



OPEN ACCESS

EDITED BY

Qinglin Meng,
South China University of Technology, China

REVIEWED BY

Rumi Hiraga,
National University Corporation Tsukuba
University of Technology, Japan
Waldo Nogueira,
Hannover Medical School, Germany

*CORRESPONDENCE

Lloyd May
✉ lloydmay@stanford.edu

RECEIVED 16 January 2024

ACCEPTED 28 August 2024

PUBLISHED 19 September 2024

CITATION

May L, Hodges A, Park SY, Kaneshiro B and
Berger J (2024) Designing audio processing
strategies to enhance cochlear implant users'
music enjoyment.

Front. Comput. Sci. 6:1371728.

doi: 10.3389/fcomp.2024.1371728

COPYRIGHT

© 2024 May, Hodges, Park, Kaneshiro and
Berger. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Designing audio processing strategies to enhance cochlear implant users' music enjoyment

Lloyd May^{1*}, Aaron Hodges¹, So Yeon Park^{1,2}, Blair Kaneshiro¹
and Jonathan Berger¹

¹Department of Music, Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA, United States, ²Center for Design Research, Stanford University, Stanford, CA, United States

Cochlear implants (CIs) provide hundreds of thousands of users with increased access to sound, particularly speech, but experiences of music are more varied. Can greater engagement by CI users in the music mixing process mutually benefit them as well as audio engineers? This paper presents an exploratory mixed-methods study that leverages insights from CI users and professional audio engineers (AEs) in the investigation of techniques to enhance CI user music enjoyment and promote empathetic practices in AEs. Analysis of data collected over a multi-step process of surveys, interviews, and iterative cycles of sound mixing and feedback revealed two mixing strategies—namely essentializing and exaggeration of musical elements—common among mixes preferred by CI users. Participant responses also highlighted systemic ableism and audism in the music industry—an industry from which CI users report feeling excluded. Specifically, AEs felt inadequately trained around aural diversity considerations and experienced anxiety around their own aural diversity. In sum, this work contributes to insights into CI user music listening preferences; how AEs approach mixing for an aurally diverse audience; and discussion around the efficacy of CI simulation, user feedback, and AE peer feedback on mix enjoyment by CI users. Based on these findings, we offer several design insights that emphasize the need for customizable music listening tools centered around user agency and enjoyment.

KEYWORDS

accessibility, aural diversity, music personalization, cochlear implant, music appreciation, disability, deaf and hard-of-hearing

1 Introduction

Cochlear implants (CIs) are a form of assistive technology that provides access to sound. As of December 2019, approximately 736,900 registered CIs have been implanted worldwide, with roughly 118,000 of these adults and 65,000 in children in the US (NIDCD, 2021). CIs are electronic medical devices that convert acoustic signals to electrical signals, which are then used to stimulate the cochlea. They consist of an internal part that is surgically implanted into the cochlea, as well as a removable, external part comprising a microphone and electronic processing unit that is attached to the outside of the skull, usually via a magnet. CIs are used as an electronic prosthetic device that provides electric stimulation directly to the cochlea by folks who have been diagnosed as “severely

hard-of-hearing or profoundly deaf”¹ and for whom standard hearing aids may not provide the desired level of hearing assistance.

Current audio processing techniques for CIs are optimized primarily for speech, and perceptual experiences of other complex auditory stimuli such as music vary greatly among CI users (Maarefvand et al., 2013; Spangmose et al., 2019). The frequency resolution of CI users—that is, the number of unique frequencies someone is able to perceive—is approximately 10–20 times lower than listeners who are not Deaf or Hard-of-Hearing (DHH), and the maximum dynamic range—the range between the loudest and most quiet sound a person can perceive—is altered from around 40–80 dB in a traditional hearing auditory nerve response to roughly 3–10 dB of coded dynamic range (Hartmann and Kral, 2004; Zeng et al., 2014). These differences often result in varying perceptions of complex auditory phenomena such as speech in noise and polyphonic music (Limb and Roy, 2014). A CI not only affects how sound and music are transmitted into the cochlea but has also been shown to affect a variety of auditory processing pathways.

CI users’ perception of certain musical features, such as tempo and rhythm, are comparable to those with traditional hearing; however, perception of melodic, harmonic, and timbral information differs (Limb and Roy, 2014). Current techniques to enhance musical enjoyment among CI users include adjusting the signal processing on the device itself; utilizing algorithmic approaches to pre-process or “optimize” a piece of recorded music for CI users; and tailoring aspects of the music at the compositional level (Au et al., 2012; Nagathil et al., 2015; Pons et al., 2016; Buyens et al., 2017; Nogueira et al., 2018; Nogueira, 2019). Additional techniques have been proposed that encode temporal fine structure, such as *SpecRes*, *HiRes*, and *FS4*, as well as utilize phantom stimulation techniques to extend the range of perceivable pitch (Nogueira et al., 2009; Riss et al., 2016; Krüger et al., 2022). While there are certainly merits to these approaches and the technical strides made, they are limited in their recognition of CI listener agency and the diversity of their listening experience, strategies, and goals; and they tend to assume that CI users’ music-listening experiences are both homogeneous and consistent over time. Additionally, these processes assume a passive CI listener with a limited desire to play an active role in their listening experience.

Overall, the topic of CI user music enjoyment has been largely approached through the lens of engineering as a *problem* to be *solved*, particularly in terms of optimization and scalability. Yet this topic could also be approached as an opportunity for co-exploration between both audio engineers and CI listeners, thereby developing tools centered around user agency and joy facilitation. This starts with a better understanding of the needs and desires of CI users, often achieved through empathetic design practices. Accordingly, we present findings from a mixed-methods study that approaches empathy generation through the lens of sound design. We recruited professional audio engineers (AEs) and CI users, who co-contributed insights into desirable combinations of audio processing strategies through a multi-step process of surveys,

interviews, and iterative sound mixing and feedback cycles. Audio engineers incorporated CI user feedback and peer feedback into their mixes, as well as an audio plug-in approximating the listening experience of a CI user. This exploratory study of music pre-processing strategies to enhance music enjoyment for CI users is the first phase of a larger project ultimately aimed toward creating a tool that highlights the autonomy, curiosity, and mastery of one’s own listening experience by allowing users to personalize said experience through multiple layers of controllable, instrument-level processing.

The contributions of this work are five-fold: encompassing insights into CI user music listening preferences; how AEs approach mixing for an aurally diverse audience; the effect of CI simulation on the mixing process; the efficacy of CI simulation, CI user feedback, and AE peer feedback on mix enjoyment by CI users; and technical insights regarding the combination of audio processing tools and techniques used to create more enjoyable mixes for CI users.

2 Background and related work

2.1 Cochlear implants

CIs are devices comprising a small wire that is surgically inserted into the user’s inner ear (cochlea) and connected to a removable receiver placed on the scalp. A microphone in the receiver picks up acoustic sound waves which are then processed into electrical currents that are sent down the wire and excite specific areas of the user’s cochlea.

Notably, the CI limits both the range and resolution of perceivable frequencies. While traditional acoustic hearing usually ranges from approximately 20–20k Hz, CIs have an input range of roughly 150–8k Hz. Additionally, the number of uniquely perceivable frequencies is greatly reduced due to the relatively small number of frequency bands, generally between 12–24, that CIs use. This difference in frequency range and resolution often leads to differences in CI users’ sound perception compared to non-CI users, particularly for broad-spectrum sound such as music (Limb and Roy, 2014). Particularly, previous work has found that CI users often report similar rhythm perception to those with traditional hearing, but that aspects of music that rely on spectral and temporal fine structure cues, such as pitch, melody, and timbre, are not well represented by the CI and CI users can have great difficulty perceiving these (Sorrentino et al., 2020).

CIs were introduced in the late 1950s and saw more extensive use—and controversy—in the USA in the 1980s and 1990s (Sparrow, 2010). Many Deaf advocates argued that CIs were the embodiment of audism, a discriminatory perspective viewing hearing as inherently better or preferential to deafness. This controversy was heightened due to many CI recipients, particularly Deaf adults, having limited improvement in their quality of life following the expensive surgery and often painful recovery. However, the device saw less controversy and use by those not exposed to Deaf Culture, such as deaf children born to a hearing family or those whose hearing changed later in life. By the 2000s, attitudes toward CIs had slowly shifted toward centering the agency

¹ According to the National Institute on Deafness and Other Communication Disorders, <https://www.nidcd.nih.gov/health/cochlear-implants>.

of the DHH individual to decide for themselves (Wheeler et al., 2007).

2.2 Recording processes, standards, and technologies

In Western music traditions, ranging from jazz to pop and classical, the vast majority of *recording* processes share common features (Hepworth-Sawyer and Golding, 2011). Once a piece of music is written and rehearsed, it will be performed acoustically and recorded with a microphone and/or programmed into a computer application to render the desired audio. Either process results in a multi-track recording of the song, where different instruments and musical elements occupy individual audio channels and can therefore be individually edited and processed.

The *mixing* process entails adjusting various properties of the individual tracks, including volume and stereo panning, and the application of various audio processing techniques such as frequency filtering or the addition of artificial reverberation. Following this, the multi-track project, often containing dozens of channels, can be rendered into *stems*, whereby all channels of an instrument family are typically collapsed into a single stereo channel. For example, while the multi-track project may include multiple individual channels for each element of a drum kit, the stem representation will likely have a single “Drums” channel. While not necessary for every mix, stems are extremely useful for archiving, remixing, and adapting purposes, such as removing singing from a certain section of a pop song used in a television commercial.

Recording, mixing, and mastering music often requires empathetic listening and perspective-taking at various points in the process (Hepworth-Sawyer and Golding, 2011). Specifically, in productions with dedicated mixing and/or mastering engineers, engineers are asked to balance many factors when making technical and aesthetic choices regarding the mixing and mastering processes, including the desires of the composer and the producer/record label; the engineer’s prior training and experience; and their assumptions about when, where, and by whom the music will be listened to. Additionally, the engineer often considers whether the mix will “translate” well through various music-listening apparatus including headphones, car stereos, and even smartphone speakers (Gibson, 2019). Viewing the mixing process through the lens of design, it becomes apparent that common design techniques are used, namely empathizing with end users and listeners; iterating; and relying on domain-specific knowledge (School, 2018). Similar to more traditional design, audio engineering also struggles with the role, degree, and type of empathy employed, as various competing factors are at play and the end users/listeners are so numerous, heterogeneous, and abstracted during the process that empathetic perspective-taking becomes challenging (Bennett and Rosner, 2019).

The current audio standards and technology climate greatly increase the possibilities of designing tools that increase the flexibility and control users have over their music listening experience. Audio is no longer confined to physical mediums and inflexible stereo formats, and the rapid growth of spatialized

sound formats like *Dolby Atmos* illustrates both the technical and experiential flexibility offered by going beyond the standard stereo mix. One such extension of potential interest is the multi-track or stem rendering of a piece of recorded audio. The prevalence of multi-track and/or stem recordings being released, leaked, or algorithmically separated illustrates a strong desire from music makers, remixers, and general listeners to engage with musical material through personalization and experimentation. Examples include artists from *Nine Inch Nails*² to *Charli XCX*³ releasing stems and encouraging remixes by fans; a popular catalog of bootleg stems leaked from cracked video games such as *Rockband* and *Guitar Hero* (McGranahan, 2010); the rapid increase in quality and availability of source-separation algorithms such as Spleeter (Hennequin et al., 2020); and the release of the Kanye West *Stem Player*⁴ and stem remixing platforms like *Myxstem*.⁵ However, so far none of these has created a standard format for stem releases, nor have they centered aural diversity in their design and implementation—even as aurally diverse communities such as CI and hearing aid users could stand to both benefit from and contribute to such systems.

2.3 CIs and music enjoyment

CI users are a diverse and heterogeneous group of people. For example, music appraisal varies greatly between CI users, with only some reporting a decline in music enjoyment and listening time post-implantation (Gfeller et al., 2000; Mirza et al., 2003). While some CI users are optimistic about navigating the constraints of the CI to enjoy music, others express frustration with their experience and a reluctance to engage with music (Gfeller et al., 2000, 2019). Therefore, for some CI users, interest and time spent listening to music have declined rapidly. However, even among these users, the reasons for this decline are not universal. As noted in a recent review study by Riley et al. (2018), some literature indicates that a reduction in music appreciation is caused by a reduction in general auditory perception fidelity, while other studies indicate that appreciation does not always reflect perceptual ability but may depend more on personal, situational, and emotional factors. Previous studies have also shown a strong correlation between music perception and higher speech perception, age, and pre-implantation experience with music. In all, it is clear that experiences of speech, music, and other sounds vary based on the individual CI user (Kohlberg et al., 2015; Riley et al., 2018).

For music specifically, a number of factors impact CI users’ music enjoyment. Familiarity with a given song is a large contributing factor to reported enjoyment (Riley et al., 2018). In addition to prior personal experience, many factors of the music itself can be adjusted to improve music enjoyment for some CI listeners. For instance, music can be simplified by accentuating features that are generally more easily perceived such as rhythm,

2 <http://www.ninremixes.com/>

3 https://twitter.com/charli_xcx/status/1279090789419814913

4 <https://www.stemplayer.com/>

5 <https://myxstem.com/>

voice, and low-frequency information while reducing other, more harmonically or spectrally complex elements (Kohlberg et al., 2015; Nemer et al., 2017). However, we still lack an evaluation of music appreciation due in part to the heterogeneity of surveying and protocols used for CI studies (Riley et al., 2018). While CI simulations have been used as a proxy in the study of audio-related phenomena with some success, it is notable that musical enjoyment cannot be accurately studied in this way (Wright and Uchanski, 2012). This is likely due to people using the CI simulator not being accustomed to the sound of the CI, thereby not finding enjoyment in such a radically different music-listening experience. Additionally, the factors that impact CI users' enjoyment of music, such as subjective and stylistic familiarity, likely do not impact people using the CI simulators in the same way.

Music pre-processing personalization was recently explored with the use of a simplified web-based mixing console where CI users could granularly adjust parameters of the music, including increasing the relative volume of percussive sounds, applying either low- or high-shelf filters, and adding artificial reverberation (Hwa et al., 2021). Approximately 80% of the 46 CI users in Hwa's study expressed interest in "having access to technology to mix [their] own music," with 90% of participants also expressing interest in "technologies that could afford [them] greater control over the music [they] listen to"—indicating a clear desire for an increased agency of musical experiences among CI users. Additionally, previous work has shown that remixing preferences among CI users are diverse, further supporting the hypothesis that personalized mixing may be helpful to CI users (Pons et al., 2016).

2.4 Empathy in design

Empathy is an affective and cognitive state, with several definitions. In a 2016 review article on empathy, Cuff et al. (2016) defined empathy as:

An emotional response (affective), dependent upon the interaction between trait capacities and state influences. Empathic processes are automatically elicited but are also shaped by top-down control processes. The resulting emotion is similar to one's perception (directly experienced or imagined) and understanding (cognitive empathy) of the stimulus emotion, with recognition that the source of the emotion is not one's own (Cuff et al., 2016).

A notable distinction is made between sympathy and empathy. Whereas empathy may be described as "feeling *as*" another, sympathy would be "feeling *for*" another (Hein and Singer, 2008). For example, if Sara's friend's father passed away, empathy would elicit a feeling of grief and sadness in Sara, even though she had never met her friend's father. However, sympathy would elicit a feeling of concern for her friend, but not feelings of grief.

While the cognitive and affective aspects of empathy often operate together, it is useful to define these aspects separately. Empathic understanding, or aspects of empathy that involve primarily cognitive components, is the ability to *understand* another's feelings and is related to theory of mind. Affective empathy is focused on the *affective experience* of the elicited

emotion, which does not have to include an understanding of the underlying feelings (Rogers, 1975; Blair, 2005).

Empathy is often discussed in design thinking literature as an important aspect of understanding the user, allowing the designer to "get closer to the lives and experiences" of the user (Koupric and Visser, 2009). Therefore, various tools and techniques were developed to elicit empathy in designers. Specifically, the "empathic horizon," the designer's "range of understanding of user experiences in different contexts," was thought to be expanded by activities such as focus groups, user interviews, user shadowing, surveys, etc. (McDonagh-Philp and Denton, 1999). In the field of designing for (and with) people with Disabilities, the technique of disability simulation for empathy generation, such as sighted designers wearing a blindfold during an activity to simulate vision loss, remains controversial. Some argue that these activities mainly elicit sympathy, pity, and a shallow, unnuanced understanding of the experiences of people with Disabilities (French, 1992). However, others state that the effect of context and contact with people from the community whose experience was being simulated has led to increased empathy generation (Burgstahler et al., 2004). In a meta-analysis by Flowers et al., simulation activities had little positive or negative effects on participants, even when controlling for various moderator variables such as age, type of disability being simulated, and method of simulation (Flower et al., 2007). In the case of CI simulation, Embøl et al. (2021) found that a virtual reality CI simulation elicited feelings of deeper understanding in non-hearing impaired parents of children with CIs. Therefore, while Disability simulation techniques can indeed cause harm when used without sufficient care or guidance, there is some evidence that these techniques can help elicit feelings of understanding or empathy in specific situations.

3 Materials and methods

3.1 Ethics statement

This study was approved by the Stanford University Institutional Review Board (Protocol # 61854). All participants provided written, electronic informed consent prior to participating in the study and were invited to ask the research team questions either verbally or via email before giving consent. Participants were recruited from 5th January 2022–1st August 2022.

3.2 Participants

We recruited 10 audio engineers (AEs) and 10 expert listeners with cochlear implants (ELs).⁶ All ELs and AEs received an electronic gift card as an honorarium for their participation in the study. AEs were screened to be 18 years or older, self-identified as proficient in the *Reaper*⁷ digital audio workstation (DAW),

⁶ The phrase "expert listeners" was chosen to highlight the lived-experience expertise that CI users have regarding CI listening and perception. As this study includes CI simulation, we found it of even greater importance to highlight the expertise of the CI users.

TABLE 1 Demographics of audio engineer (AE) participants.

ID	Age	Pronouns	Professional experience	Primary genre(s) of work
AE1	43	He/Him	Mixing, Recording engineer; Composer	Rock, metal
AE2	26	He/Him	Mixing, Recording engineer; Composer	Rock, jazz, church sermons, short films
AE3	43	He/Him	Mixing engineer	Hip-hop, TV, trailers
AE4	46	He/Him	Mixing engineer; Composer	Pop, rock, spoken word, video tutorials
AE5	33	He/Him	Mixing engineer	Hip-hop, RnB, live church services
AE6	39	He/Him	Mixing, Mastering, Recording engineer; Composer	Hip hop, funk, afrobeat
AE7	57	He/Him	Mixing engineer	Jazz, funk
AE8	42	He/Him	Mixing engineer; Sound designer	Video game sound tracks
AE9	35	He/Him	Mixing and Mastering engineer; Composer; Sound designer	Video game sound tracks
AE10	40	He/Him	Mixing engineer	Punk

comfortable with written and spoken English, and had mixed and/or mastered at least two hours of publicly released music. Recruited AEs worked in varied genres ranging from hip-hop to metal, ambient, and video-game soundtracks; were an average of 37 (SD 7.7) years old; and all self-identified as male.⁸ Information about AEs is summarized in Table 1.

ELs were screened to be 18 years and older, comfortable with reading written English and communicating through either verbal English or American Sign Language, and had at least 1 year of experience using their CI for an average of over 5 h per day. Of the 10 ELs recruited, one used a CI on both ears, eight used a CI on one ear and a hearing aid on the other, and one used a CI on one ear and no hearing aid on the other. ELs were an average of 59 (SD 19.4) years old; four self-identified as male, six as female, and none as non-binary or non-conforming. Seven of the ELs had previous experience playing musical instruments for more than 5 years. Demographics, CI information, and music-listening and instrument experience of ELs are detailed in Table 2.

3.3 Stimuli

The music segments for the mixing tasks were drawn from the following genres: folk, hip-hop, electronic dance music, soft pop, and funk. This selection was aimed at representing a broad range of diverse popular musical styles, as well as ensuring a wide range of dynamics, frequencies, timbres (both acoustic and synthetic), and tempi were present. An additional single recording of speech-in-noise was presented. The durations and genres of the clips were chosen to balance timbral and stylistic diversity while managing practical considerations of the AEs' time. The duration of each music or speech excerpt was approximately five seconds. This was selected to allow for a variety of genres to be used for stimuli, while remaining feasible for the AEs to mix multiple times, as longer clips would place an additional burden on the volunteer AEs. The funk clip was selected to be extended to 30

seconds to allow for at least one of the clips to allow for additional contextualization of timbre and instrumentation changes within the same clip. with the exception of the funk clip, which was chosen to be 30 seconds to fit the phrase length and to contextualize timbre and instrumentation changes. Stimuli were acquired from the *Cambridge Music Technology's "Mixing Secrets For The Small Studio—Free Multitrack Download Library"*.⁹ A full list of stimuli is available in Supplementary Table 1. Participants were asked to listen to the stimuli using the technologies and CI mappings they would ordinarily use when listening to music for enjoyment. These differed from participant to participant and are captured in Table 2.

3.4 Study design and data collection

After screening and consenting, each participant was given a study identification number that they used to complete all surveys and submissions. Participants began the study by completing an online pre-study survey and initial interview. AEs were then asked to complete the mixing and survey submission portion of the study before completing an exit interview and final survey. ELs were asked to provide written feedback on an AE's mix and complete an online mixing preference survey. The steps of the procedure are summarized in Figure 1 and discussed in more detail below.

3.4.1 Surveys

All surveys were administered digitally using *Qualtrics*.

The EL screening survey consisted of questions regarding participants' demographic information, their history of hearing assistive technology use, music listening habits, genre preferences, prior experience playing music, and their preferred mode of music playback (headphones, direct-to-CI streaming, stereo speakers, etc.). The AE screening survey contained demographic questions as well as questions relating to their audio engineering practice and history such as the genre they primarily work in, the audio equipment they primarily use, the role of feedback in their mixing practice, and how often they listen to their mixes on different audio systems before finalizing a mix.

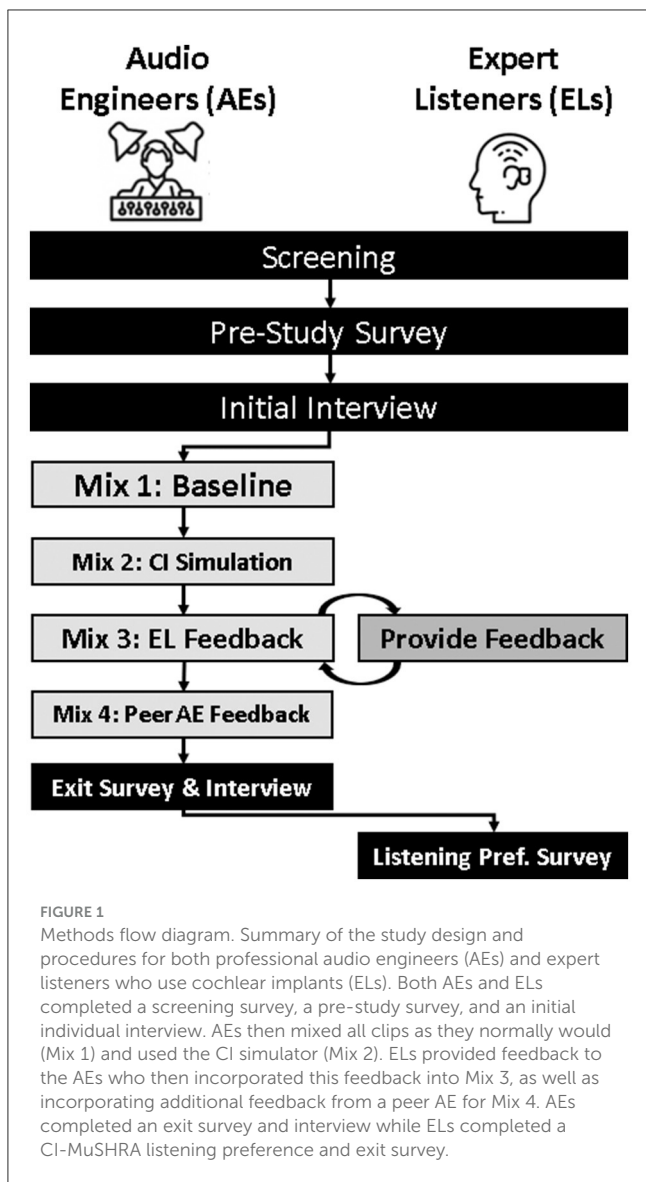
⁷ <https://www.reaper.fm/>

⁸ The unfortunate and notable gender disparity in the field of audio engineering made it extremely challenging to recruit a more gender-diverse cohort (Brereton et al., 2020).

⁹ <https://cambridge-mt.com/ms/mtk>

TABLE 2 Demographics of expert listener participants.

ID	Age	Gender	Hearing assistive tech	Current weekly music listening	Pre-CI weekly music listening	Preferred genre	Preferred listening method	Instrument experience
EL1	61	She/Her	CI (L), CI (R)	1–2 h	1–2 h	80s, Beatles, classical	Stereo speakers	Piano, and Flute (20+ years), No longer play
EL2	75	He/Him	HA (L), CI (R)	3–7 h	3–7 h	Classical	Direct to CI streaming, Stereo speakers	Piano (high school), No longer play
EL3	19	She/Her	CI (L), HA (R)	1–2 h	3–7 h	60–80s pop, classical, jazz	Direct to CI streaming, Stereo speakers	Piano and drums, not played since CI
EL4	62	He/Him	CI (L), HA (R)	0 h	8–12 h	Church music, other simple, single-melody pop music. Used to listen to classical.	Direct to CI streaming, Stereo speakers	Guitar, piano, voice, No longer play
EL5	72	He/Him	HA (L), CI (R)	0 h	1–2 h	Classical, 60s pop	Direct to CI streaming, Stereo speakers	Trombone and euphonium (64 years), stopped for 13 years, restarted recently
EL6	78	She/Her	HA (L), CI (R)	1–2 h	8–12 h	Opera, Pop	Direct to CI streaming	Piano as a child, No longer play
EL7	51	She/Her	CI (L), HA (R)	8–12 h	12+ h	Jazz, classical guitar, solo flute or violin, adore 80s pop but need to listen to it as instrumental-only now	Laptop or phone speakers, Stereo speakers	Double bass (38 years), still play, occasional piano and ukulele
EL8	36	He/Him	CI (R)	3–7 h	3–7 h	Pop, rock, contemporary rock	Direct to CI streaming, Stereo speakers	None
EL9	77	She/Her	HA (L), CI (R)	0 h	1–2 h	Pop or opera but only music they already remember	Stereo speakers	None
EL10	68	She/Her	CI (L), HA (R)	3–7 h	3–7 h	Pop and R&B	Direct to CI streaming	None



In their pre-study surveys, both AEs and ELs completed the two hearing attitudes questions, namely “How do you feel when thinking about your hearing changing in a noticeable way?” and “How do you feel when thinking about other people whose hearing is different from yours?” Additionally, AEs were asked to rate “How important is feedback to your process?” on a 1–5 scale, with 1 being “not important at all” and 5 being “very important.” ELs were asked to rank between four mixing conditions that would produce “the best audio experience” for (1) “CI users” and (2) “a broader audience.” Both groups of participants were also asked to rate their level of agreement with three statements reflecting different emotions and potential anxieties, namely:

1. I am nervous when I think about my hearing changing in a substantial way.
2. I am concerned for my career when I think about my hearing changing in a substantial way.
3. I am concerned for my emotional well-being when I think about my hearing changing in a substantial way.

As these statements were designed to evaluate hearing-ability-related anxieties across the study, AEs completed these questions in the pre-study survey as well as after each of the four mixes (Baseline, CI Sim, EL Feedback, Peer AE Feedback). Participants responded on a 1–5 scale, with 1 being “not at all” and 5 being “extremely.” AEs also reported how deeply they considered the diversity of hearing abilities in the mix they most recently completed in their previous mix on a scale of 1 (“not at all considered”) to 5 (“very much considered”) upon submission of the mix.

3.4.2 Mixes

In this context, “mix” refers to the practice of using various processes to adjust the volume and other acoustic parameters of different instruments and elements of a single piece of recorded music in an attempt to produce a final result that is technically and aesthetically aligned with the artist or composer’s intent. The AEs were asked to mix each of the six short clips of audio in the following rounds:

Mix 1, Baseline: Create a baseline mix, as if mixing for commercial release.

Mix 2, CI Sim: Re-mix “Mix 1” through a CI simulation plug-in.¹⁰

Mix 3, EL Feedback: Re-mix “Mix 2” after receiving written feedback from an EL with the option for additional email correspondence.

Mix 4, Peer AE Feedback: Re-mix “Mix 3” after receiving written feedback from a peer AE who is aware of all the details of the study.

The CI simulation employed in the study used a 16-band noise vocoder, a simulation technique used in previous studies (Mehta and Oxenham, 2017; Goldsworthy, 2019). This method was selected as it has been shown that noise vocoder simulations increase the difficulty of speech perception tasks (Whitmal et al., 2007; Chen and Lau, 2014), and that this increased difficulty may better reflect the uncertainty of the simulation better than other techniques, such as a tone vocoder. While it was emphasized to AEs that these simulations are coarse and do not perfectly model the perceptual experience, the resultant mixes in round 2 would likely be different if an alternate simulation strategy were used. Future work is needed to better understand this relationship.

To increase the consistency of data for analysis, the AEs were asked to mix using only default *Reaper* audio effects and processes; no third-party plugins or analog outboard effects units were permitted. Once an AE was satisfied with the mix in a given round, they submitted the *Reaper* project file to the research team (who analyzed the project file and rendered the audio) and completed a short survey regarding the most recent round of mixing.

3.4.3 Feedback

ELs and the peer AEs were sent anonymized *SoundCloud* links to stream each mix for which they were asked to give feedback.

¹⁰ While it was emphasized to AEs that this was a very coarse approximation, the results of mixing in round 2 may differ if AEs used a different simulation method, such as a tone vocoder, and future work is needed to explore the effect of simulation choice on AEs mixing decisions.

They then provided feedback via an online survey by answering three open-ended prompts: (1) “What I enjoyed,” (2) “What could be improved in the clip for me to enjoy it more,” and (3) “Additional questions and comments I have.” This feedback was compiled into an anonymized document and sent to the corresponding AE, who then had the option to reply or ask additional questions via email. All email exchanges were anonymized and facilitated by the research team.

3.4.4 Interviews

All interviews were conducted by researchers via *Zoom*; each interview lasted approximately 30–45 min and consisted of one to three members of the research team and an individual participant. Interviews were transcribed and anonymized for analysis, after which the original recordings were deleted.

All participants were interviewed at the beginning of the study. In EL interviews, the research team first communicated the goal of the study and various steps involved, after which ELs were invited to ask clarifying questions to ensure they understood all steps of the study protocol. We then used a semi-structured approach with questions to cover the following topics with the EL: Personal relationship to hearing/hearing status; perceived effectiveness of CI simulation or feedback in aiding the AEs to produce more pleasing mixes; perceptions of the music industry and the role of an AE; and the degree to which they anticipated the AEs would sincerely take their feedback into account.

AE initial interviews followed a similar structure, with additional onboarding time allocated for them to watch a 5-min video overview of the mixing project files, with emphasis placed on the custom audio routing within the project files as well as the restrictions placed on third-party plug-in use. After AEs confirmed they understood the tasks of the study, they were also engaged in a semi-structured interview with questions from the following topics: The AE’s usual mixing process including defining goals, reference tracks, and use of external feedback; the AE’s guess of what the simulator might sound like; how they anticipate the feedback process will personally feel; and how helpful they think the feedback will be.

Exit interviews were completed by AEs to gain a deeper understanding of both the technical and empathetic processes they underwent during the study. These followed a similar format as the initial interviews, however with the questions rephrased as reflections of the AEs’ experience in the study. Additional time was given to AEs at the end of the study to explain the broader purpose of the study, explain that the lived experience of CI users is very different from what the simulator could approximate, and finally invite them to ask any remaining questions they had.

3.4.5 Assessing listening preference using a modified CI-MuSHRA

Following the completion of all mixes by all of the AEs, ELs were asked to rate their enjoyment of a subset of these mixes using a re-purposing of the CI-MuSHRA (Multiple Stimuli with Hidden Reference and Anchor for CI users) framework for CI users (Roy et al., 2012). MuSHRAs are used in the field of auditory perception, specifically in the assessment of audio quality (Series, 2014; Caldwell et al., 2017). The structure of the test allows for

the establishment of a known point of comparison (the labeled reference) as well as a built-in check to ensure participants are adequately attending to the task (the hidden reference and anchor). In the present paradigm employed, ELs were asked to rate each of five unlabeled clips against a reference clip on a 0–100 scale, where 50 indicated enjoying the clip as much as the reference, 100 mapped to “enjoy substantially more,” and 0 indicated “enjoy substantially less” as shown in Figure 2. The modified MuSHRA contained two trials for every genre and the order of mix presentations was randomized. The “matched-feedback” trial contained only mixes by the AE to whom the EL gave feedback, with mix 1 serving as both the labeled and hidden reference. The “unmatched-feedback” trial contained the same reference and hidden reference as the “matched-feedback” trial; however, mixes 2, 3, and 4 in this trial were each selected from another AE, to whom the specific EL completing the survey did not give feedback.

The modified CI-MuSHRA, therefore, contained one reference and five clips per trial, namely the hidden reference, the hidden anchor, and mixes 2, 3, and 4. The anchor was constructed by applying a 1 kHz lowpass filter to the reference mix and then introducing uniformly distributed white noise until a signal-to-noise ratio of 16 dB was reached, the same anchor generation strategy used by Roy et al. (2012). All audio was normalized to -30 dB RMS to control for loudness differences between the unmastered audio mixes. In addition to the modified MuSHRA ratings, ELs were asked to identify specific attributes of the mixes that they rated most highly.

3.5 Analysis

3.5.1 Analysis of survey data

We visualized quantitative survey results for interpretation. Given the small sample and exploratory nature of the present work, we chose not to perform statistical analyses. In visualizing ordinal response data, we added a slight jitter to individual data points for greater visibility of individual observations.

3.5.2 Analysis of interview and feedback data

Three reflexive thematic analyses were conducted (Braun and Clarke, 2019): one for combined AE interviews, one for EL interviews, and another for EL written feedback they provided to the AE during the third mix. Researchers first familiarized themselves with all EL and AE data and then generated codes for each of the three analysis groups, separately identifying all relevant topics and sentiments that emerged. These codes were then grouped into initial themes. The researchers went through several iterations of comparing developed themes against the existing data to ensure that all relevant aspects were captured and that themes were adequately combined or separated to best represent the data. The two sets of data were coded once the themes and descriptions for each were finalized. All quoted materials are given with reference to the participant identifiers listed in Tables 1, 2.

3.5.3 Analysis of modified MuSHRA data

All clips were rated with respect to the baseline reference mix using a modified MuSHRA framework on a scale of 0–100, with

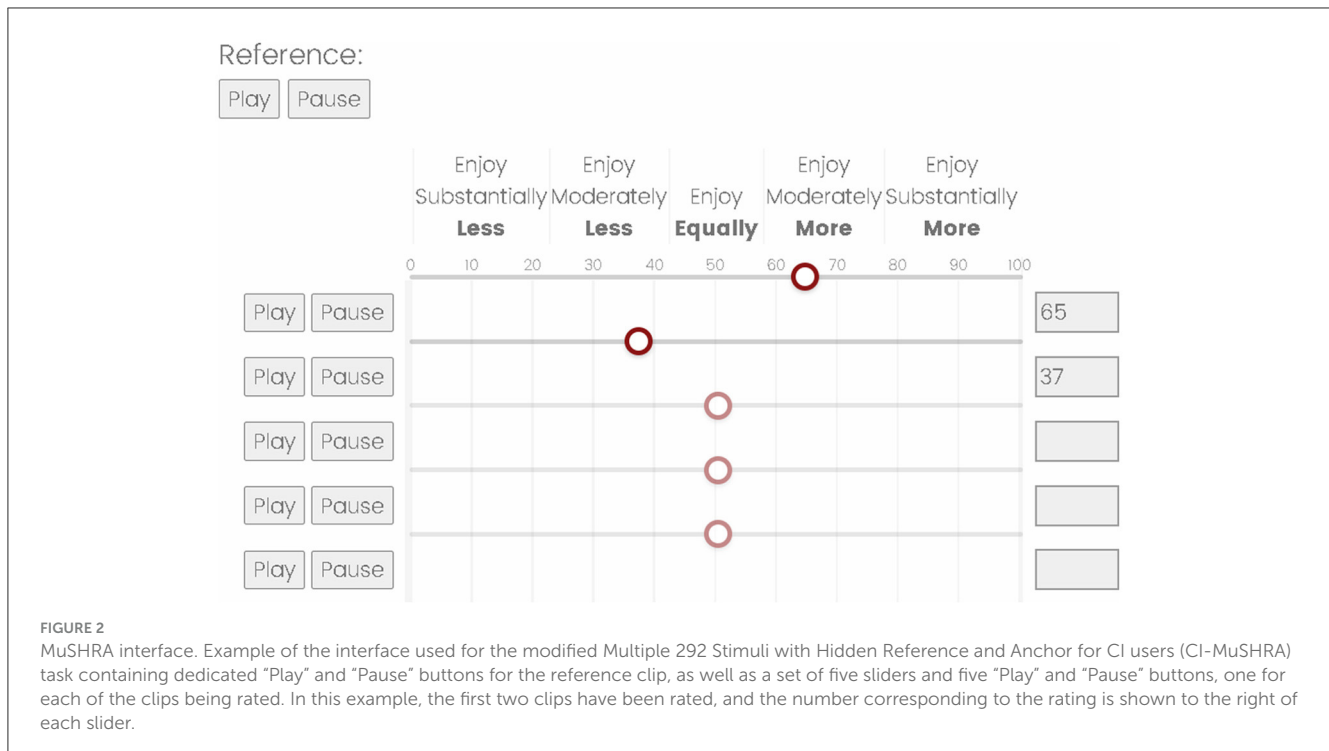


FIGURE 2 MuSHRA interface. Example of the interface used for the modified Multiple 292 Stimuli with Hidden Reference and Anchor for CI users (CI-MuSHRA) task containing dedicated “Play” and “Pause” buttons for the reference clip, as well as a set of five sliders and five “Play” and “Pause” buttons, one for each of the clips being rated. In this example, the first two clips have been rated, and the number corresponding to the rating is shown to the right of each slider.

0 being “enjoy substantially less,” 50 being “enjoy equally,” and 100 being “enjoy substantially more,” as outlined in Methods and Materials. Eight of 10 ELs completed the mix preference survey, resulting in a total of 440 rated clips. Of these, 50 ratings were incomplete and removed, resulting in 390 valid clip ratings. After removing ratings for the hidden baseline and anchor mixes, which were used as analytical reference points, 264 valid ratings were made spread across the simulator mix, EL feedback mix, and peer feedback mix categories. Relative improvement for each mix was calculated as the difference between the rating for each clip and the hidden reference. Mixes that were rated higher than the hidden reference were labeled as *preferred mixes* and separated for additional analysis.

For the preferred mixes, data regarding the mix condition and the relationship of EL to the AE who produced the mix (paired for feedback or non-paired) were tabulated and visualized.

The project files for all preferred mixes were opened in *Reaper* and analyzed by a research team member trained as an AE. The researcher documented the plug-in chain used on individual tracks within each mix and the approximate settings of each plug-in. These were then compared to the elements of the mix identified by the ELs as particularly successful, and a list of mixing trends in preferred mixes was synthesized.

4 Results

4.1 EL findings

4.1.1 EL pre-study survey

In pre-study surveys, ELs ranked which of the four mixing conditions (baseline, CI sim, EL feedback, peer AE feedback) they

thought would lead to the best audio experience for CI users as well as a broader audience; results are summarized in Figure 3.

Regarding the best experience for CI users, all 10 ELs ranked the baseline mixing condition last (fourth). Rankings for both the peer feedback and CI simulation conditions were mixed, with the CI simulation condition ranked slightly higher. Notably, EL feedback was ranked first by eight ELs and second by the remaining two. Regarding the best audio experience for a broader audience, ratings were more mixed: Six of ten ranked baseline fourth, and two ranked CI simulation and peer feedback as fourth; EL feedback received a mix of first-, second-, and third-place rankings.

4.1.2 Thematic analysis of EL interviews

We interviewed ELs at the beginning of the study to learn more about their views on topics framing this study. Through thematic analysis of interview transcripts, we identified three main themes.

First, the theme of **Relationship to hearing assistive technology** encompassed responses related to changes in technology, changes in physiological hearing, and CI sound processing characteristics. Changes in technology involved ways in which innovations, updates, and modifications to hearing assistive technologies—including CIs—impacted users’ music-listening experience. Examples included direct-to-device streaming via Bluetooth as well as the ability for users to change and adjust device programs through the use of an app. ELs also commented on how changes in physiological hearing—such as a change in residual acoustic hearing years after implantation, or “when I stopped being able to use my hearing aid [on my non-implanted side]” (EL8)—impacted their experiences. These reports highlighted that the combination of all hearing modalities, not just the CI itself, informs an EL’s music listening experience. However, ELs did note the many

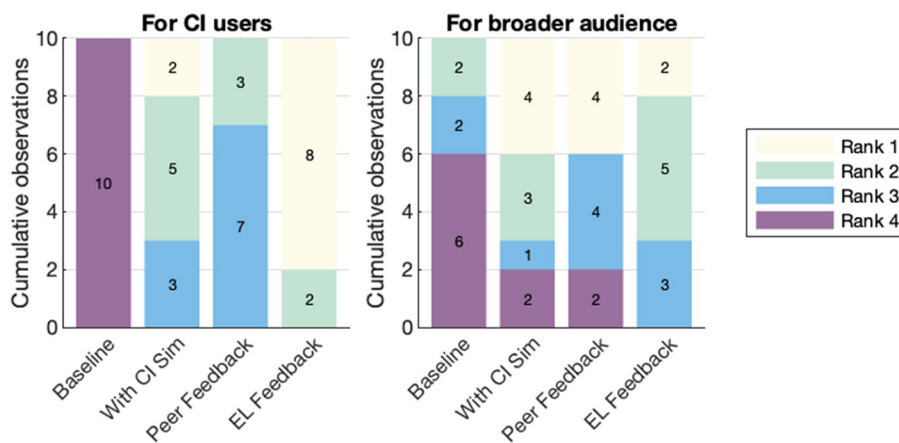


FIGURE 3

ELs' pre-study mixing condition rankings. Stacked bars summarizing ELs' pre-study rankings of which mixing conditions they thought would lead to the best audio experience for CI users and a broader audience. For both contexts, the initial baseline mixes received the lowest rankings overall. For CI users, mixes informed by EL feedback were considered more valuable, while for a broader audience rankings of CI sim, feedback, and EL feedback were more evenly distributed.

ways in which CI sound processing characteristics impacted their music listening experience. Here, interviewees highlighted changes in the timbral qualities of the sound, describing the sound as “robotic and tinny” (EL7), as well as the dynamic qualities of sound, such as the compression and noise gating of broad-spectrum sound such as concert-hall applause (EL2).

The second theme involved ELs' varied **Relationship to music**. Like other groups of listeners, ELs mentioned aesthetic and genre preferences. For example, EL4 stated: “Church music is the only [genre] I voluntarily listen to, I don't bother with anything else anymore”. Also noted by several other ELs, EL4 stopped engaging with particular genres due to changes post-implantation: “After using [the CI], I hear something musical in what you might call single-threaded [or monophonic] music.” Moreover, familiarity with a piece of music played a big role in their music enjoyment, mentioning that if a song is “... familiar, I [can] hear it much better” (EL4). Finally, the role of context of listening was apparent, with some ELs citing preferences for live music and even particular venues: “Old concert halls get the acoustics right” (EL3). Other ELs expressed that multi-modal engagement contexts were particularly helpful such as “dancing with [my partner], [they] lip-synced the words. It really helps” (EL1).

Third, ELs' **Perceptions of the music industry** largely focused on the role of AEs, given the context of the study. One sub-theme expressed by ELs was bi-directional empathy, namely an assumption of empathy on the part of the AEs, stating that “[AEs] must have a lot of empathy to do what they do” (EL9). ELs additionally expressed empathy for AEs, highlighting their knowledge of the challenges present in the AE profession: “[AEs] just have so many people to please” (EL10) and “It seems really complicated, what they're doing” (EL4). However, ELs also expressed a strong feeling of exclusion from the functioning of the music industry. When asked who they suspect the target audience is for recorded music, all 10 ELs expressed that they did not feel part of that target demographic, stating that music is made “not [for] me, not [for] people with hearing impairments” (EL4). Finally, some ELs

expressed a strong desire for increased agency or control over their music listening experience, with one EL stating, “If someone would just attach an equalizer [to my CI] and let me play with it, that might be something productive” (EL3).

4.1.3 Thematic analysis of EL feedback

We conducted a reflexive thematic analysis of both the initial EL feedback and the facilitated email correspondence between AEs and ELs. Only two AE–EL pairs engaged in facilitated correspondence, as the other eight AEs opted to not ask any clarifying or follow-up questions after receiving the initial feedback.

The first main theme identified was **Subjective preferences and affective perceptions**. Here, ELs often referred to the emotion or to affective qualities they experienced or perceived to be communicated by the music, such as “Relaxing to listen to” (EL3), “Lively music... good vibe” (EL1). In addition, ELs expressed preferences for specific styles or genres of music, such as “I like folk [music] better than some of the other genres.” (EL5). ELs also expressed a disinterest in certain styles of music, such as “[Hip-hop] is just not my kind of music” (EL10), with some even opting to not provide feedback for mixes of unpreferred genres.

Another theme in ELs' written feedback was **Sonic and musical characteristics**. ELs made explicit mention of pitch, frequency, or register of musical material stating that “Perhaps the speaker's voice might come through better if [it was] in a different frequency” (EL10), “I do better with instruments in the range of male voices” (EL7). The musical function (melody, harmony, and rhythm) of sounds was another factor in many EL suggestions, such as “Rhythm is a part of music that does get through properly, but this had nice musical threads in the [melody] in the background as well” (EL4). The sonic characteristic of an instrument or sound, i.e., the timbre, was important to many ELs as evidenced in requests that sounds be “less twangy” (EL10), and comments that they “sounded ‘scratchy’” (EL8) and “I'm picking up lots of treble in the strike of the string” (EL7). Veridicality—the perceived correctness of a

recorded sound and its connection to an acoustic instrument—was a factor for some ELs, such as “... [Is] that how the instrument really sounds?” (EL7). Some even expressed uncertainty in instrument classification, e.g., “I believe I heard a guitar” (EL6). ELs made explicit reference to spatialization and their perception of depth and 3D space, commenting that the “beat sounds shallow—depth is lacking” (EL10), and noting “I’m surprised at how much this mix makes the voice stand out in a 3-D way” (EL8). Finally, many overt mentions of the influence of hearing assistive technologies were made, and their impact was very salient to some ELs: “With both aid and CI, I liked this best. With only the CI, it is pretty awful.” (EL10); “Details are usually lost in the noise of the CI” (EL4).

The third and final theme, **Separation of sonic elements**, implicated audio stream segregation as well as clarity and masking. The ELs’ ability to meaningfully distinguish between two or more instruments or musical elements playing at the same time in a mix through audio stream segregation was reflected in favorable comments that “The separation of the instruments was easy to listen to” (EL7), or that “Here I was able to enjoy the complexity of multiple parts intertwined and contrasting, with musicality in the background” (EL4). This was often achieved through balancing the perceived loudness of certain instruments relative to one another, e.g., “The rhythm guitar seems too loud in the mix” (EL7). In addition to separation through perceived loudness, the perceived intelligibility of a musical element relative to the mix through clarity and masking was an important factor: “The voice was easy to identify” (EL7).

4.2 AE findings

4.2.1 AE survey data

In the pre-study survey, AEs answered demographic questions whose answers are captured in Table 1. Additionally, AEs rated “How important is feedback to your process?” on a 1–5 scale, with 1 being “not very important” and 5 being “very important”. Responses to this question ranged from 2 to 5, with a median value of 4.

AEs were asked several questions in both the pre- and post-study surveys. When asked “How effective [will] the ELs feedback be in supporting you to mix music that is more enjoyable for the person giving feedback?”, pre-study ratings ranged from 3 to 5 (median 4) and post-study ratings ranged from 2 to 5 (median 4). This question was repeated but “for a broader audience,” ratings to this ranged from 1 to 5 (median 3) pre-study and ranged from 1 to 5 (median 2) post-study. When asked how effective they think/thought the CI simulation will be/was in “supporting [them] to mix music that is more enjoyable for the person whose hearing is being simulated,” on a 1–5 scale, pre-study ratings ranged from 3 to 4 (median 3) and post-study ratings ranged from 2 to 4 (median 3). The same question was asked, but “for a broader audience,” and pre-study ratings ranged from 2 to 4 (median 3) and post-study ratings ranged from 1 to 4 (median 2).

When discussing their usual process of mixing, AEs highlighted five common approaches they generally took to chronologically organize their mixing efforts. Three AEs used a “bottom-up” mixing approach, focusing on instruments in order of frequency

content from low to high, while one AE used a “top-down” approach, mixing from high- to low-frequency content. Three AEs generally used a “drums then vocals, then everything else” approach, while two used a “broad strokes, put out fires” approach, which started with a rough mix of all the instruments, followed by mixing instruments in order of perceived problems. Finally, one AE described their mixing chronology as “lead instrument first, then everything else.”

Throughout the study, AEs used a 1–5 scale to report their agreement with a series of statements about anxiety and concern related to hearing loss. Individual participants’ agreement at the pre- and post-stages are given in Figure 4. For the statement “I am nervous when I think about my hearing changing in a substantial way,” the median response shifted from 3 to 4 over the course of the study, with agreement increasing for six participants. Median agreement with the statement “I am concerned for my career when I think about my hearing changing in a substantial way” increased from 3.5 to 4, with five participants increasing their agreement between pre- and post-measures. Finally, for the statement “I am concerned for my emotional well-being when I think about my hearing changing in a substantial way,” median agreement again increased from 3.5 to 4; for this statement, seven participants reported higher agreement in the post-measure.

At the conclusion of each of the four rounds of mixing, AEs were asked to use a 1 (“not at all considered”) to 5 (“considered extremely deeply”) scale to report the degree to which they considered the diversity of hearing abilities in the mix they most recently completed. Individual AE ratings are shown in Figure 5. At the conclusion of the first baseline mix, ratings ranged from 2 to 5 (median 3) while in all subsequent mixes, all 10 AEs gave a rating of 5.

4.2.2 AE interviews

We identified five main themes from a thematic analysis of combined pre- and post-study AE interview transcripts. The first theme, involving AEs’ **Attitudes toward aural diversity**, manifested in several ways. Some AEs expressed anxiety around hearing change, highlighting worry in relation to their own hearing changing noticeably, concerns regarding age-related hearing loss, and the longevity of their audio engineering career. AE3 highlighted, “My hearing is good, and it’s got to be, that’s how I [can] work [in the music industry].” These anxieties were heightened by CI simulation before the final debriefing in the last interview. A subset of AEs expressed satisfaction or pride in not having any diagnosed hearing impairments, while others discussed their own personal aural diversity, such as acquiring tinnitus and high-frequency hearing loss related to playing the drum kit (AE7). AEs’ feelings about and opinions of others with hearing impairment consisted of a combination of empathetic feelings—of trying to imagine or embody their experience, including frequent mentions of the CI simulation as both a mechanism and motivation for this (AE4)—and of sympathetic feelings—of pity, sadness, or mourning the assumed impact of others’ hearing impairment on their enjoyment of music (AE2, AE5, AE9), with AE5 stating, “I wish everyone could hear perfectly.” Several assumptions regarding AEs’ perceptions of music enjoyment of DHH people included the

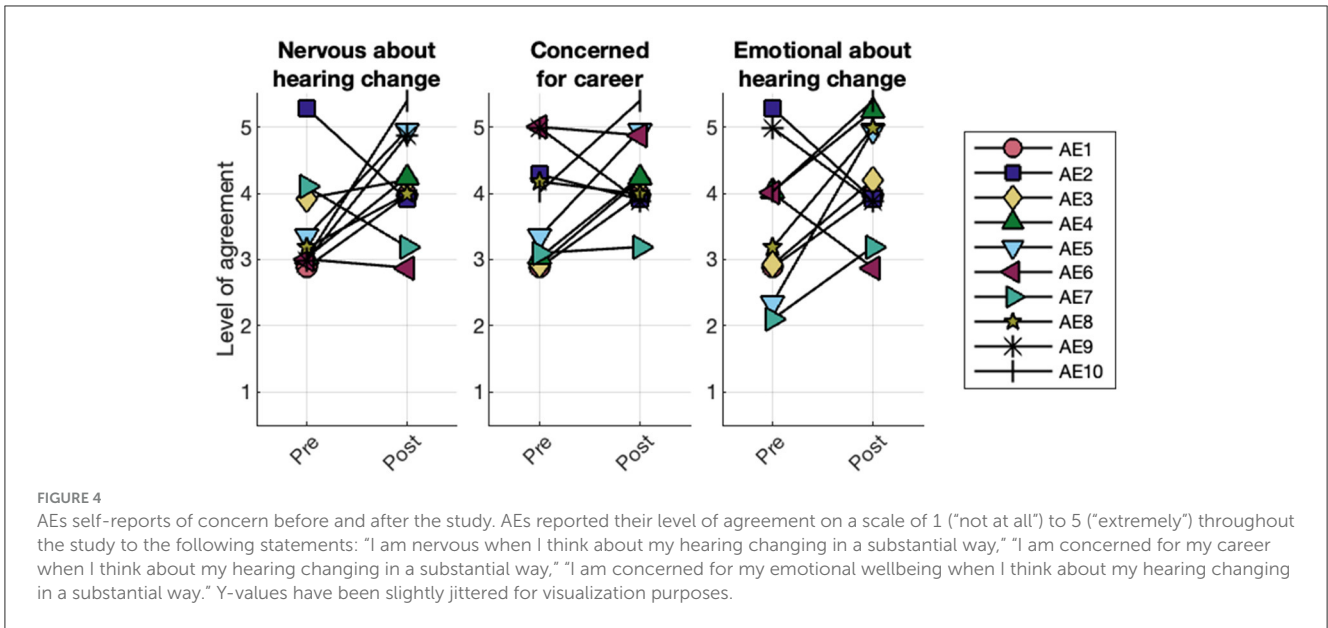


FIGURE 4
 AEs self-reports of concern before and after the study. AEs reported their level of agreement on a scale of 1 (“not at all”) to 5 (“extremely”) throughout the study to the following statements: “I am nervous when I think about my hearing changing in a substantial way,” “I am concerned for my career when I think about my hearing changing in a substantial way,” “I am concerned for my emotional wellbeing when I think about my hearing changing in a substantial way.” Y-values have been slightly jittered for visualization purposes.

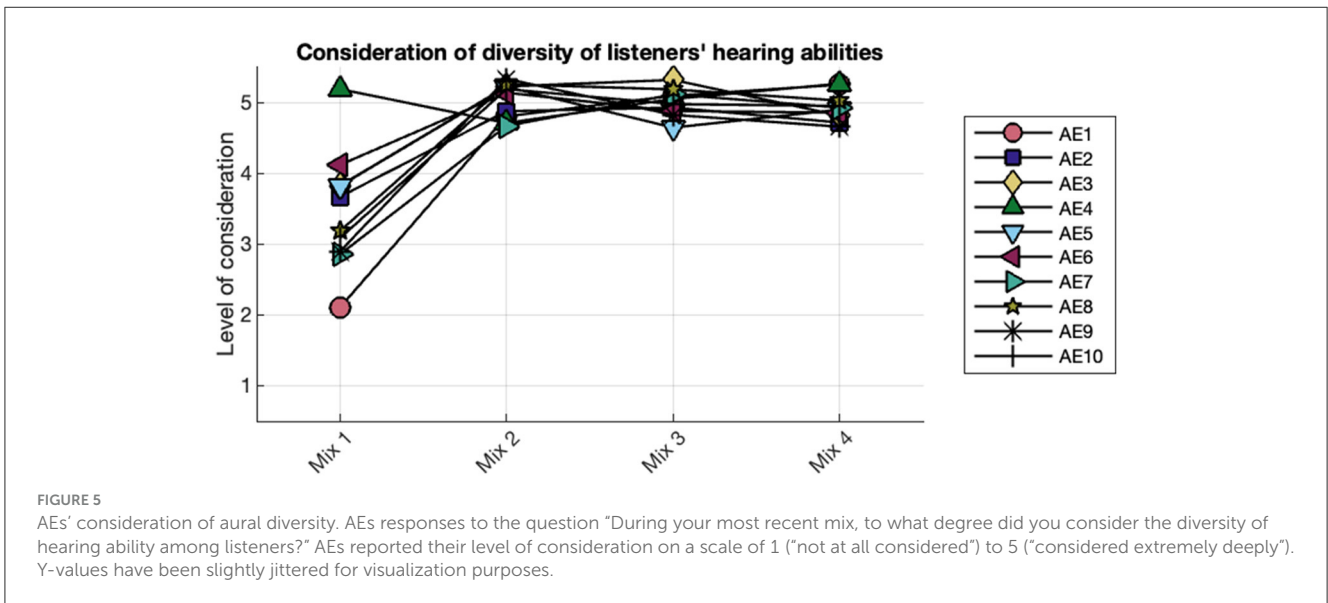


FIGURE 5
 AEs' consideration of aural diversity. AEs responses to the question “During your most recent mix, to what degree did you consider the diversity of hearing ability among listeners?” AEs reported their level of consideration on a scale of 1 (“not at all considered”) to 5 (“considered extremely deeply”). Y-values have been slightly jittered for visualization purposes.

assumed techniques, technologies, and musical curation process used by DHH people in their music engagement practice, such as the use of haptic technologies to augment musical experiences, as well as the assumption that music with strong rhythmic information and/or a large amount of low-frequency information would be preferred, summarized by AE3 stating “I’d guess [DHH folks] just like the bass really blasting.”

The second theme involved the **Novelty of the tasks** being asked of the AEs in the study. Responses in this area highlighted, for example, AEs’ concerns that their prior experience would not fully transfer into this new technical and aesthetic context. AE4 noted, “The rules of my hearing don’t seem to apply. So all the things I’ve learned about mixing go out the window.” The nature of the CI simulation was unexpected by every AE, with reactions ranging from surprise to disbelief: “My initial reaction was like, I don’t think I can make this make sense to me, to my hearing” (AE4). Two AEs contacted the research team to ensure that a technical error

had not occurred. AE2 notably balanced the tension between the simulator and EL feedback, stating that “The simulator is still just [the researcher’s] interpretation of somebody else’s hearing. Having it go through somebody who that’s their reality, that feels more real to me.” Many AEs adopted an active change in perspective to deliberately attempt to shift their listening practices and mixing goals, noting that “It’s like a complete re-orientation and navigating what is the goal” (AE3).

Third, AEs discussed **Technical aspects of the mix**. They expressed concerns regarding the mix translation—namely if the mix would sound as desired across non-assistive technology playback devices, such as phone speakers and car sound systems, and across listening situations, such as in a club or as background music in a restaurant. Technical limitations associated with using only *Reaper* stock plug-ins were noted by AEs, with notable omissions such as the lack of saturation or de-noiser plug-ins (AE5, AE9). Regarding the mixing process itself, two primary techniques

were used by the majority of AEs, namely to *essentialize* the mix by highlighting the most important elements and removing other elements, and to *exaggerate* those important elements through strategies to make the highlighted elements sound larger, clearer, louder, or in a wider range of frequencies. This was balanced by a tension concerning the natural, veridical nature of elements in the mix and a desire to not alter them past the point where they are no longer perceived as that instrument or sound.

Regarding **Listener feedback**, AEs notably highlighted the relationship between technical language and lay language. Individual and situational factors impacted the preference of AEs to receive feedback that is either specific and uses technically precise language, or more general and uses lay language or metaphor. Technical feedback was largely provided through the peer AE feedback and included discussions of feeling validated about both the task's difficulty and perceived effectiveness of their strategy, as well as tendencies to defer to the peer's expertise and implement their advice, with AE2 noting that *"As much as it feels good to be validated by peers, that's just for us."* The AEs' opinions of EL feedback included general deference to the EL's opinion and a strong desire to implement all of the EL feedback. This was held in tension with the perceived vagueness of some feedback and the level of interpretation required when implementing this non-technical feedback. Additional tensions existed between navigating the assumed motivation behind certain feedback as stemming from personal, aesthetic preference or as being a direct result of CI usage.

Lastly, the role of industry and genre **Conventions** on AEs' process manifested in the discussion of the purpose/goal of the mix. This included assumptions and internal negotiations regarding the perceived goal of the mix, which ranged from conveying emotion or artistic intent to pleasing the client or artist, a general audience, and/or oneself as an AE. AE2 noted that *"When mixing, I think of other people but my taste is the deciding factor."* Genre, technical, and cultural norms additionally have a large influence on the learned aesthetic conventions, perceived expectations, and technical strategies and techniques AEs used during the mixing process: *"If it's Hip-Hop, I already know I'm gonna be boosting the bass, boosting the kick"* (AE5). Additionally, AEs noted distinct moments when they knowingly broke from these conventions, such as AE1's observation that *"[without the simulator] my changes would have been more subtle and less drastic."*

4.3 Mix analysis

After all four mixing stages were complete, eight ELs completed the listening preference survey using the CI-MuSHRA paradigm detailed in §2. After removing incomplete ratings, 264 valid ratings of non-reference, non-anchor mixes were made. Preferred mixes were selected if a mix was rated higher than the hidden reference and hidden anchor to which it was being compared. The results of the rating portion of the survey are summarized in [Figure 6](#). Of the 264 valid ratings of non-reference, non-anchor mixes obtained, 100 mixes were found to be preferred over their hidden anchor. Of these 100, 34 were from the EL feedback condition, 29 were from the CI simulation mix condition, and 37 were from the AE peer feedback

condition. Of the 34 preferred mixes of the EL feedback condition, 20 were rated as preferred by the EL who was paired with the AE for that mix, while 14 were rated as preferred by an EL who did not give feedback on that specific mix.

4.3.1 Mixing trends in preferred tracks

At the end of each genre section of the mix preference survey, ELs were asked: "Of the clips of music that you most enjoyed on the previous page, which aspects of the sound or music did you most enjoy?" The reasons for high mix ratings given by ELs were grouped into nine categories, namely:

1. The musical elements were easily distinguished.
2. The volume balance between musical elements was pleasing or appropriate.
3. The beat or rhythm of the music was pleasing.
4. The timbral quality of one or multiple elements was pleasing.
5. The clarity of a particular musical element was pleasing.
6. A sense of depth or space was conveyed.
7. The overall aesthetic coherence of the mix was pleasing.
8. The overall sound of the piece was less dissonant than others.
9. The mix was the "least worst", but nothing was particularly enjoyable.

After analysis of the project files of preferred mixes, several common processing techniques were identified. The relative frequency of these techniques across mixes are shown in [Table 3](#):

1. Heavy compression and additive, band-pass EQ on percussive elements.
2. Drums and lead vocals or lead melodic instruments mixed relatively loud compared to other instruments.
3. The removal or muting of numerous musical elements, particularly those playing pads or chords.
4. Large boost of sub-100Hz frequencies using a low-Q filter or multi-band compression on the bass and/or kick drum.
5. Large (+12dB or more) EQ boosts in the upper mids of lead vocals or lead melodic instruments.
6. Multi-band compression to reduce mid-range frequencies on the drum bus.
7. Octave shifting of bass lines, kick drums, or background vocals one octave higher.
8. Octave shifting of lead vocals or lead melodic elements one octave lower.
9. A pseudo-rearrangement using a long, single-tap delay on a vocal line, essentially moving it to a different, less-busy time in the song, similar to having a backup singer repeat a vocal line after it has been sung by the lead singer.

5 Discussion

Several of the mixing strategies implemented in preferred mixes supported findings from previous work. Musical elements that were deemed non-essential were either removed completely, had their volume reduced, or were subtractively equalized to reduce their spectral complexity ([Nagathil et al., 2015](#); [Pons et al., 2016](#); [Hwa et al., 2021](#)). A preference for retaining percussive elements

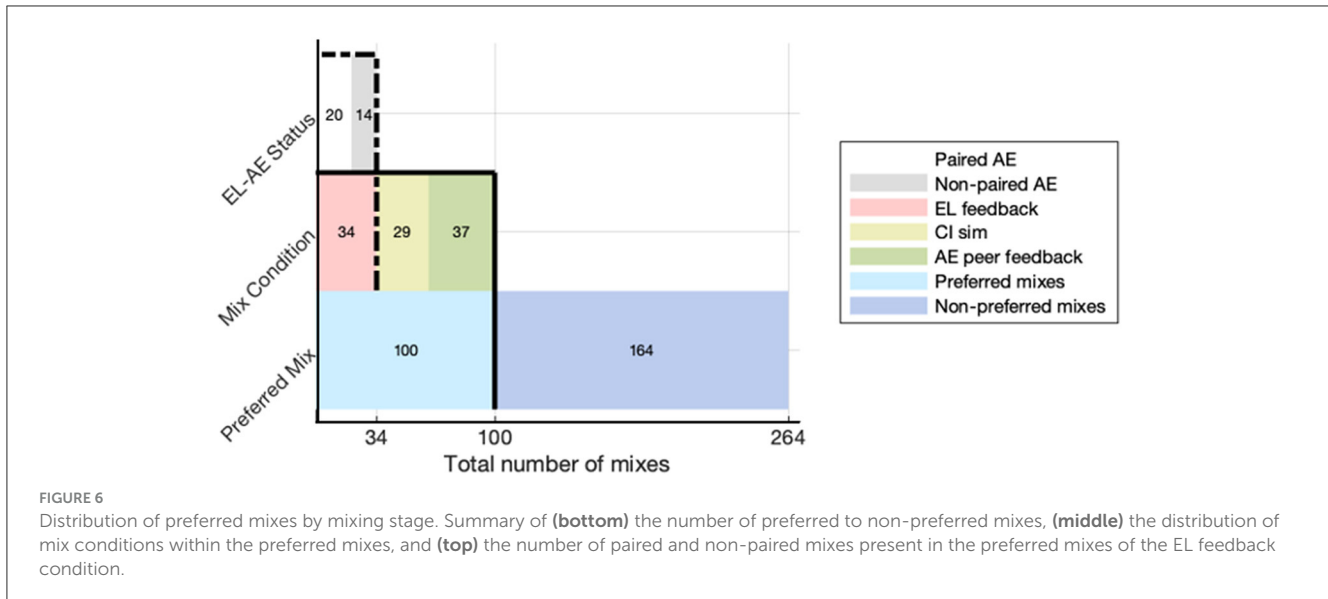


FIGURE 6 Distribution of preferred mixes by mixing stage. Summary of (bottom) the number of preferred to non-preferred mixes, (middle) the distribution of mix conditions within the preferred mixes, and (top) the number of paired and non-paired mixes present in the preferred mixes of the EL feedback condition.

TABLE 3 Frequency of identified mixing technique used in each round of mixing (50 mixes per round), and across preferred mixes (100 preferred mixes).

Mixing technique	Mix 1	Mix 2	Mix 3	Mix 4	Preferred mixes
Compression and EQ on percussion	84%	88%	86%	86%	92%
Drums and/or lead mixed loudly	0%	86%	90%	90%	87%
Mute or remove musical elements	0%	84%	86%	86%	84%
Bass boost	42%	64%	70%	70%	63%
Large EQ boost in upper mids	0%	34%	46%	46%	41%
Multi-band compression, mid-range reduction	22%	24%	20%	20%	19%
Bass or kick, pitch-shifted up	0%	14%	14%	14%	11%
Lead melody pitch-shifted down	0%	6%	10%	10%	7%
Delay vocal line	0%	12%	16%	16%	1%

as well as the voice (if present) was also seen. This supports previous findings in which CI users demonstrated a preference for music with more sparse instrumentation, strong rhythmic features, and musical elements with less spectral complexity (Kohlberg et al., 2015; Nagathil et al., 2015; Nemer et al., 2017; Gauer et al., 2018). Similarly, techniques to enhance certain musical elements, such as bass boosting, were seen in previous work and the majority of the preferred mixes (Buyens et al., 2014; Hwa et al., 2021; Krüger et al., 2022). However, other enhancement strategies, such as the inclusion of additional harmonic saturation and distortion as well as pitch-shifting signals, were not featured prominently in previous literature, suggesting that these techniques may be beneficial in certain circumstances, although further work is required to understand the factors that influence the efficacy of these techniques.

The influence of stylistic familiarity and preference in this study was clear, with three ELs not providing feedback for the Hip-Hop clip and rating all of the Hip-Hop mixes relatively low. Strong individual differences and preferences resulted in greatly varying mixes of the same musical clips, but also in varied ratings among ELs. This supports previous findings that musical factors, such as musical complexity, number of instruments, and presence of voice,

as well as individual factors of the listener, such as previous musical experience and subjective familiarity, influence the music listening experience of CI users (Kohlberg et al., 2015; Riley et al., 2018; Gfeller et al., 2019).

A notable mixing trend that emerged in this study – and has been explored in previous literature—is the addition of artificial reverb. While Hwa et al. (2021)’s work demonstrated a preference for additional reverb for musical sounds among CI users, Roy et al. (2015) identified that reverb had a negative impact on musical sound quality. Notably in our study, no EL requested reverb or used non-technical language to allude to wanting more reverb, space, or echo on any musical elements. Additionally, no AE elected to use reverb in any of the mixes in rounds 2–4. In contrast, several AEs expressed a desire to use a de-reverber plug-in, which would remove reverb and echoes from a signal.¹¹ AEs stated that additional reverb reduced the clarity of some instruments once processed through the CI simulator in the second round of mixing, and no AE elected to use artificial reverb in response to feedback

¹¹ This type of plug-in was not made available to AEs during the study due to the study design of using only *Reaper* stock plug-ins for improved consistency in the analysis of the mixes.

in rounds three or four, even if they had used artificial reverb in their baseline mix. This difference may be due to signal or stylistic differences in the stimuli, the influence of the CI simulator on AEs, or ELs not being aware of reverb as a parameter that could be adjusted in the mixing process.

The CI simulation had a notable impact on AEs with many expressing disbelief or confusion upon hearing the CI simulation.¹² Four AEs stated in interviews that the CI simulation was effective in prompting them to read the ELs' feedback more earnestly and apply more radical mixing strategies as they did not previously understand the nature of audio perception through CIs. This partially supports Embøl et al. (2021)'s finding of increased feelings of understanding among non-hearing impaired parents of children with CIs. The reasons for this may be similar as despite the AEs in the study not having extensive experience or training regarding CIs, they have a non-cursory relationship to hearing ability and audio technology as well as a vested interest in pleasing the "client," in this case the EL. Additionally, the listening experience of CI users is something that many people, including AEs, are generally very unfamiliar with. This unfamiliarity with a specific assistive technology may also be a modulating variable that may impact the efficacy of Disability and assistive technology simulations for empathy generation.

5.1 Design insights

5.1.1 Employing essentialization and exaggeration strategies

Based on a combination of thematic analysis of AE feedback, EL requests, and EL exit interviews, two notable mixing strategies were particularly apparent, namely to *essentialize* the mix through removal or lowering the volume of non-essential musical elements and subtractive equalization, as seen in previous work (Kohlberg et al., 2015), and to *exaggerate* the remaining elements through volume boost, compression, saturation, boosting equalization, and addition of higher or lower octaves. However, deciding which musical elements to preserve and/or exaggerate is not a trivial task as the factors that influence this decision are largely subjective (Pons et al., 2016). The peer AE, who provided feedback on all ten mixes in the final stage summarized this trend: *"This is a different mixing philosophy. I find when mixing you're trying to marry sounds together... When the simulator was on, it would just blur so cohesion wasn't helping, it was making it worse. The key seemed to be to essentialize."*

The common groupings of mixing techniques used in favorable mixes could be leveraged to add additional control and nuance to prior work that explored empowering CI users to control aspects of their music listening experience (Hwa et al., 2021). A larger evaluation study that employed statistical methods would be needed to better understand the utility of the mixing strategies identified in this study. To facilitate the inclusion of non-experts

into these future studies, the identified mixing techniques could be grouped into individual settings and given non-technical labels, such as "clarity," "fullness," or "punch," as illustrated in Figure 7.

The breadth of digital tools available that could be used to personalize, augment, and translate musical experiences for DHH users shows great potential (Pons et al., 2016; Hwa et al., 2021; Althoff et al., 2024). Future work may also explore the design space of multi-track music distribution to remove the need for potentially lossy stem separation that may benefit listeners for who source separation artifacts impact their listening enjoyment (Pons et al., 2016). With remixing cultural and generative artificial intelligence tools heavily problematizing the notion of a recorded piece of music as an untouchable urtext, new distribution formats that empower creative listening and remixing practices could be created. As new formats, tools, and interfaces for music listening are created, there is an opportunity to begin addressing accessibility concerns earlier in the design phase by engaging with D/deaf and Disabled communities from the beginning.

5.1.2 Highlighting user agency

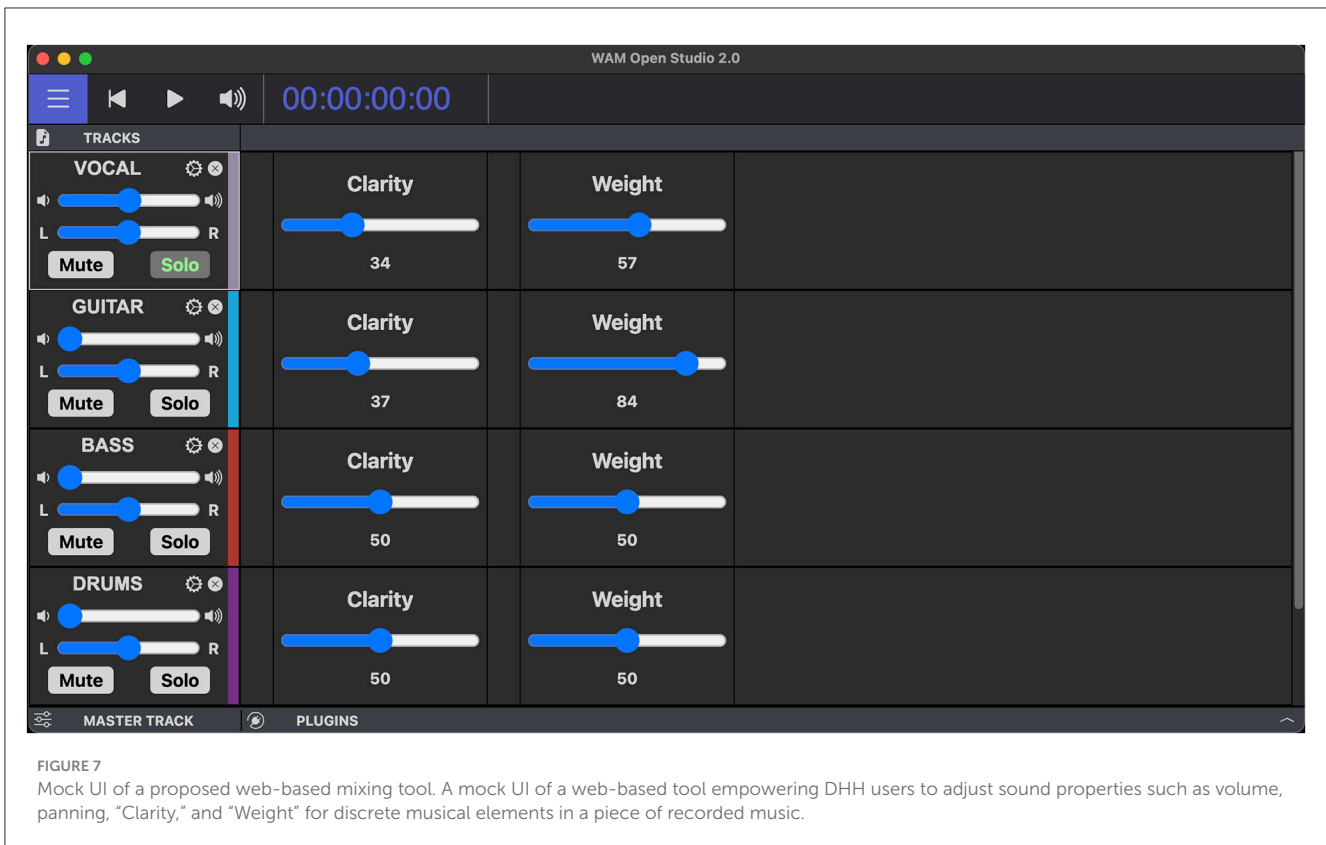
The complex factors impacting the music enjoyment of CI users, as well as a strong desire for agency and control over music listening experiences expressed by some CI users, suggest that customizable, interactive, or personalized mixes may form a helpful addition to purely algorithmic approaches to address common frustrations around music enjoyment for CI users. Employing methods that encapsulate, infer, or generalize listening behaviors, preferences, and strategies (e.g., machine learning and deep learning techniques) can be particularly helpful in situations where users desire a more streamlined experience that does not require using an app or other tool. However, if the intention behind these methods is incorrectly communicated, these tools run the risk of creating and re-enforcing norms around music listening and appreciation among CI users, possibly suggesting that there is a "correct" music mix for CI users and that if a CI user does not enjoy that mix, they are at fault.

Additionally, while well-defined research targets like speech-in-noise scores and general audio quality measures may help address specific questions, using a more practical and contextual metric, such as *enjoyment* or *desire to listen again* may provide invaluable guidance and encourage more holistic design and implementation strategies. Tools that carefully consider the diversity of hearing abilities, listening strategies, aesthetic preferences, and situational and cultural contexts, combined with algorithmic approaches, hold exciting potential to improve music enjoyment for CI users.

5.1.3 Improving feedback systems

The results of this exploratory study also highlight the need for improved feedback mechanisms for both AEs and ELs. Several AEs' expressed a general deference to the ELs' opinion and showed a strong desire to implement all of the feedback given by the ELs. While some AEs appreciated less technical feedback that provided room for them to devise creative solutions in response to the ELs' feedback, a tension was present between the perceived vagueness of some feedback and the level of interpretation required when implementing this non-technical feedback. Therefore, ELs' could

¹² In all exit interviews, AEs were extensively briefed that the CI simulation is a very coarse approximation of the experience of a CI user and that the CI user's lived experiences and acquired expertise of listening through the device greatly impact their music listening experience.



be empowered to provide more robust feedback that more clearly communicates what they desire, supporting the development of streamlined approaches to develop default initial mixes and facilitate the customization of subsequent mixes. Examples of this may include a library of example mixes ELs could reference or a real-time collaborative mixing interface where an EL could respond to the AE’s mixing decisions in real-time and additionally allow ELs to *show* AEs what they want rather than simply *telling* them. Given both the diversity of preferences among CI users as well as the lack of training around auditory diversity for most AEs, uncertainty around “when to stop” and “how much is too much” when implementing EL feedback could be alleviated through a more robust feedback system.

5.2 Limitations

The study contained several limitations including the relatively small sample size and biased recruitment demographics. While this initial study indicated a breadth of music remixing and personalization preferences among CI users, with only 10 ELs, further research is needed to explore and characterize this diversity of preferences, and statistically evaluate the mixing techniques identified in this study. Additionally, the recruited AE cohort represented a lack of gender diversity. Results may differ with a more gender-diverse AE cohort and future work could use additional AE recruitment strategies, such as direct recruitment from AE training programs, to

facilitate the recruitment of a more gender-diverse participant pool.

Seven of the ten ELS had played musical instruments for more than five years, and all were primarily familiar with Western tonal music, therefore the trends seen here may not hold for populations with different musical traditions and backgrounds, and future work utilizing statistical methods would be needed to confirm and further explore the findings of these studies, as well as their translatability across cultural and musical traditions. It was noted that the five-second audio excerpts proved challenging for some ELs, as the additional musical context of longer pieces of music is often helpful. Due to the time constraints present in the study, several genres were not included, such as jazz, rock, and heavy metal and should be explored in future work. However, even with the short clip lengths and limited number of genres, some AEs noted that the study’s workload was challenging to complete. Additionally, some feedback provided by ELs addressed aspects of the music—such as song selection and compositional factors—that are out of scope for this study.

6 Conclusion

We have presented a mixed-methods study that leveraged insights from both AE professionals and EL CI users to investigate possible technical techniques to enhance CI user music enjoyment and promote empathetic listening practices in AEs. Two mixing strategies, namely *essentializing* a mix

by removing non-crucial musical elements, and *exaggerating* the remaining musical elements, were found to be common among mixes preferred by ELs. Common themes found throughout the interviews include systemic ableism/audism in the music industry that resulted in CI users feeling excluded from the music industry's target audience, AEs not receiving adequate training around aural diversity consideration, and AEs experiencing anxiety around their own aural diversity. Finally, this work corroborates previous findings of a strong desire among some CI users for increased agency over their music listening experience.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Stanford University Institutional Review Board (IRB). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LM: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. AH: Data curation, Project administration, Writing – original draft, Writing – review & editing. SP: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. BK: Funding acquisition, Resources, Supervision, Visualization, Writing – review & editing. JB: Funding acquisition, Supervision, Writing – review & editing.

References

- Althoff, J., Gajecki, T., and Nogueira, W. (2024). Remixing preferences for western instrumental classical music of bilateral cochlear implant users. *Trends Hear.* 28:23312165241245219. doi: 10.1177/23312165241245219
- Au, A., Marozeau, J., Innes-Brown, H. (2012). "Music for the cochlear implant: audience response to six commissioned compositions," in *Seminars in Hearing* (Thieme Medical Publishers), 335–345. doi: 10.1055/s-0032-1329223
- Bennett, C. L., and Rosner, D. K. (2019). "The promise of empathy: design, disability, and knowing the "other"," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. doi: 10.1145/3290605.3300528
- Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious. Cogn.* 14, 698–718. doi: 10.1016/j.concog.2005.06.004
- Braun, V., and Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualit. Res. Sport, Exer. Health* 11, 589–597. doi: 10.1080/2159676X.2019.1628806
- Brereton, J., Daffern, H., Young, K., and Lovedee-Turner, M. (2020). "Addressing gender equality in music production," in *Gender in Music Production*, 219. doi: 10.4324/9780429464515-18
- Burgstahler, S., Doe, T. (2004). Disability-related simulations: If, when, and how to use them in professional development. *Rev. Disab. Stud.* 1:385. Available at: <https://rdsjournal.org/index.php/journal/article/view/385>
- Buyens, W., Moonen, M., Wouters, J., and van Dijk, B. (2017). "A model for music complexity applied to music preprocessing for cochlear implants," in *2017 25th European Signal Processing Conference (EUSIPCO)* (IEEE), 971–975. doi: 10.23919/EUSIPCO.2017.8081352
- Buyens, W., Van Dijk, B., Moonen, M., and Wouters, J. (2014). Music mixing preferences of cochlear implant recipients: a pilot study. *Int. J. Audiol.* 53, 294–301. doi: 10.3109/14992027.2013.873955
- Caldwell, M. T., Jiam, N. T., and Limb, C. J. (2017). Assessment and improvement of sound quality in cochlear implant users. *Laryngos. Invest. Otolaryngol.* 2, 119–124. doi: 10.1002/lio2.71
- Chen, F., and Lau, A. H. (2014). "Effect of vocoder type to mandarin speech recognition in cochlear implant simulation," in *The 9th International Symposium on Chinese Spoken Language Processing* (IEEE), 551–554. doi: 10.1109/ISCSLP.2014.6936705

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was done with the support of a grant from the GRAMMY Museum® Grant Program.

Acknowledgments

We are grateful for the administrative support of Nette Worthey. We are particularly grateful to the participants for their time, patience, and thoughtful input.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomp.2024.1371728/full#supplementary-material>

SUPPLEMENTARY TABLE 1

List of stimuli used in the mixing portion of the study.

- Cuff, B. M., Brown, S. J., Taylor, L., and Howat, D. J. (2016). Empathy: a review of the concept. *Emot. Rev.* 8, 144–153. doi: 10.1177/1754073914558466
- Embol, L., Hutterers, C., Junker, A., Reipur, D., Adjorlu, A., Nordahl, R., et al. (2021). Hearmevirtual reality: using virtual reality to facilitate empathy between hearing impaired children and their parents. *Front. Virtual Real.* 2:691984. doi: 10.3389/frvir.2021.691984
- Flower, A., Burns, M. K., and Bottsford-Miller, N. A. (2007). Meta-analysis of disability simulation research. *Remed. Special Educ.* 28, 72–79. doi: 10.1177/0741932507028002060
- French, S. (1992). Simulation exercises in disability awareness training: a critique. *Disab. Handicap Soc.* 7, 257–266. doi: 10.1080/02674649266780261
- Gauer, J., Nagathil, A., and Martin, R. (2018). “Binaural spectral complexity reduction of music signals for cochlear implant listeners,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE), 251–255. doi: 10.1109/ICASSP.2018.8461675
- Gfeller, K., Christ, A., Knutson, J. F., Witt, S., Murray, K. T., and Tyler, R. S. (2000). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *J. Am. Acad. Audiol.* 11, 390–406. doi: 10.1055/s-0042-1748126
- Gfeller, K., Mallalieu, R. M., Mansouri, A., McCormick, G., O’Connell, R. B., Spinowitz, J., et al. (2019). Practices and attitudes that enhance music engagement of adult cochlear implant users. *Front. Neurosci.* 13:1368. doi: 10.3389/fnins.2019.01368
- Gibson, D. (2019). *The Art of Mixing: a Visual Guide to Recording, Engineering, and Production*. London: Routledge. doi: 10.4324/9781351252225
- Goldsworthy, R. L. (2019). Temporal envelope cues and simulations of cochlear implant signal processing. *Speech Commun.* 109, 24–33. doi: 10.1016/j.specom.2019.03.003
- Hartmann, R., and Kral, A. (2004). “Central responses to electrical stimulation,” in *Cochlear Implants: Auditory Protheses and Electric Hearing* (Springer), 213–285. doi: 10.1007/978-0-387-22585-2_6
- Hein, G., and Singer, T. (2008). I feel how you feel but not always: the empathic brain and its modulation. *Curr. Opin. Neurobiol.* 18, 153–158. doi: 10.1016/j.conb.2008.07.012
- Hennequin, R., Khlif, A., Voituret, F., and Moussallam, M. (2020). Spleeter: a fast and efficient music source separation tool with pre-trained models. *J. Open Source Softw.* 5:2154. doi: 10.21105/joss.02154
- Hepworth-Sawyer, R., and Golding, C. (2011). *What is Music Production? A Producer’s Guide: the Role, the People, the Process*. New York: Taylor & Francis. doi: 10.1016/B978-0-240-81126-0.00023-8
- Hwa, T. P., Tian, L. L., Caruana, F., Chun, M., Mancuso, D., Cellum, I. P., et al. (2021). Novel web-based music re-engineering software for enhancement of music enjoyment among cochlear implantees. *Otol. Neurotol.* 42, 1347–1354. doi: 10.1097/MAO.0000000000003262
- Kohlberg, G. D., Mancuso, D. M., Chari, D. A., and Lalwani, A. K. (2015). Music engineering as a novel strategy for enhancing music enjoyment in the cochlear implant recipient. *Behav. Neurol.* 2015, 1–7. doi: 10.1155/2015/829680
- Koupric, M., and Visser, F. S. (2009). A framework for empathy in design: stepping into and out of the user’s life. *J. Eng. Des.* 20, 437–448. doi: 10.1080/09544820902875033
- Krüger, B., Büchner, A., and Nogueira, W. (2022). Phantom stimulation for cochlear implant users with residual low-frequency hearing. *Ear Hear.* 43, 631–645. doi: 10.1097/AUD.0000000000001121
- Limb, C. J., and Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear. Res.* 308, 13–26. doi: 10.1016/j.heares.2013.04.009
- Maarefvand, M., Marozeau, J., and Blamey, P. J. (2013). A cochlear implant user with exceptional musical hearing ability. *Int. J. Audiol.* 52, 424–432. doi: 10.3109/14992027.2012.762606
- McDonagh-Philp, D., and Denton, H. (1999). Using focus groups to support the designer in the evaluation of existing products: a case study. *Des. J.* 2, 20–31. doi: 10.2752/146069299790303570
- McGranahan, L. (2010). *Mashnography: creativity, consumption, and copyright in the mashup community*. PhD thesis, Brown University.
- Mehta, A. H., and Oxenham, A. J. (2017). Vocoder simulations explain complex pitch perception limitations experienced by cochlear implant users. *J. Assoc. Res. Otolaryngol.* 18, 789–802. doi: 10.1007/s10162-017-0632-x
- Mirza, S., Douglas, S., Lindsey, P., Hildreth, T., and Hawthorne, M. (2003). Appreciation of music in adult patients with cochlear implants: a patient questionnaire. *Cochl. Implants Int.* 4, 85–95. doi: 10.1179/cim.2003.4.2.85
- Nagathil, A., Weihs, C., and Martin, R. (2015). Spectral complexity reduction of music signals for mitigating effects of cochlear hearing loss. *IEEE/ACM Trans. Audio, Speech Lang. Proc.* 24, 445–458. doi: 10.1109/TASLP.2015.2511623
- Nemer, J. S., Kohlberg, G. D., Mancuso, D. M., Griffin, B. M., Certo, M. V., Chen, S. Y., et al. (2017). Reduction of the harmonic series influences musical enjoyment with cochlear implants. *Otol. Neurotol.* 38:31. doi: 10.1097/MAO.0000000000001250
- NIDCD (2021). *NIDCD Fact Sheet: Cochlear Implants*. National Institute of Deafness and Other Communication Disorders.
- Nogueira, W. (2019). “Evaluation of new music compositions in live concerts by cochlear implant users and normal hearing listeners,” in *14th International Symposium on Computer Music Multidisciplinary Research*, 294.
- Nogueira, W., Litvak, L., Edler, B., Ostermann, J., and Büchner, A. (2009). Signal processing strategies for cochlear implants using current steering. *EURASIP J. Adv. Signal Process.* 2009, 1–20. doi: 10.1155/2009/534059
- Nogueira, W., Nagathil, A., and Martin, R. (2018). Making music more accessible for cochlear implant listeners: recent developments. *IEEE Signal Process. Mag.* 36, 115–127. doi: 10.1109/MSP.2018.2874059
- Pons, J., Janer, J., Rode, T., and Nogueira, W. (2016). Remixing music using source separation algorithms to improve the musical experience of cochlear implant users. *J. Acoust. Soc. Am.* 140, 4338–4349. doi: 10.1121/1.4971424
- Riley, P. E., Ruhl, D. S., Camacho, M., and Tolisano, A. M. (2018). Music appreciation after cochlear implantation in adult patients: a systematic review. *Otolaryngol. Head Neck Surg.* 158, 1002–1010. doi: 10.1177/0194599818760559
- Riss, D., Hamzavi, J.-S., Blineder, M., Flak, S., Baumgartner, W.-D., Kaider, A., et al. (2016). Effects of stimulation rate with the fs4 and hdcis coding strategies in cochlear implant recipients. *Otol. Neurotol.* 37, 882–888. doi: 10.1097/MAO.0000000000001107
- Rogers, C. R. (1975). Empathic: an unappreciated way of being. *Couns. Psychol.* 5, 2–10. doi: 10.1177/001100007500500202
- Roy, A. T., Jiradejvong, P., Carver, C., and Limb, C. J. (2012). Assessment of sound quality perception in cochlear implant users during music listening. *Otol. Neurotol.* 33, 319–327. doi: 10.1097/MAO.0b013e31824296a9
- Roy, A. T., Vigeant, M., Munjal, T., Carver, C., Jiradejvong, P., and Limb, C. J. (2015). Reverberation negatively impacts musical sound quality for cochlear implant users. *Cochlear Implants Int.* 16, S105–S113. doi: 10.1179/1467010015Z.00000000000262
- School, D. (2018). *An introduction to design thinking - process guide*. Stanford: Stanford University
- Series, B. (2014). *Method for the subjective assessment of intermediate quality level of audio systems*. International Telecommunication Union Radiocommunication Assembly.
- Sorrentino, F., Gheller, F., Favaretto, N., Franz, L., Stocco, E., Brotto, D., et al. (2020). Music perception in adult patients with cochlear implant. *Hearing, Bal. Commun.* 18, 3–7. doi: 10.1080/21695717.2020.1719787
- Spangmose, S., Hjortkjær, J., and Marozeau, J. (2019). Perception of musical tension in cochlear implant listeners. *Front. Neurosci.* 13:987. doi: 10.3389/fnins.2019.00987
- Sparrow, R. (2010). Implants and ethnocide: learning from the cochlear implant controversy. *Disab. Soc.* 25, 455–466. doi: 10.1080/09687591003755849
- Wheeler, A., Archbold, S., Gregory, S., and Skipp, A. (2007). Cochlear implants: the young people’s perspective. *J. Deaf Stud. Deaf Educ.* 12, 303–316. doi: 10.1093/deafed/enm018
- Whitmal, N. A., Poissant, S. F., Freyman, R. L., and Helfer, K. S. (2007). Speech intelligibility in cochlear implant simulations: effects of carrier type, interfering noise, and subject experience. *J. Acoust. Soc. Am.* 122, 2376–2388. doi: 10.1121/1.2773993
- Wright, R., and Uchanski, R. M. (2012). Music perception and appraisal: cochlear implant users and simulated cochlear implant listening. *J. Am. Acad. Audiol.* 23, 350–365. doi: 10.3766/jaaa.23.5.6
- Zeng, F.-G., Tang, Q., and Lu, T. (2014). Abnormal pitch perception produced by cochlear implant stimulation. *PLoS ONE* 9:e88662. doi: 10.1371/journal.pone.0088662