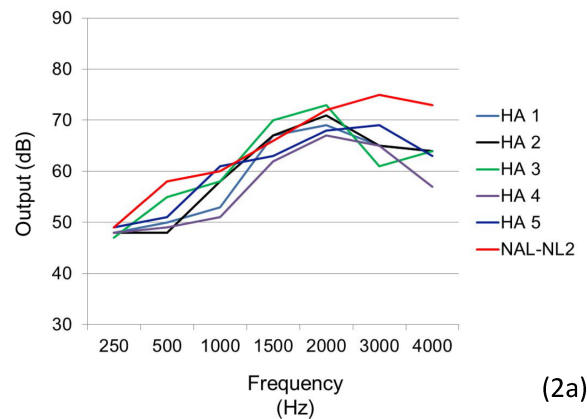
Many individuals fitting and dispensing hearing aids want a "click and go" solution. That is, one click on "First Fit" and the wearer is happy. What makes the typical new user happy on the day of the fit? Something that doesn't sound like a hearing aid, something that sounds "natural," and certainly, something that doesn't sound "tinny". It is therefore to no one's surprise, that propriety fittings under-fit (compared to generic methods), and in particular, roll off gain above 2,000 Hz.

In 2015, a group of us compared the proprietary fittings for the premier products of the five leading manufacturers [22]. Our mean results (REARs; 16 ears) are shown in Figure 2, for inputs of 55, 65 and 75 dB SPL. This was for a downward sloping hearing loss, going from 25 dB in the lows, to 70 dB in the highs. The NAL-NL2 fitting targets also are included for comparison purposes. Granted, the proprietary methods aren't geared to meet NAL-NL2 targets (you could simply use the NAL algorithm if they were), but this provides a reference.
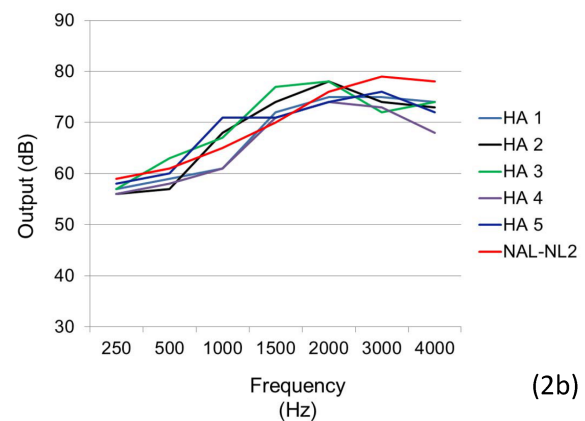
Notice that we do see a 5–8 dB difference among manu-facturers, but the pattern of the output for all the propri-etary fittings is similar, and considerably different than that of the NAL-NL2. For the higher frequencies, output above 2,000 Hz falls 10 dB or more below the NAL pre-scription for the 55 dB SPL input (a level just slightly be-low average speech [23]). This could be a holdover from the days when feedback reduction systems were not very effective. We see it today, however, even for moderate losses in the high frequencies – the high-frequency loss for the sample audiogram in this study was only 70 dB, a level where feedback would not be an issue for *most* modern hearing aids, even with an open fitting.

If you follow the mean output values (1,500–4,000 Hz range) for a given manufacturer for the 55 to the 75 dB input levels (20 dB input difference), you see a change in output of ~17–20 dB – in other words, these are es-sentially linear fittings. This helps explain why they under-fit for soft inputs, and over-fit for high intensity levels.
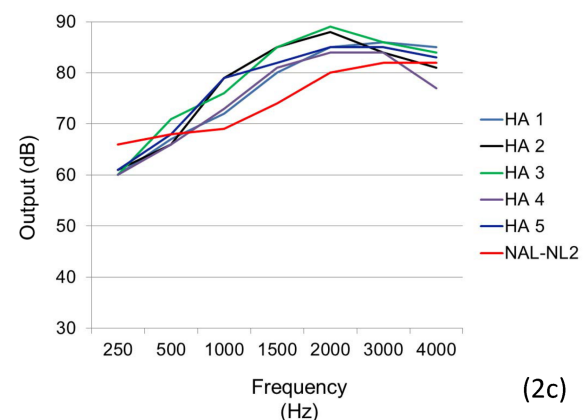
We also recorded the SII that was present for each parti-cipant. The group mean values for the three input levels compared to a NAL-NL2 fitting are shown in Figure 3. There is a sizeable difference between the SII of the pro-prietary fittings compared to the NAL-NL2. As the input goes up, the differences become smaller. The most con-

(2a)

(2b)

(2c)

Figure 2: Mean ear canal SPL output for the proprietary fitting (experienced user, bilateral fitting) of five leading manufacturers for a mild-to-moderate downward sloping hearing loss. Data shown for inputs of 55 dB (2a), 65 dB (2b), and 75 dB (2c) SPL. The NAL-NL2 targets are displayed as reference. Adapted from Sanders et al. [22].

servative fitting for the 55 dB input was HA-4, with an SII of only 0.25, compared to 0.47 for the NAL-NL2. For some listening situations, going from an SII of 0.25 to 0.47 can improve speech recognition substantially.

For the 75 dB input, the SIIs are similar to that of the NAL, but this value is misleading for real-world use (assuming the wearer has a method to lower gain). If we go back to Figure 2, note that for some instruments, the output in the mid frequencies (1,000 to 2,000 Hz) is about 10 dB greater than the NAL prescription. Given the amount of research by the NAL to determine preferred loudness
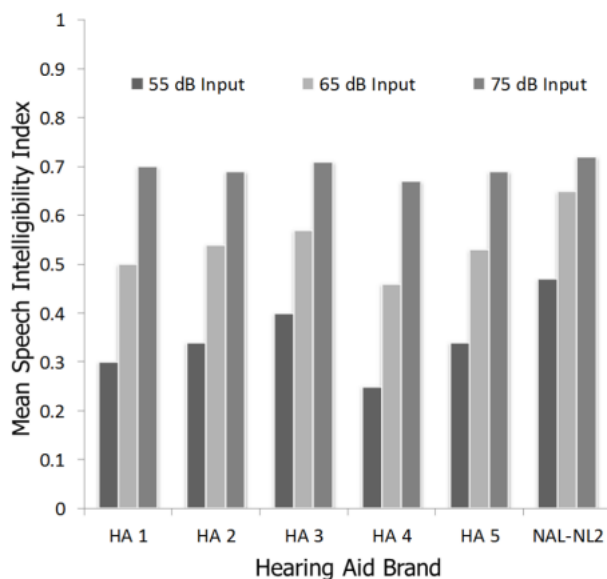
Figure 3: Mean SII values (as calculated by Audioscan Verifit) for the proprietary fitting (experienced user, bilateral fitting) of five leading manufacturers for a mild-to-moderate downward sloping hearing loss. Data shown are for inputs of 55, 65, and 75 dB SPL. The SIIs for a precise fitting to NAL-NL2 targets are shown as a reference. Adapted from Sanders et al. [22].

levels, it seems likely that a hearing aid user would find this too loud, and reduce gain. This would then lower the SII for this input level, and of course would make the SII for soft speech even worse than it already is.

The reduced audibility for the manufacturer's default fitting was illustrated more recently by Valente et al. [24]. These researchers reported that for soft speech, the mean gain for the proprietary fitting fell 15 dB below NAL-NL2 targets for 3,000 Hz and 21 dB for 4,000 Hz. For average level speech, the mean differences were 9 dB and 13 dB respectively.

ZAUD: This would seem to have an effect on speech recognition.

Mueller: I think even a beginning student of audiology would predict that this minimal audibility would reduce speech understanding for individuals fitted in this manner. They would be correct. Ron Leavitt and Carol Flexer [25] fit hearing-impaired individuals (typical downward sloping losses) who were experienced hearing aid users with seven different pairs of hearing aids and conducted aided QuickSIN testing [26]. The QuickSIN sentences were presented at 57 dB SPL (roughly average speech [23]). Six of the seven pairs of hearing aids were the premier models from the leading hearing aid companies. Special features such as directional microphone technology and noise reduction were activated. The seventh pair were 10-year-old analog, single-channel, omnidirectional hearing aids with no noise reduction features. Each of the six premier hearing aids were first evaluated while programmed to the manufacturer's first fit, and then also when programmed to the NAL-NL1. The old analog hearing aids were only tested programmed to the NAL-NL1.

The results of this study are plotted in Figure 4, which are the mean QuickSIN scores for the participants for all the aided conditions; the QuickSIN is scored in "SNR Loss" – the difference between the SRT-50 value for a given individual and that of the QuickSIN norms for individuals with normal hearing. In Figure 4, a –10 dB SNR would indicate that mean performance is 10 dB worse than expected for normal hearing individuals (in other words, down is bad).
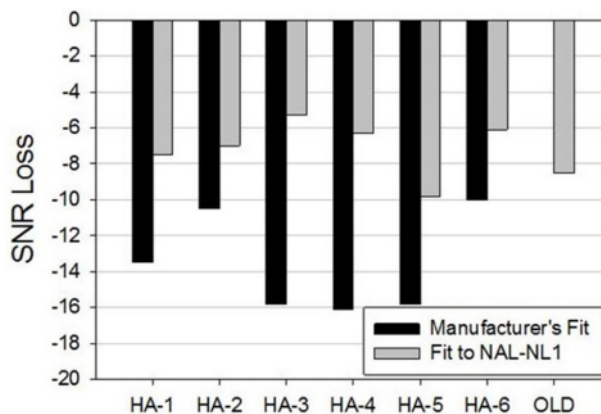

Figure 4: Mean QuickSIN scores (SNR-Loss) for hearing impaired individuals fitted with six different pairs of hearing aids. One pair (labeled "old") were 10-year-old analog single-channel products; the other six pair were the premier model of the six leading hearing aid manufacturers.

If you first look at the far right bar, you see that the mean SNR-Loss for the old analog instruments was around 8 dB SNR-Loss. Compare this to the mean performance for the manufacturers' recommended fitting for the six different new premier hearing aids (dark bars). The results for HA-6 are fairly similar to those of the old analog hearing aids, but note that when the participants used HA-3, HA-4 or HA-5, their scores were about 8 dB worse.

As we would predict, when the premier hearing aids were programmed to the NAL-NL1 rather than the manufacturer's recommended first fit, you now see that most of the new products are performing 2 dB or so better than the old analog instruments. Note that with HA-3, for example, the mean QuickSIN score improved by over 10 dB simply by changing the programming from manufacturer's fit to NAL. This is clearly a good example that it's not the brand of the hearing aid that matters so much, it's the person who programs it. A 10 dB SNR improvement could be a life changing difference for some hearing aid users. Similar findings for speech recognition in quiet were reported by Valente et al. [24]. Shown in Figure 5 are the speech recognition scores for a NAL-NL2 fitting compared to the manufacturer's proprietary fit. Observe that the 25th percentile of the proprietary fitting exceeds the 75th percentile of the manufacturer's fit.

As you might expect, these speech recognition advantages for the NAL-NL2 carry over to the real-world, as evidenced by self-assessment inventories. Valente et al. [24] report there was a significant advantage for the NAL-NL2 fitting (compared to the proprietary) observed with the self-
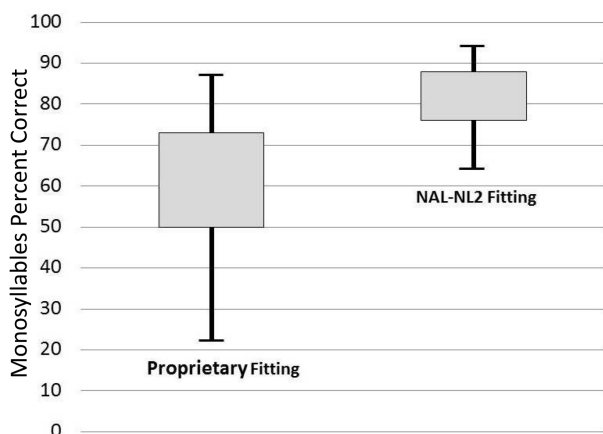
Figure 5: Percent correct performance for speech recognition in quiet for the manufacturer's proprietary fitting compared to a verified NAL-NL2 fitting. Data from Valente et al. [24].

assessment ratings of the Abbreviated Profile of Hearing Aid Benefit (APHAB) – it was a cross-over design, so all 24 participants had both conditions. The APHAB findings from Valente et al. [24] are plotted in Figure 6, and for interest, the APHAB norms for elderly individuals with normal hearing from Cox [27] have been added. Two things are apparent: The NAL-NL2 fitting is superior to the proprietary default, and when fitted to the NAL algorithm, the real-world performance for this group of hearing aid users was equal to individuals with normal hearing.
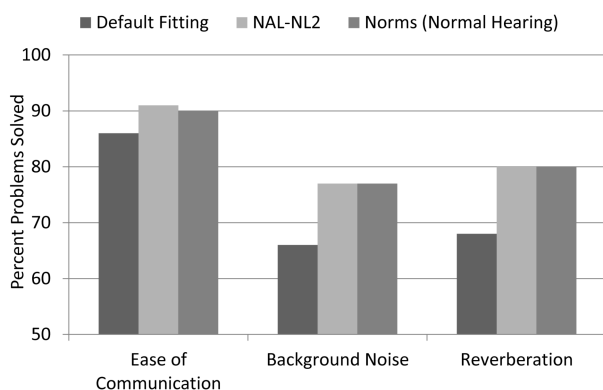

Figure 6: APHAB performance (percent problems solved) for the Ease of Communication, Background Noise and Reverberation subscales comparing the proprietary (default) fitting and the NAL-NL2 fitting, from Valente et al. [24]. For comparison, also shown are the normative data from Cox [27] for elderly with normal hearing.

Is all this talk about proprietary fitting really necessary? Are they really being commonly used? We have a good idea that this is the case, at least in the U.S., and probably in Europe as well. Here is a snapshot. Leavitt et al. [28] reported on probe-mic measures for a total of 97 individuals (176 fittings) who had been fitted at 24 different facilities within the state of Oregon. The participants were current hearing aid users and were wearing hearing aids that came from 16 different manufacturers; the average age of the product was 3 years. These researchers found that in general, all the hearing aid users were under-fit.

When RMS errors were computed, they found that 97% of the wearers were >5 dB from NAL-NL2 targets, and 72% were >10 dB.

ZAUD: You have been talking about proprietary fittings, but all manufactures have the option of using either the NAL-NL2 or DSLv5.0 in their fitting software. Does this reduce the need for verification?

Mueller: In a word – make that two words – absolutely not! "A rose is a rose is a rose," is a commonly used phrase dating back to the early 1900s. I can assure you that the NAL is *not* the NAL is *not* the NAL, regardless what you might see on a manufacturer's fitting screen. First, we would not expect the software fitting to the prescriptive target to be a perfect match in the real ear. We would expect variance above and below based on the individual's RECD. That is, if the average RECD for a given frequency is 8 dB, and the individual's RECD when fitted with a given earmold is 11 dB, then we would expect the output to be 3 dB above target. But the problem is bigger than this. Much bigger. Shown in Figure 7 are data from our comparative lab study mentioned earlier [22].
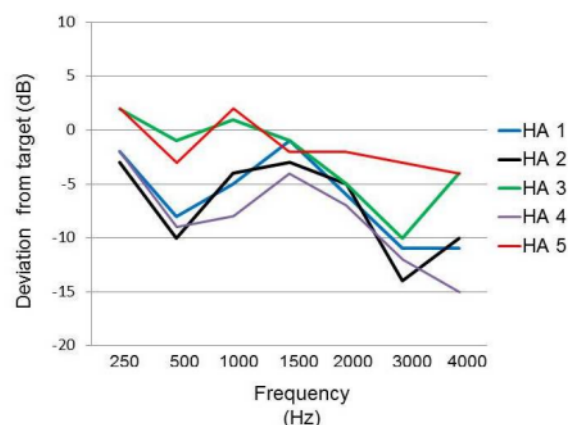

Figure 7: Mean deviation from NAL-NL2 target based on real-ear measured output (55 dB input signal; n=16) for the premier hearing aids from the five leading manufacturers. Hearing aids programmed to the manufacturer's NAL-NL2.

These are the results for a 55 dB SPL speech-signal input, averaged over 16 ears for the premier product of five different manufacturers, programmed to NAL-NL2 (according to the software). What you see is the mean measured REAR deviation from the NAL-NL2 target. The deviations are similar to what we saw for the proprietary fittings. It is important to mention, that in all cases, *the deviation from target on the manufacturer's fitting screen was no more than 1 dB*. Imagine audiologists, looking at the software fitting screen simulation and patting themselves on the back for being a good person and fitting to target, when it's very possible they could be missing target by 10 dB or more in the high frequencies.

Amlani et al. [29] reported very similar results. They found that the manufacturer's NAL-NL2 fitting, on average, fell nearly 10 dB below real-ear NAL targets, for both soft and average speech inputs. This led to speech recognition

(Connected Speech Test) to be significantly poorer when their participants were fitted with the manufacturer's NAL-NL2 compared to a real-ear-verified NAL-NL2.

Are things better in 2020? As part of a larger study, we recently conducted probe-mic measures on 2020 premier hearing aids using 2020 software. Figure 8 shows the REAR findings for soft speech inputs (ISTS) for 16 NAL-NL2 fittings, all to the same mild-to-moderate hearing loss sloping from 30 dB in the lows to 70 dB in the high frequencies. We were careful to match the settings of the software to that of the probe-mic equipment; bilateral fitting, experienced user, gender neutral, closed earmold. As shown, and similar to previous reports, REARs fall well below NAL-NL2, with mean deviations of 10 dB or more. We did not sample all brands, so it's possible some manufacturers have a better match than this, but in the past under-fitting for the NAL-NL2 has been common.
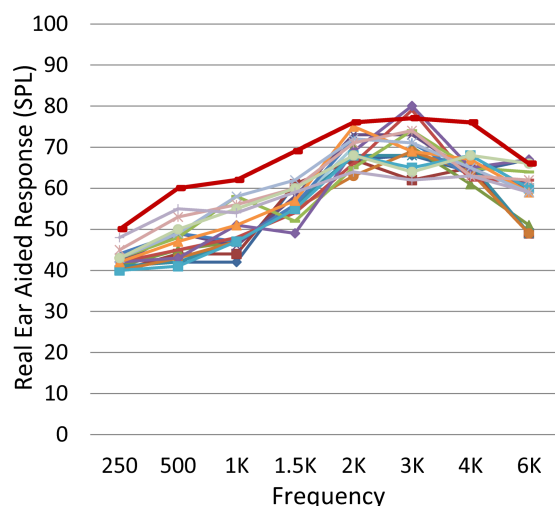

Figure 8: REARs of 16 fittings, programmed to the manufacturer's NAL-NL2 (Input: 50 dB SPL ISTS). Upper line is NAL-NL2 prescriptive target.

All of this is not really news. Going back to 2003, Hawkins and Cook [30], testing 12 hearing aids from different manufacturers, reported that gain in the high frequencies was 8–10 dB below the software simulation. Aazh and Moore [31] used four different types of hearing aids and programmed them to the manufacturer's NAL-NL1 using the software selection method. When probe-mic verification was conducted, only 36% of fittings were within ±10 dB of NAL targets. Aazh et al. [32] conducted a similar study with open fittings. They reported that of the 51 fittings, after programming to the manufacturer's NAL in the software, only 29% matched NAL-NL1 targets within ±10 dB. And this problem doesn't appear to be unique to the NAL methods. Folkeard et al. [33] reported that the manufacturer's DSLv5.0 fell ~7 dB below target for both soft and average inputs.

The bottom line is pretty simple. If we consider that our primary fitting goal is to optimize speech recognition, hearing aid users do the best when fitted to a validated prescriptive fitting. We also know that the software fitting screen is not correct, and therefore, the only way to know

the fitting is appropriate is to conduct probe-mic verification.

ZAUD: Perhaps your discussion here might encourage some clinicians to make probe-microphone verification a more routine part of their hearing aid fitting. For readers who have been away from probe-mic measures for a few years, is there anything new?

Mueller: If we look at the last 10 years or so, there have been some definite trends. While I believe audiologists still use the REIG for verification in some parts of the world, nearly everyone in the U.S. uses REAR targets; this provides a clear picture of audibility, and of course eliminates the need to conduct an REUR. Another area of change is that we finally have all agreed on a good speech signal for testing, the ISTS [34], [35], which is available on most all equipment and commonly used.

In more recent years, some changes include:

- The use of simultaneous bilateral measurements. Each hearing aid of course still needs to be programmed independently, but the bilateral measures do save some time. For example, if AGCo kneepoints were originally adjusted correctly, only one REAR85 presentation may be necessary.
- Some systems now have an automated method to inform the examiner if the probe tube is within the desired 5 mm of the eardrum. This relieves some apprehension for inexperienced examiners, and helps ensure valid measures.
- Most hearing aid companies have partnered with one or more probe-mic manufacturer to offer autoREMfit [36], [37]. This is when the hearing aid software and the probe-microphone equipment communicate with each other, and the fit to target is automated – only a few mouse clicks from the audiologist are needed, This isn't really new, as it's been available for 20 years [38], but only recently has it become widespread. There are still some minor issues to work out, but research shows that it is valid, and reduces the time spent on verification by about ½ [33].
- Finally, from a procedural standpoint, more and more audiologists are conducting an initial RECD, and then using these values for the HL-to- ear-canal-SPL conversion along with the RETSPL (rather than the average RECD stored in the probe-mic equipment). This adds accuracy to the displayed ear canal audiometric thresholds (for audibility decisions), and in turn, adds to the accuracy of the prescriptive targets, which are calculated from these thresholds.

ZAUD: Is there a specific verification protocol that you recommend?

Mueller: In our most recent book, we have step-by-step guidelines for all types of probe-mic measures, which include verification of direction technology, noise reduction, frequency lowering, the occlusion effect and other fun-to-do measures [16].

Regarding basic verification, I would first pick my favorite prescriptive method (either the NAL-NL2 or DSLv5.0) from the fitting software and initially program to that. Then make programming changes to obtain a match to target for soft (55 dB SPL), average (65 dB SPL) and loud (75 dB SPL). Some people use 50–65–80 dB SPL, which is okay too. The key is to start with the soft input level – for some reason, people like to start with average, which I think is a mistake. Soft will nearly always be under-fit, so if you start with average, and obtain a target match for average, and then go to soft and make appropriate adjustments (increase gain), average will then be too loud, and you'll need to go back and re-program average. So why not start with soft? After soft, I would then go to loud (75 dB SPL). After programming loud, average should be pretty close to okay, as it falls in the middle of the two levels that already have been programmed correctly. At some point, I would probably also do an REOR, just to ensure that the degree of openness or tightness of the eartip meets my fitting goals.

It is then important to do an REAR85 to ensure that the MPO is set correctly. We used to be concerned that the MPO would be too high, but recently manufacturers have become pretty conservative in their default MPO settings, and now it's more common that we need to move our AGCo kneepoints up rather than down. Hopefully the audiologist doing the fitting has conducted pure-tone LDLs, and entered them into the probe-mic software, so we then have targets for the REAR85 measure. Finally, I'd present some of the obnoxious noises available on the probe-mic equipment at 85 dB SPL, to ensure that the output is "Loud, But Okay" and did not reach the hearing aid user's LDL (using the Cox 7-point loudness chart [39]). That's about it.

ZAUD: In closing, let's go back to the underlying issue, that many audiologists do not see the need for real-ear verification of gain and output. Is this something that should be addressed by professional organizations?

Mueller: To some extent, it already has been. Most organizations do have best practice guidelines regarding the fitting of hearing aids, such as those of the EUHA. All the guidelines I have seen over the past 25 years state that probe-microphone verification should be conducted. In guidelines published in 2005, the International Society of Audiology went so far as to state what variation from prescriptive target was allowable [18]. But – these are *guidelines*; there really is no penalty if they are not followed. I did hear that in the province of British Columbia in Canada, disciplinary action can be taken for not conducting verification routinely – perhaps the loss of a person's dispensing license. But this, unfortunately, is not common.

In the near future, over-the-counter (OTC) hearing aids will be available in the U.S. It's going to be important that audiologists differentiate themselves from what can be purchased in Aisle 7 of the neighborhood Big Box store. It's expected that some of these OTC products will come with smartphone apps for the prospective user to fit themselves. As I mentioned earlier, at least one study has reported that hearing aid users will fit themselves to gain that is quite similar to the NAL-NL2 prescription [13]. If audiologists are not conducting real-ear verification, and not fitting to the NAL-NL2, logic would suggest that for individuals who can navigate the fitting app successfully, they would be better off to fit themselves!

Several years ago, Catherine Palmer, now the President of the American Academy of Audiology (AAA), wrote an article describing how the failure to do verification is an ethical violation [40]. I agree. Consider that most professional audiology organizations and licensure boards have a Code of Ethics. Here is an example from the Ethics Code of the AAA [41]: *Principle 4: Members shall provide only services and products that are in the best interest of those served*. I fail to see how charging someone a sizeable amount of money, and then sending them out the door with a fitting that has little or no audibility for soft speech, is in the best interest of the patient. I think we can do better.

Abbreviations

- AGCo: automatic gain control for output
- APHAB: Abbreviated Profile for Hearing Aid Benefit
- DSL: Desired Sensation Level
- ISTS: International Speech Test Signal
- JND: just noticeable difference
- LDL: loudness discomfort level
- MPO: maximum power output
- NAL: National Acoustic Laboratories
- NL1/NL2: non linear
- OTC: over-the-counter
- REAR: real ear aided response
- RECD: real ear coupler difference
- REIG: real ear insertion gain
- REM: real ear measures
- RETSPL: reference equivalent threshold in sound pressure level
- REUR: real ear unaided response
- REOR: real ear occluded response
- RMS: root mean square
- SII: speech intelligibility index
- WDRC: wide dynamic range compression

Notes

Competing interests

The author declares that he has no competing interests.

References

1. Schwartz D, Walden B, Mueller H, Surr R. Clinical experience with probe-microphone measurements. In: Conference of the American Speech-Language-Hearing Association (ASHA); 1980; Detroit, Michigan.

2. Byrne D, Tonnison W. Selecting the gain of hearing aids for persons with sensorineural hearing impairment. Scand Audiol. 1976;5:51-9. DOI: 10.3109/01050397609043095

3. Dillon H. The research of Denis Byrne at NAL: Implications for clinicians today. AudiologyOnline. 2001 Dec 24. Available from: https://www.audiologyonline.com/articles/research-denis-byrne-at-nal-1200

4. Johnson E. Same or different: Comparing the latest NAL and DSL prescriptive targets. AudiologyOnline. 2012 Apr 9. Available from: https://www.audiologyonline.com/articles/20q-same-or-different-comparing-769

5. Byrne D, Dillon H. The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear. 1986 Aug;7(4):257-65. DOI: 10.1097/00003446-198608000-00007

6. Byrne D, Dillon H, Ching T, Katsch R, Keidser G. NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. J Am Acad Audiol. 2001 Jan;12(1):37-51.

7. Keidser G, Dillon H, Flax M, Ching T, Brewer S. The NAL-NL2 Prescription Procedure. Audiol Res. 2011 May;1(1):e24. DOI: 10.4081/audiores.2011.e24

8. Mueller HG. Fitting hearing aids to adults using prescriptive methods: an evidence-based review of effectiveness. J Am Acad Audiol. 2005 Jul-Aug;16(7):448-60. DOI: 10.3766/jaaa.16.7.5

9. Mueller HG, Hornsby BWY. Trainable hearing aids: the influence of previous use-gain. AudiologyOnline. 2014 Jul 14. Available from: https://www.audiologyonline.com/articles/trainable-hearing-aids-the-influence--12764

10. Palmer C. Implementing a gain learning feature. AudiologyOnline. 2012 Aug 8. Available from: https://www.audiologyonline.com/articles/siemens-expert-series-implementing-gain-11244

11. Keidser G, Alamudi K. Real-life efficacy and reliability of training a hearing aid. Ear Hear. 2013 Sep;34(5):619-29. DOI: 10.1097/AUD.0b013e31828d269a

12. Mueller HG, Hornsby BW, Weber JE. Using trainable hearing aids to examine real-world preferred gain. J Am Acad Audiol. 2008 Nov-Dec;19(10):758-73. DOI: 10.3766/jaaa.19.10.4

13. Sabin AT, Van Tasell DJ, Rabinowitz B, Dhar S. Validation of a Self-Fitting Method for Over-the-Counter Hearing Aids. Trends Hear. 2020 Jan-Dec;24:2331216519900589. DOI: 10.1177/2331216519900589

14. Johnson EE, Dillon H. A comparison of gain for adults from generic hearing aid prescriptive methods: impacts on predicted loudness, frequency bandwidth, and speech intelligibility. J Am Acad Audiol. 2011 Jul-Aug;22(7):441-59. DOI: 10.3766/jaaa.22.7.5

15. Keidser G, Dillon H. What's new in prescriptive fitting down under? In: Palmer C, Seewald R, editors. Hearing Care for Adults Conference 2006. Phonak AG; 2006. p. 133-42.

16. Mueller HG, Ricketts TA, Bentler RA. Speech Mapping and Probe Microphone Measures. San Diego, CA: Plural Publishing; 2017.

17. Mueller HG, Ricketts TA. Open-canal fittings: Ten take-home tips. Hear J. 2006;59(11):24-39. DOI: 10.1097/01.HJ.0000286216.61469.eb

18. International Society of Audiology. Good practice guidance for adult hearing aid fittings and services - Background to the document and consultation. 2005. Available from: https://www.isa-audiology.org/ members/pdf/GPG-ADAF.pdf

19. British Society of Audiology. Guidance on the verification of hearing devices using probe microphone measurements. 2018. Available from https://www.thebsa.org.uk/resources/

20. Caswell-Midwinter B, Whitmer WM. Discrimination of Gain Increments in Speech-Shaped Noises. Trends Hear. 2019 Jan-Dec;23:2331216518820220. DOI: 10.1177/2331216518820220

21. Cornelisse LE, Seewald RC, Jamieson DG. The input/output formula: a theoretical approach to the fitting of personal amplification devices. J Acoust Soc Am. 1995 Mar;97(3):1854-64. DOI: 10.1121/1.412980

22. Sanders J, Stoody T, Weber J, Mueller HG. Manufacturers' NAL-NL2 fittings fail real-ear verification. Hear Rev. 2015; 21(3):24.

23. Pearsons K, Bennett R, Fidel S. Speech levels in various noise environments. Report No. EPA-600/1-77-025. Washington DC: US Environmental Protection Agency; 1977.

24. Valente M, Oeding K, Brockmeyer A, Smith S, Kallogjeri D. Differences in Word and Phoneme Recognition in Quiet, Sentence Recognition in Noise, and Subjective Outcomes between Manufacturer First-Fit and Hearing Aids Programmed to NAL-NL2 Using Real-Ear Measures. J Am Acad Audiol. 2018 Sep;29(8):706-21. DOI: 10.3766/jaaa.17005

25. Leavitt R, Flexer C. The importance of audibility in successful amplification of hearing loss. Hear Rev. 2012;19(13):20-3.

26. Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2004 Oct;116(4 Pt 1):2395-405. DOI: 10.1121/1.1784440

27. Cox RM, Alexander GC. The abbreviated profile of hearing aid benefit. Ear Hear. 1995 Apr;16(2):176-86. DOI: 10.1097/00003446-199504000-00005

28. Leavitt R, Bentler R, Flexer C. Hearing aid programming practices in Oregon: fitting errors and real ear measurements. Hear Rev. 2017;24(6):30-3.

29. Amlani AM, Pumford J, Gessling E. Real-ear measurement and its impact on aided audibility and patient loyalty. Hear Rev. 2017;24(10):12-21.

30. Hawkins D, Cook J. Hearing aid software predictive gain values: How accurate are they? Hear J. 2003;56(7):26-34. DOI: 10.1097/01.HJ.0000292552.60032.8b

31. Aazh H, Moore BC. The value of routine real ear measurement of the gain of digital hearing aids. J Am Acad Audiol. 2007 Sep;18(8):653-64. DOI: 10.3766/jaaa.18.8.3

32. Aazh H, Moore BC, Prasher D. The accuracy of matching target insertion gains with open-fit hearing aids. Am J Audiol. 2012 Dec;21(2):175-80. DOI: 10.1044/1059-0889(2012/11-0008)

33. Folkeard P, Pumford J, Abbasalipour P, Willis N, Scollie S. A comparison of automated real-ear and traditional hearing aid fitting methods. Hear Rev. 2018;25(11):28-32.

34. Holube I, Fredelake S, Vlaming M, Kollmeier B. Development and analysis of an International Speech Test Signal (ISTS). Int J Audiol. 2010 Dec;49(12):891-903. DOI: 10.3109/14992027.2010.506889

35. Holube I. Getting to know the ISTS. AudiologyOnline. 2015 Feb 9. Available from: https://www.audiologyonline.com/articles/20q-getting-to-know-ists-13295

36. Mueller HG, Ricketts TA. Hearing aid verification: Will autoREMfit move the sticks? AudiologyOnline. 2018 Jul 9. Available from: https://www.audiologyonline.com/articles/20q-hearing-aid-verification-226-23532

37. Mueller HG. Is autoREMfit a reasonable verification alternative? AudiologyOnline. 2019. Available from: https://www.audiologyonline.com/

38. Mueller HG. Probe microphone measurements: 20 years of progress. Trends Amplif. 2001 Jun;5(2):35-68. DOI: 10.1177/108471380100500202

39. Cox RM. Using loudness data for hearing aid selection: The IHAFF approach. Hear J. 1995;48(2):10, 39-44. DOI: 10.1097/00025572-199502000-00001

40. Palmer C. Best practice: It's a matter of ethics. Audiology Today. 2009;21(5):31-5.

41. American Academy of Audiology. Code of ethics. 2011. Available from: http://www.audiology.org/resources/documentlibrary/Pages/codeofethics.aspx

Corresponding author:

Prof. H. Gustav Mueller, PhD
1940 Harbor Drive, Bismarck, North Dakota 58504, United States
gus@gusmueller.net