Evaluating the Efficacy of Music Programs in Hearing Aids

Recorded February 15, 2023

Presenter: Joshua M. Alexander, PhD, CCC-A; Emily Sandgren, BS

- It's my pleasure to welcome back Dr. Joshua Alexander and his counterpart Emily Sandgren. Today's course is going to be about evaluating the efficacy of music programs in our hearing aids. And so this is just a great course, because it's about innovation in our industry. And during this month in February, AudiologyOnline is celebrating just that, innovation in our hearing industry and it's our third summit. For those of you who have missed the previous recordings from this month, they are available to you in the course library. We also like to thank our title sponsor CareCredit and all of the other corporate partners who are participating in this summit. Next week is our big finale and that's our roundtable. Please join us next Wednesday. And at this time, I hand the mic over to you, Dr. Alexander.

- Thank you, thank you so much for inviting me to present again this year. Thank you for everyone in attendance. And what I'll be presenting today, we'll be presenting a some of the results of AuD's Capstone Project that was co-conceived with my third year AuD student, Emily Sandgren. So Emily was responsible for coming up with this idea based on some of our discussion in the course about how some of the limitations of hearing aids and we conceived of this project. She is responsible for doing all of the recordings that we'll be talking about today through the hearing aids. And then being the nerd that I am wanting to know a little bit more, not just what we found, but while we found, I will be presenting the second half of the presentation about the analysis of the results and some of the acoustic findings that we have found.

Some disclosures, these are the course learning outcomes. For those of you who will be taking this course for CEU credits, just wanna point your attention to the upper right corner of the slides. We'll have a Q1, 2, 3, 4, 5 up there for you for where to find the content that's related to the CEU questions. At this point, I will hand off the presentation to Emily.

- Thank you for the introduction and thank you to AudiologyOnline and to Christy and Kimberly for hosting us today. I'm so grateful for the opportunity to present today and thank you to all of you listening who are interested in this topic. So jumping right into our topic for today, we'll be talking about music and hearing aids. So a good starting point for that is to look at why music is important and the data on this slide provides some numbers to demonstrate just that, music is really important to a lot of people. Because it's such a significant source of entertainment for so many Americans, which is associated with quality of life, Greasley et al found that being deprived of, or separated from a valued source of entertainment results in negative emotional consequences.

So to provide holistic care focused on improving quality of life, not just communication, improving access to music is important for patients with hearing loss. And patients with hearing loss are just like everyone else. They still value and enjoy being able to listen to music. And Alinka Greasley found recently that a large majority of hearing aid users do use their devices to listen to music. Unfortunately, even though hearing aid users want to use their devices to listen to music, hearing aid users have consistently expressed dissatisfaction with the sound quality of music when using hearing aids. So as you can see here, multiple studies have found that distortion was a significant contributing factor to the problem of sound quality.

So to back up just a little bit, why do we need to consider music differently than speech with hearing aids? On this slide, we'll have a list with just a few examples. There are more differences than this, but to highlight just a few, preferred listening levels are higher for music than for speech. This can be a problem if it results in oversaturating the hearing aid, which could contribute to that distortion and poorer sound quality. The study listed here is also the information we used to determine recording levels for our study, which we'll discuss later. Secondly, music has higher

crest factors than speech, although this is not as much of an issue with recorded music, because it's already compressed.

So the crest factors for the samples for our study averaged about 10 dB, while speech is that crest factor of 12 dB. In general this difference in crest factors occurs because the vocal track is heavy on damping, which reduces peaks, unlike musical instruments. But again, recorded music is already compressed from the studio editing. And thirdly, music is wideband while speech is narrowband, for a few reasons. With music, there can be multiple instruments of high and low pitches playing at the same time, unlike having just one voice and one signal for speech. And even with one instrument, the signal may contain a low fundamental frequency and more higher frequency harmonics than for speech which has lower frequency speech sounds or higher frequency speech sounds in rapid succession, but not at the same time.

So with music, the frequency response gets flattened if we control game channels in the same way we do for speech. And please note the Q1 in the upper right-hand corner of this slid. For those of you completing the CEU quiz, this corresponds to that content and the quiz question number and that will be noted on other slides. One other fourth reason not listed here that I'll also mention briefly later is that while speech has a long-term average speech spectrum, which allows hearing aid algorithms to recognize it and pick it out from the environment, there's no long-term average spectrum for music. So it's difficult for hearing aid algorithms to recognize and differentiate music from background noise.

Because of these differences, the hearing aid processing features that have been designed to benefit the hearing aid user when listening to speech have the opposite effect when listening to music. So instead of improving sound quality, they have a negative effect. Like the previous slide, this will be an incomplete list, but again, to mention a few of the highlights, compression is problematic, as I just mentioned,

because recorded music is already compressed, so this results in it being doubly compressed. And this can alter the spectral balance, which changes the harmonics, which are one of the things that differentiate one instrument from another. So if a cello, a french horn and an oboe all play the same note in the same octave, those harmonics are how we tell them apart and how those sounds are different.

Directional microphones can affect the temporal envelope and that temporal envelope, as a review, is slowly varied, change in amplitude over time and they can also result in a roll off of low frequencies and noise reduction can also distort the temporal envelope and with impulse or transient reduction, while we want to reduce unwanted impulse sounds for speech, music has transient sounds such as percussion instruments that we do want to maintain. And we don't have quite enough time today to go into feedback suppression and frequency lowering, but you can see those notes here. So if the hearing aid processing features designed for speech are not helpful for music, what changes could improve the music listening experience? Lots of past research has focused on what compression strategies might improve sound quality for music.

So two studies have shown that linear compression is preferred over WDRC. And slow-acting compression has also shown to be preferred over fast-acting compression. So as you notice that Q3 in the corner, one important factor is slow-acting compression. And secondly, the frequency range has been associated with improved sound quality for music as well. So ultra low and ultra high frequencies here mean frequencies below and above that 200 to 8,000 hertz range where speech mapping and programming is usually performed. For the ultra high frequencies, hearing loss itself is a limiting factor, but including ultra-low frequencies is important to focus on because it allows us to capture the fundamental frequencies of those low pitches in music, which with larger instruments can be lower than fundamental frequencies in the human vocal tract.

Without capturing those fundamental frequencies, we can still recognize a pitch based on its harmonics, but these studies indicate that including those fundamental frequencies improve sound quality. And this could be because including that low frequency energy recreates the sound quality that listeners experience when listening to recorded music in a semi-reflective room, because reflections in a room increase low-frequency energy and it creates a sound quality described here as spaciousness and naturalness. And that's what we're accustomed to hearing. So the final step in our review of background literature is the step from research to clinical application. In light of hearing aid users' dissatisfaction with music, the differences between music and speech, and listeners' preferences, manufacturers, as we know, have created special programs for listening to music, but there's lots to discuss about the use and effectiveness of these programs.

So as a first step in that discussion, music program use is inconsistent. You can see here that a low percentage of hearing aid users report having a music program. However, Greasley et al in 2020 reported that for a large portion of hearing aid users, music was never discussed in clinic. So lots of questions remain here. Is music not discussed because patients don't mention music when clinicians ask them about their lifestyle, unless they are in fact a musician? Would they mention it if they were directly asked a question about it? In the same article, Greasley et al who had interviewed both patients and audiologists found that clinicians reported that they don't feel confident discussing music in clinic or programming a hearing aid for music, because there just isn't enough training.

But would this improve if there was enough evidence that music programs are effective without needing extra programming changes? There's also a statistic here that 86% of the users who reported having a music program report using it, but it's possible that this could be reporting bias. If they knew they had one, they likely reported using it. If they didn't know they had a music program, they likely didn't report using their music

program. And if patients were never asked in clinic if music is important to them, they most likely would never know that music programs are available and if they want one or not. So there's not a clear picture about specifically why music programs are or aren't being used, but even if only 25% of patients want a music program, it will still be important to those 25% of patients to have a good music program.

So secondly, are the music programs working? Can clinicians create music programs for patients without needing extra programming changes? And can we expect that the music programs will improve sound quality of music for patients or are music programs and sound quality still unsatisfactory? So research has shown mixed results. As you can see here in both 2014 and 2019, users reported little difference between general programs and music programs. In 2022, some brands showed an improvement from the speech to music program while other brands did not. But there was a more significant change between brands or manufacturers than there was between programs. And here again, we see another piece of evidence for the importance of that low frequency information. So now that we've covered the background and other research relating to music programs and using hearing aids with music, the motivation for our study was to evaluate what programming features or other factors make a music program successful and evaluate the performance of music programs currently being used.

Even after previous research, which often focused on isolating one factor, like compression, there's still a need for research directly applicable to products used clinically today. So even if previous studies have established that slow-acting compression and extended low frequency response are preferred, how do all the programming algorithms interact together when they aren't isolated? And are the factors that hearing aid users prefer being applied in devices in the market today? And is it improving sound quality for the listener? So our first step in setting up study was to program the hearing aids used for the study. We used the highest level technology

receiver and canal hearing aid as of June, 2022 from the seven major manufacturers. And you can see those models listed here.

I program them to the N3 audiogram, which is one of the standard audiograms representing audiograms commonly met in clinical practice. And I verified the speech programs in the verified test box to represent best practice. Then after verifying the speech program, I created the default music program for each hearing aid. I didn't make any changes to the music program. Whatever features are turned on or off by the manufacturer's default is the music program that was used in the study. I also didn't make any changes to the gain for the music programs, because manufacturers emphasize how important their proprietary gain formulas are for their music programs. So this table shows the features in the music program for each manufacturer.

As you may note, all the microphone modes are either omnidirectional or at least omnidirectional in the low frequencies. So that virtual pinna is omnidirectional in the lows and directional in the highs. Almost all of the noise features are turned off for all the manufacturers. And we also noted bit depth here as we were considering whether that contributed to a hearing aid becoming oversaturated and a music sample sounding distorted. So once the hearing aids were programmed, I then used the test box to take recordings of a wide variety of music samples with each hearing aid in both the speech and music program. So 14 total conditions, speech and music for seven manufacturers. If the default general program for the manufacturer was an environmental classifier, I created a duplicate of the speech and noise program to lock it into speech and noise, because as mentioned earlier, music can be difficult for hearing aids to recognize even with an environmental classifier program, because there isn't one long-term average spectrum like there is for speech.

So locking them into the speech and noise program allowed us to evaluate the difference between speech and music programs in case a hearing aid in a real life

situation didn't recognize music as music and stayed in a speech and noise program. This is the list of music samples used. As I mentioned before, we used 80 dB for our recordings, which has been shown to be a preferred listening level for music for both normal hearing listeners and listeners with hearing loss. To choose these music samples, I contacted the cultural centers on Purdue's campus and asked them for suggestions of both currently popular songs and timeless classics from their culture. And from that group of suggestions, I selected samples with a good variation of instruments used and ensemble size that would represent a wide range of genres.

The first three samples in gray here were used as a training round during data collection to allow participants to become familiar with the range of sound qualities before beginning data collection with the remaining 10 samples. As another note, if you're interested in listening to our recorded samples, at the end of the slide deck, we'll have a link to a webpage where all of the recordings are posted. So finally, once we had the recordings, we collected data by asking participants to rate the sound quality of the recordings on a seven-point scale with the labels you see listed here. We did not specifically recruit musicians and non-musicians separately, but we used a musical experience questionnaire to evaluate musical experience.

And the 30 participants classified as musicians had seven or more years of formal training while non-musicians had four or less years of formal training. And with that I will hand it off to Dr. Alexander here to discuss our results and our investigation and analysis.

- Thank you very much, Emily, for the wonderful setup and excellent work in getting this project underway. I'm gonna go ahead and share some of the outcomes that we found. So what we found and then a search for looking for the acoustics and seeing if we can find some explanations for some of the results that we found. In addition to the 14 samples from conditions where we have the seven hearing aid brands with the

default speech program, default music program, we presented a control stimulus to the participants. And so what I wanna note here, this is the input, this is the long-term average spectrum of all of the music samples. This was the input to the hearing aids with the solid black line being the average, the red line being the peaks as defined as the 99th percentile over the short duration time windows, and then the blue line being the, what's operationally defined as the valleys or the 30th percentile.

One thing I wanna just kinda point your attention to is, as Emily had indicated, one of the key differences between speech and music is music has a much more low frequency component than speech. And we could see that here in the spectrum as shown in the green here, where the most intense parts of the signal are coming from the very lowest frequencies, those ultra-low frequencies that Emily had indicated, people in controlled studies prefer to have some of that access to that signal, even as low as 50 or 55 decibels. Knowing, however, that the frequency response can sometimes be an important determinant in people's preferences For listening to music, we did not present the sounds on the left, which were the original sounds.

What we wanted to do is we shaped the original samples of music to match the long-term average spectrum coming from all of the seven hearing aids in their default speech program. And that's what we see on the right. So what listeners heard was the music that was shaped on the right, so from a frequency perspective, frequency shaping perspective, the control stimuli were identical in the long-term average as coming from the hearing aids. And so our first results slide, as we saw the ratings varied from one being unacceptable all the way to excellent. And what we have shown here is for the seven different brands, notice the color coding scheme, this color coding scheme will be used throughout all of the results.

The conditions where the music program is on is indicated by the darker bars. When we say off, that is music program off, just simply meaning that is the default speech

program. Notice compared to the control stimulus, which is in the dark gray, that the hearing aids performed suboptimally, meaning that there was a significant degradation in the perceived sum quality associated with all of the hearing aid recordings. This is kind of hard to tell a little bit what's going on. So we did a little bit of a further cleaning up of the data and I'm just gonna show you that next. So what we did here is we wanted to control for, make each person their own control.

So the idea here is we gave no specific directions on how listeners were to use the scale. They had plenty of, they had 45 trials of just practice trying to kinda get acclimated to the overall variation in sound quality and then they were tested on the subsequent stimuli. That being said, you know, we were aware that some people might be a little bit more critical than others and have an overall lower average rating. In addition, some people might be a little bit more strict in terms of how they were, they put the ratings were, then also in terms of use of the scale, some people used just a smaller segment to the scale. The overall ratings may be only varied by three or four, others maybe used the entire scale.

So what we did, as shown here, is we normalized each participant's ratings according to their own use of the scale. So we converted all of their ratings into what we would use in statistics as Z-score. So we kinda centralize everything to their own mean and then everything is scored relative to their standard deviation. And so here we see the same results as before, but now taking into consideration each individual's own use of the scale. So on the Y-axis, we have Z-score rating. And just to kinda re-familiarize you with how Z-scores work, one Z-score, so if we look at the gray bar associated with the control stimulus, it's at one. That means that the control stimulus on average is one standard deviation higher than the individual's mean rating.

And mean rating here then corresponds to zero on the scale. From this perspective, it's pretty clear that there were some favorites amongst the brands and the long-term, the

overall stories is that consistent with previous findings, there's a bigger difference between individual brands, that is the different colors, than between the music program and the speech program between the different shadings. So we see here an overall preference for brands B with the music program on and off. In fact, music brand B in the default speech program was rated better, was a higher quality than with the music program on. The next favorite was brand F over on the right there. In addition then, one of the interesting things to note about this graph is so the light colored bars can correspond to default speech program.

But if you look at the change in the ratings from the light to the dark with the music program on, we see more of a similarity. Differences become smaller between the brands with their music programs on. So they kinda get closer to that mean rating. Trying to examine then how the music programs affected individuals' overall ratings, what we see here is just simply a difference in their ratings. Compare it with music being, preference for music being in the positive part of the graph here, so towards upper part of the graph, indicates a greater preference for the music program and towards the bottom part would indicate that there was more of a sound, the speech program was rated higher than the music program.

What we see here, as indicated, are brands A, D and E. They had significant improvement with the music program compared to their default speech program. It should be noted however that these are the three conditions, three brands that were the rated the overall lowest in terms of the sound quality. So A, D and E in the default speech program were the three lowest ratings. So going to the music program, we see a significant improvement. Interestingly are the brands that were overall rated highest, B and F and as indicated before, we see, interestingly, a significant preference for the speech program over the music program, and again. So we kind of see a regression towards the mean, I think in both of these cases where the music program is kinda bringing things closer together.

And program G or manufacture G and C, we do not see any significant differences between the ratings for the speech program or the music program. Comparing the musicians versus the non-musicians. This was actually of high interest to Emily in terms of, you know, whether or not musicians would be a little bit more sensitive to the differences between the different hearing aid brands, the differences between the music programs versus the speech programs. And here is a breakdown by musicians and non-musicians. The musicians in blue, the non-musicians in the yellow-orange color, just simply indicating that there was no major differences between musicians and non-musicians in terms of when we kinda centralize everything according to an individual's own use of the scale.

So we don't see that in fact that musicians were any different than non-musicians in this case. So that's the what. And as I've kinda indicated, I have a high interest in learning and trying to figure out why we see this differences between the brands and maybe between the different, the speech and the music program. One thing that kinda came to mind late in the project was as well, you know, we've obtained these recordings in a test box using a 0.4cc coupler and we wanted to, you know, while we know that these hearing aid analyzers are very good in specifying level in frequency, maybe they're not good for presenting music. So what we did here is we obtained recordings in the test box.

We just simply played the music as we did before, but there was no hearing aids. So we kinda, using a small group of listeners, kinda looking at how they, and we just basically reran the experiment. We added these additional stimuli. So they were still evaluating all of the hearing aids just as they were before, but we've put in this extra condition where they was from the test box recordings. It should be noted that the test box does roll off its frequency response below 200 hertz. This was obvious from the original recordings and maintained even when we normalized 'cause when I normalized

for the frequency response, I normalized over the typical speech range from 208,000 hertz. So you see, down here in the low frequency end, you see a little bit of a low frequency drop in the red for the recordings from the test box compared to the original stimuli.

And I think when we look at the results, what we see here is that the test box condition is rated just slightly lower, reaches statistical significance compared to the original recording. So if we look at our scale, the control condition, the original ones that we presented, fall somewhere between good and very good. So as I call it, pretty good. The test box recordings fall right about at the good level. Now it should be noted that the average rating for the hearing aids was about a four, so fair. So a couple of conclusions. First conclusion, I think the difference between the control stimulus and the test box stimuli, and I'm gonna come back to this, might be this preference for the ultra low frequencies.

'Cause we've seen that in the previous literature, just even giving an extra five or six decibels below 200 hertz can make a difference in individual's preferences for one hearing aid or, sorry, just one sample of music over the other, okay. Now you be careful, I've been saying the word preference, but we never actually asked individuals preference. We just ask 'em to over overall rate the individual song quality on their own merits. So this is not a comparison stimuli, which one do you like better or the other? It's just an idea there was is, you know, we opened up the possibility that all of the hearing aids would be perfectly acceptable. So this is, I've gotta be, just gonna make a note of my terminology here.

The other take-home message, however, is that even when we control for any sort of effects of the frequency response or associated with the test box itself, I think we can be fairly confident of concluding that the sound quality from hearing aids does leave some room for improvement over what we might otherwise have. Just maybe for

example, simply listening through our commercial earbuds. As Emily indicated, the research has shown that people, that when music is presented through hearing aids or through simulations, listeners prefer linear or over WDRC mostly. And with its multi-channel, wide dynamic range compression, people prefer slower compression ratios over faster compression, I'm sorry, slower time constants over faster time constants. And what we have shown here are actual measures of the compression ratios.

So in literature you might see this as effective compression ratios, so not what was in the software, but what was actually measured in each of the frequency bands. And we simply did that by looking at the dynamic range of the input and dividing it by the dynamic range of the output and that's what we have shown here. So these are real effective compression ratios. Again the color coding is the same. The compression ratios from the samples that were recorded through the music program are shown by the siloed lines. Those with the speech program are shown by the dotted lines. A few interesting observations is shown by these down arrows next to the scale for four of the brands.

In the music program, the compression ratios decreased. So this idea that, you know, if they're doing anything intentional between the different music programs, that if anything, maybe trying to make them more linear in the output and that we do see some indication of that. The other take-home from this is is that if we kinda look across the preferred brands, B and F, compared to some of the others, I don't see any systematic relationship between the overall compression ratios. And in fact if we run correlations between these compression ratios and individual's ratings, there is no relationship. You might, if we expected, one, it would be the lower the compression ratios, the closer they are to one that individuals might indicate a higher sum quality rating.

And that does not seem to be the case in this. The other question then is, well, let's consider, we know from headphone studies that individuals expect a certain frequency response, they prefer, I'm sorry, a certain frequency response. Overall frequency response or specter balance can be an influencing factor. And so what we have shown here is the overall frequency response for the speech programs is the music off compared to the average. So just took those seven recordings with just the speech programs and then compared each individual brand next to that. And what we see here overall, with a few notable exceptions, we see everything is pretty close in terms of their output plus or minus four decibels are still across the brands.

And that makes sense because Emily programmed all of these in the test box to match prescriptive targets. So it makes sense in that regard that they should be fairly close to one another. We see a few oddities, see the dark blue, would be brand E, have this weird peak up here at about between 200, 500 hertz. And the other one would be brand D, the light blue having, I see a pretty sharp roll off around at 8,000 hertz and above. I point this one out, because this was the one brand that was rated the lowest. It's on the high frequency end. Keeping in mind that our listeners, our participants had normal hearing, so they had access to these higher and ultra higher frequencies that certainly could've influenced their overall sound rating.

In an actual hearing aid case usually, as indicated, the hearing loss is gonna kinda be the determinant of the upper frequency limit. So it very well could be conceivably one of the influencing factors, at least for this particular brand, could've been some of this lack of high frequency energy. And the overall take home will be is that there's one overriding, I think, contributing factor and then there's maybe a few one-off explanations where we could just, might conceivably explain that one hearing aid or condition is slightly worse than the other because of slight differences like this. Another thing I wanna point out is that we know from previous studies with hearing aids,

hearing peer listeners and even normal hearing listeners doing headphone preference ratings is this, an importance of those ultra low frequencies below 200 hertz.

And we're gonna take a little bit more, in a few slides here, a little bit more depth look at maybe some of the effects of this extra low frequency energy. So just kinda point this out where we can see some differences already and maybe what some of the consequences or some of the what's happening underneath the hood there. Now we do know that brands A, D, and E showed a significant improvement in their sound quality ratings going from the speech program to the music program. What's shown here is a difference in the overall long-term average spectrum as before, but comparing the music program to the speech programs. And again just kind of, you know, looking at the curves and looking at, you know, again, keeping in mind A, D and E, so the red, the light blue and the dark blue.

Not seeing any systematic differences between the overall frequency response from the music program to the speech program that could explain why they're performing better with the music versus the speech program. Interestingly too, I think we could also look at, you know, how similar those two responses are. There's not a lot of differences in the overall frequency response going from one to the other, except for brand G, it looks like, you know, maybe a very deliberate attempt to boost, in the orange line, to boost the extra high frequency gain. I see that around 8,000 hertz. Another thing that we've looked at was the MPO from the Verifit, so the actual measured MPO, and comparing that to the prescriptive estimates NAL-NL2s, the prescription that we used, the estimated LDL shown by the solid black line.

One of the things here, and this should be noted, these are the manufactured default MPO settings. And, you know Gus Mueller has, many times, in this forum and other forums, has emphasized the importance of considering getting loudness right. And to the extent that we can measure those LDLs and make full use of the user's dynamic

range. Oftentimes I teach a lot of students and the mindset is, hey, as long as the MPO curve doesn't exceed the estimated LDLs, that their job is done, but we have to point out the importance of making the most of their dynamic range, getting those outputs close to their LDLs and that's important for loud inputs, whether it be the patient's own voice or loud inputs like music, which is what we're talking about here.

And so there's gonna be take-home messages, is one of the things I think we should consider, is that we don't wanna be shooting ourselves in the foot, you know, in the sense that if we want patients to be able to enjoy these loud inputs, we gotta give that signal room to grow in the output. And one of the things that I see here is that these four conditions that I just highlighted, brand A, D and E, those should be familiar to you. They had the lowest MPOs in the mid frequencies compared to the other brands, particularly in the default speech program. However, there's a caveat, I'm not quite sure this actually may be the explanation, and I'll tell you why, okay?

One of the things I do notice too is interestingly, one of the big differences, if you look at A, the red, and E, the dark blue is that going from the music, sorry, from the speech program, the data lines to the music program, the solid lines, we see a significant improvement in the overall MPOs or an increase in the MPOs, which may, may be one explanation for why music was a little bit better than, rated better than speech. However, if we try to do a strict measure of association like correlation, it's not the case that more is better, that higher MPO is gonna result in higher sound quality. I think the point is is that, you know, we have to get beyond that certain minimum, 'cause if we look at the favored brand, the one that was rated the highest is yellow here, brand B and that kinda falls right in the mid range there.

So it's not necessarily the case that more is better. The other reason I'm a little, not so sure if MPO is the explanatory factor here, because if you look at the light blue, brand D, where we know we saw an improvement for music and speech and there is no

difference in the overall MPOs between those two programs. So I'm not entirely sure that is the explanatory factor. It may be one contributing factor, for sure. The other reason I'm not so sure is because if we look at just say manufacture F, which had one of, the music off, which had some of the lowest MPOs, default MPOs, what I've shown here is that this is a long-term average spectrum, okay?

This is 128 millisecond analysis windows, the same as you get out of our hearing aid analyzers. The peaks are showing in red. That's operationally defined as the 99th percentile. The magenta line that's at the top is, I just drew that in at 90, which was the lowest of the MPOs. And we see here is that the peaks are getting close to but they're not at or, well, they couldn't exceed the MPO. So I'm not entirely sure, although further analysis is in order, particularly in light of what I'll share with you in a moment where we'll be looking at those instantaneous levels. Not levels that have been averaged over certain time segments, but the individual time wave form, instantaneous peaks.

Maybe a few of those easily could have conceivably exceeded MPO, but a further analysis is in order to see if that is indeed the case. So this is in terms of the output, so just kinda thinking about, you know, what it is that might be contributing to patients' overall ratings. We asked individuals at the conclusion of the study, it was a open-ended question to provide some descriptors for how they made the ratings, what sort of factors influenced it? And so this is a word cloud, so the bigger the word, so we kinda, we categorized similar terms. But the bigger the word is, the more often that it showed up. You can see some obvious ones like clarity, sound quality and balance.

Interestingly too though is these other terms, static, fuzzy, distortion, robotic, grainy. I've highlighted those, because if you remember some of those introductory slides from Emily, that distortion does seem to be a commonly reported problem when listening to music with hearing aids. And so it's not just simply that people are saying, I like this one better than the other or that this one has more base than the other. There's actually

an, again, at the conclusion of this, there is gonna be a link. Everybody will be able to listen to these recordings yourself. There is a little bit something extra in these signals that was not there in the input which might be contributing to these qualifiers that relate to distortion.

One of the things that was looking at, in light of the prior studies, is this importance of those ultra low frequencies. So lemme just kinda orientate you to this graph here. On the y-axis are those are the ratings in the Z-score scale that we've talked about before. So higher on the graph, there's gonna be a better rating than lower on the graph. And then what's on the x-axis is the overall gain at 160 hertz and below. So these are frequency ranges that we would never look at for speech, okay? But knowing how important they might be for music, let's go ahead and see if there's some sort of relationship. And that's what we have shown here.

So what's shown here on the x-axis is the gain, and as you can see they're all negative. So further to the right is actually not more gain, it's actually more loss. And if we look at those brands where we had the highest on quality ratings, B and F, they had essentially no gain reduction, no loss of sound as measured from the output relative to the input. For the brands and the conditions that were rated, the lowest, they had the greatest amount of ultra low frequency reduction. So it's a negative gain. The output is actually coming out lower than the input here. This is a significant relationship. If you're familiar with R squared, if you wanna get to the traditional Pearson correlation coefficient, you just take the square root of this number.

R squared is kind of, roughly means that how much proportion of the variance in the data can be explained by the factor. So about 47% of the data. So we see a pretty significant strong measure of association where higher or, in this case, less reduction in the peaks is resulting in higher sound quality ratings. I don't think, per se, that this is a preference for ultra low frequencies. We know users prefer, if given the option, to have

a little bit extra ultra low frequency output. I think there's something else going on here. And so what I've done to this is a measure of the distortion, the temporal envelope distortion, temporal envelope being the slowly modulating amplitude variation over time which corresponds to ratings of quality.

So what we have here, in this case, a higher number on the y-axis is more distortion, it's worse. And so what I think is happening, so let's go ahead and look at, those that score the lowest, that have the least amount of reduction in the ultra low frequencies, have the least amount of envelope distortion at 160, for the 160 hertz band. If we look at those hearing aids and those conditions, they had the greatest amount of low frequency, ultra low frequency reduction. They had the most temporal envelope distortion. So I think it's not just simply, hey, that there's less ultra low frequency energy here. I think that there is some, it's actually distorting the signal in that area, the temporal envelope, particularly in those ultra low frequencies, that could be explaining why individuals preferred one hearing aid brand or condition over the other.

Let's take it a step further. So what I've done here, so what we see here is, is this is the instantaneous levels. These are the samples. So we see these clear white gaps and this is the time wave form in seconds. And on the y-axis we have is the decibel level, the instantaneous decibel level. And these gaps here then represent each of those music samples. What we have shown here is the, in the green, is for the 160 hertz low frequency band, the blue is for the composite signal, the overall signal. And what I wanna show you here is, is that even when we quantify things in terms of their peaks over a time window, those instantaneous levels do exceed those peaks and could be over-driving the input to the hearing aid.

So what I did was I took, knowing that this ultra low frequency ban may be a critical importance, I subdivided the time signal, okay? So where we had these, where the instantaneous levels exceeded the peaks, this 99th percentile, so what's shown in red,

I indexed, time indexed all of those time intervals where the 160 hertz input was at or above the peaks and compared that to all of those time intervals that were somewhere in between. So the 99th to the 30th percentile. Here's the thing as, this is interesting. So what we have here is input/output curves for the higher frequency bands separated into, so what's shown in blue are those instances when the 160 hertz band, I chose 160 hertz because it was the most intense of those ultra lows, so I chose that one intentionally.

So blue is just when the lowest frequency component was below the peaks, that's what's shown in blue. And when it exceeded the peaks, those are the data points shown in red. And what I'm seeing here is is that when the ultra low frequency band at 160 hertz was presumably in saturation at peak levels, we see a reduction in the output at higher frequencies all the way from 250, well, from 200 to all the way to 2,500 hertz. Okay, so that's where we see the red lines lower than the blue lines. This is indicating to me that there's some sort of saturation on the input side and that's why I pointed out earlier that I was noticing that the most intense portion of the signal is those ultra low frequencies.

So it's as if we're overdriving the input. Saturating, not just what's at that frequency range, but also affecting the higher frequency components which could be giving rise to this perceived distortion in the overall signal. And so if we just look at relationship here again on the x-axis is how much gain reduction or the gain, actually if you want the gain, gain reduction, and the ultra low frequency band comparing on the y-axis how much that is affected than the output at the higher frequencies. Okay, so again, it's as if we're somehow limiting or saturating the input to the hearing aid that's kinda bringing everything down, not just at that frequency range but the others. So kinda summarize the explanatory factors.

So what I've identified just now is what I think is ultra low frequency input limiting. I think this is an input problem in those low frequencies, not necessarily an output problem. And this is why Emily introduced this concept of the bits, what I'll get to in a moment, 'cause a lot of experts in the field, particularly Marshall Chase and who's done a lot of discussion on this topic, emphasizes, hey, if you have a limited number of bits, that's going to compromise your ability to encode those higher frequency signals. I have four thumbs up for this, because of the lines of evidence, we see that a reduced low frequency output, we see, associated with the limiting and the relationship to those ratings.

I see that that low frequency reduction, ultra low frequency reduction giving rise to envelope distortion, which is also related to the ratings. I've just documented, it looks like there's downstream effects on the higher frequencies when those ultra low frequencies are presumably in saturation. And then knowing what I know a little bit from some research is that our favored brands that had the highest sound quality ratings, B and F, have a different hardware design and they have low-cut microphones that roll off those low frequencies in order to prevent saturation at the digitizer. And then they compensate for that reduction in the signal processing. But on the input side of things, they intentionally bring down those ultra low frequencies for the point of minimizing saturation.

So these are the four lines of evidence why I'm thinking at least this accounts for some of our data. And if we wanna look at those R squared, it's about 50% of the variation in the data. Okay, it's not the only one, right? 'Cause I've been pulling out little of this. All it takes is one thing to be different about any of these hearing aids or conditions to give rise to these negative sound quality ratings. So I don't think it's any one factor per se, but I think that's the leading contributing factor. So this bit, this is what I was just alluding to earlier. Is it necessarily that more is better? Well, the two lower performers had the lowest bit depth.

Okay, bits, if you recall, every bit allows you another. It allows you six decibels of dynamic range. So fewer bits means fewer decibels to code sounds on the input. Manufacturers have clever ways of getting around that. But we do see that the two brands with 16 and 18 bits were on the lower end in terms of the sound quality ratings, but so was D, which has 24 bits. But as I indicated before, it also had, you know, not as high of a frequency response. So it could be a mixed bag in terms of some of these other one-off conditions. High frequency response, there we go, the brand D, had the lowest ratings, and the lowest output above 8,000 hertz.

So it's just ideas. It could be a little of this, could be a little of that. So that's why I give one thumbs up, one thumbs down. MPO certainly could be a limiting factor, but I don't think it's the only one, again because where we saw brand D between the music and the speech programs, there was no difference, but we saw a difference in the overall ratings between speech and music for D, but no difference in the MPO. But we do know that if you do have a overly restrictive MPO, it is gonna affect, not just sound quality ratings for speech, or for music, but it's certainly for speech. And then looking at all of the differences and the features with the music program, the compression ratios, overall frequency shape.

I'm giving thumbs down on that. I don't see any indication or systematic relationship or even a one-off explanation for why certain brands or conditions are favored over others. Limitations, some next steps. We intentionally used younger, normal hearing listeners just to kinda get the ball rolling and we didn't wanna have any sort of influencing factors of hearing loss. Obviously this could, research could be extended to using individuals with hearing loss. This is important, kind of glossed over, but, hey, everybody's going, well, we know with open fits, significant venting, we're gonna allow those direct low frequencies in, so why does this matter? Well, according to my suspicions, if it's an input side of things, even when you are getting the low frequencies

direct through event, that those low frequencies, if it's an input problem, is going to, could lead to further problems I've already indicated.

Not just in those low frequencies, but also some of those higher frequencies. So I don't think we're out of the woods if we wanna say, well, there's nothing happening in those low frequencies. They're getting it direct through the vents. So I don't think we're out of the woods for sure, but we can explore that. I can think of some easy simulations we could do by using a combination of filtering with the recordings and the original signal to just kinda see what, you know, changing the mixture ratio, just kind of see what might the contribution of having' a direct signal unamplified might look like. We only had one rating in terms of overall sound quality and acceptance.

We could've got gotten a little bit deeper, more in depth if we had looked at, you know, given, a lot of these studies will look at different dimensions related to sound quality and loudness and stuff. That might give a little bit of insight as well. And then finally, as I hold it here, 'cause this is my next step is to go ahead and look at the differences in the genres. We had all of those samples and some samples had a higher crest factor than others. So this is a pretty trivial thing just to kinda look at. Let's go see if some of these music samples are responsible for the differences in ratings, whereas other samples, individuals might not have been as sensitive to the differences in the conditions.

That'll be an interesting look. Finally, conclusions, the quality of recorded music via hearing aid, microphone input, leaves room for improvement. I think that's pretty clear. We saw bigger differences between brands than between the speech and the music programs. We're not the only ones who have shown that. Brands with the worst ratings and the speech programs significantly improved in the music program, and vice versa. Those that had the highest ratings in the speech program had the, so we saw a significant drop with the music program. We didn't see any differences between

musicians and non-musicians. We have evidence that ultra-low frequency input limiting is a major factor. So again, we're not out of the woods if we consider open fits.

And I think one thing, if you wanna say, what can I take home out of this? Okay, while we know MPO may or may not have been a contributing factor for some of these conditions, again, you know, if you do real measures, measure LDLs, maximize that patient's usable dynamic range, because you can ruin a perfectly good fitting by having too low of an MPO. So as I'm identifying here, what I think is part of the problem is an input problem. Well, you don't wanna make it an output problem. So if the patient has room to grow, do maximize those MPOs to give that signal room to grow and overall, presumably result in better overall sound quality and speech quality as well.

Here is our contact information. I do not wanna gloss over this. If you go to this website here below, you will have an opportunity to listen to all of the music samples by brand, by the controls. We had one of our participants who, after listening to these music samples for an hour and a half, created his own Spotify list of the full music samples. You really do take some time to go through it, cause you'll appreciate the amount of effort that Emily went into to have a variety of genres and different overall quality in terms of locals and instrumentals, orchestra. Do take some time to do that. I actually did put a place for you to leave some comments if you wanna share any comments that you have about listening to those samples or any sort of experiences that you may have had as a clinician or as a person with hearing loss.

At that, I know we're at the top of the hour. I'm going ahead and look ahead at, it looks like I have one question in the Q&A. Oh, interesting, the question comes from Shannon, I might have missed this earlier, but can you explain oversaturated for a hearing in more detail? Actually I saw this question come up on Emily's side, but I tried to hit it on my side, but I don't think I've done a fair job. The saturation, we could think of it, coming in

from a few places,, the consensus is that it's not the hearing aid microphone. So the idea is is that at some point the input signal is not allowed to grow naturally, if you will.

So we have to use peak clipping or, well, peak clipping if you take any control measures. But if you wanna take control measures, you would have output compression limiting, which a high compression ratio with a high knee point where you're really bringing down those peaks once they exceed a certain input. Okay, so that being said, most hearing aid microphones, the consensus is not gonna be the limiting factor. What I've seen, peak inputs from microphones, about 117 to 119. We know that on the output side of things, the receiver side of things, those are largely controlled by the MPO of your receiver, well, actually the OSPL of the receiver, the physical maximum of the receiver as well as then how the clinician is programming it.

Where I think the saturation's coming in and what has been identified before, particularly with the 16-bit hearing aids is is once the analog signal leaves the microphone, it has to get converted into a digital format, into a series of ones and zeros. And that's what we mean by the digitizer. And the bits, each bit adds six decibels. So wherever you start that, you need a certain minimum bit depth to encode soft speech. So it's not like we started at 0 dBH or SPL. The point is if you only have 16 bits, you can only encode 9,600 whatever, 105, depending upon where you start dB/SPL inputs. And if you don't take control measures to make sure your input doesn't exceed that, then the digitizer is going to encode all signals at that level or the peak level as the same, and basically create those square waves we're familiar with and create a distortion component.

So that's what they mean by saturation. And that's where I think if there is saturation, lots of evidence has been, or lots of arguments have been made, once you go to 24 bits, you eliminated the problem. But I'm not sure, it does seem to, from my analysis, seem to be an ultra low frequency input limiting problem. I do not see any additional

questions at this point, but you do see my contact information, Emily's contact information, do take some time to spend on that website. I think you'll have a much finer appreciation for the differences between the hearing aids, just being able to hear them and what your patients might experience. Thank you so much for your time today and have a great rest of your afternoon, everybody.