The Final Published Article of this manuscript has been published and is available in EAR and HEARING, Jan-Feb 2024.

Mealings K, Valderrama JT, Mejia J, Yeend I, Beach EF, Edwards B. Hearing aids reduce self-perceived difficulties in noise for listeners with normal audiograms. Ear and Hearing (**2024**) 45, 151-163. doi: 10.1097/AUD.0000000000001412.

https://doi.org/10.1097/AUD.0000000000001412

Hearing aids reduce self-perceived difficulties in noise for listeners with normal audiograms

Abstract

Objective: This study assessed hearing aid benefits for people with a normal audiogram but hearing-in-noise problems in everyday listening situations.

Design: Exploratory double-blinded case-control study whereby participants completed retrospective questionnaires, ecological momentary assessments, speech-in-noise testing, and mental effort testing with and without hearing aids. Twenty-seven adults reporting speech-in-noise problems but normal air-conduction pure-tone audiometry took part in the study. They were randomly separated into an *experimental* group that trialled mild-gain hearing aids with advanced directional processing and a *control* group fitted with hearing aids with no gain or directionality.

Results: Self-reports showed mild-gain hearing aids reduce hearing-in-noise difficulties and provide a better hearing experience (i.e., improved understanding, participation, and mood). Despite the self-reported benefits, the laboratory tests did not reveal a benefit from the mild-gain hearing aids, with no group differences on speech-in-noise tests or mental effort measures. Further, participants found the elevated cost of hearing aids to be a barrier for their adoption.

Conclusions: Hearing aids benefit the listening experience in some listening situations for people with normal audiogram who report hearing difficulties in noise. Decreasing the price of hearing aids may lead to greater accessibility to those seeking remediation for their communication needs.

Keywords

Speech-in-noise hearing difficulties; Hidden hearing loss (HHL); hearing aids; self-report; Reaction time; Ecologically momentary assessment (EMA).

1. Introduction

Although pure-tone audiometry is the most widely accepted test to assess hearing performance, it is generally accepted that the audiogram does not reliably reflect a person's real-world listening abilities. It is estimated that 10–15% of adults with a normal audiogram have greater than expected difficulty hearing speech in noise (Kumar, Amen, & Roy, 2007; Spankovich, Gonzalez, Su, & Bishop, 2018; Tremblay et al., 2015). When a client's hearing difficulties cannot be explained by the audiogram, clinicians are often unsure how to best manage these difficulties (Zhao & Stephens, 2007).

Mealings et al. (2020) conducted a design thinking study to discover the unmet needs of people with difficulties understanding speech in noise and a normal or near-normal audiogram. Two of the needs that emerged from this study were "to evaluate treatment options to provide an evidence base of what interventions may help this population" and "to gather evidence about the effectiveness of different treatment options to increase clinician confidence in addressing clients presenting concern". Less than a quarter of the clients surveyed in the study recalled being offered treatment options by their clinician. For those who were offered a treatment option, it was most often a hearing aid. Therefore, there is a need to understand if hearing aids help the real-world listening experience of this population, as well as their laboratory speech-in-noise and listening effort performance.

To date there have been only a handful of studies into the effectiveness of hearing aids in this population. Roup, Post and Lewis (2018) investigated the use of mild-gain amplification over four weeks in adults with a normal audiogram (thresholds ≤ 25 dB HL 250–8000 Hz) but difficulty hearing in complex listening situations and a control group with no reported hearing difficulties. The main rationale for using mild-gain hearing aids was to provide 5–10 dB of gain in the mid-to-high frequencies to enhance soft consonants (Roup et al., 2018). Participants completed subjective questionnaires and the Revised Speech Perception in Noise Test pre- and post-intervention. Roup et al. (2018) found significantly improved speech perception in noise and a reduction in self-reported hearing difficulties when the participants were using the device compared to pre-device. However, only three of the 17 participants purchased the hearing aid after the trial. The authors note that a limitation of the study is that it did not include a placebo control group.

Singh and Doherty (2020) investigated the effect of mild-gain hearing aids on middle-aged adults' hearing handicap, motivation, and attitudes toward hearing and hearing aids. There were two participants groups: those with and without difficulty hearing in background noise. The participants in both groups trialled open-fit hearing aids for two weeks. The aids were fitted bilaterally and amplified at 2–4kHz by approximately 5 dB gain with adaptive directionality and noise suppression. Participants with difficulty hearing in noise had higher levels of hearing handicap, were more motivated to address their hearing problems, and had higher personal distress and inadequacy scores than those without difficulty hearing in noise, but these scores did improve for those with difficulty hearing in noise after wearing the hearing aids for two weeks. However, after the trial only two of the 10 participants with difficulty hearing in background noise would consider purchasing a hearing aid. Again, the authors note that a limitation of the study was that it did not include a placebo control group.

Interestingly, both of these studies by Roup et al. (2018) and Singh and Doherty (2020) asked the participants to wear the hearing aids all day, rather than just in the environments that they had difficulty in (i.e., where there was background noise). It may be that many of the situations the participants were in during the trial were not noisy situations, therefore they may have realised that it is only a few situations that they benefitted from having the hearing aids (Singh & Doherty, 2020). This may have contributed to why a low number of participants in both studies would purchase the hearing aids.

In addition to not including a placebo control group, these two studies have other limitations. Retrospective questionnaires were the main outcome measure in both studies (though Roup et al. (2018) did also use a speech-in-noise test – the limitations with it will be discussed next).

While retrospective questionnaires can provide useful insights, because they are filled out after the event, they rely on the participant's memory of the experience being investigated so may be subject to memory/recall biases (Althubaiti, 2016; Bradburn, Rips, & Shevell, 1987). Technological advances and the ubiquity of smartphones have now provided a solution to this problem – ecological momentary assessment (EMA). EMA data are collected in real-world environments, focus on participants' current state, are conducted quickly in a strategically selected moment, and gather data over multiple time periods allowing the researcher to measure how the participant's experiences and behaviour change across time and situations (Shiffman, Stone, & Hufford, 2008). EMA has been used in the past by various health science researchers, and is starting to be used more in audiology research (e.g., Galvez et al., 2012; Hasan, Chipara, Wu, & Aksan, 2014; Henry et al., 2012; Holube, von Gablenz, & Bitzer, 2020; Timmer, Hickson, & Launer, 2017; Wu, Stangl, Zhang, & Bentler, 2015).

Measuring a listener's ability to perceive speech via a speech-in-noise test is one good way to quantify their listening difficulties (and this method was used by Roup et al. (2018)). However, many of the commonly used tests are not realistic replications of everyday listening situations as they use carefully constructed sentences with clearly pronounced speech and processed noise (e.g. Bench, Kowal, & Bamford, 1979; Hagerman, 1982; Nilsson, Soli, & Sullivan, 1994). Recently, Miles et al. (2020) created the Everyday Conversational Sentences in Noise (ECO-SiN) test which uses stimuli drawn from natural conversations presented in real-world recorded background noise. Using this test is likely to give a more accurate picture of a person's listening abilities. While speech-in-noise testing has traditionally been used to quantify a person's speech-in-noise difficulties in terms of the speech-to-noise ratio necessary for accurate perception, recently there has also been an emphasis on how much listening effort is needed to do so.

"Effort" is defined by Pichora-Fuller et al. (2016) as "the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task". When listening to a degraded acoustic signal, such as speech masked by other speech or noise, listeners have to rely more on cognitive systems to extract meaning from the speech compared to if they were listening to a clear signal (Peelle, 2018). In these conditions, listeners take longer to respond, have greater difficulty remembering the speech, and make more errors when processing syntactically complex sentences (see Peelle (2018) for a review). Listening in difficult acoustic conditions is not only an auditory challenge but is also likely to affect verbal working memory and attention-based performance monitoring (Peelle, 2018).

The aim of the current study was to assess the real-world benefit of hearing aids as a potential intervention to help people who have a normal audiogram but difficulties hearing speech in noise hear and process speech better in complex listening environments. This study extends previous research by including a control group with an acoustically-transparent (or placebo) hearing aid. It also uses a more realistic speech-in-noise test and assesses listening effort as well as speech-in-noise performance. Additionally, it uses EMA as well as traditional questionnaires to give real-time, real-world data. The real-world component only required the participants to wear their hearing aids in situations that they would have difficulty in (i.e., where there was background noise), and the EMA app recorded the noise level to include in the analysis.

It was hypothesised that the group wearing mild-gain hearing aids would have an improved hearing experience in the real-world and better speech-in-noise performance with reduced listening effort in the laboratory.

2. Methods

2.1. Ethics and data sharing

The study was conducted at the National Acoustic Laboratories (NAL, Sydney, Australia) following protocols in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. These protocols were approved by the Hearing Australia Human Research Ethics Committee (EC00109, Ref. AHHREC 2019-15).

Consistent with the Findable, Accessible, Interoperable, and Reusable (FAIR) Data Principles (Wilkinson et al., 2016), the raw data is available as supplementary material.

2.2. Participants

Participants were recruited from the NAL Research Participant Database (a register of people who have given their consent to be invited to participate in NAL research); clients from Hearing Australia (an Australian government-funded hearing service provider); hearing clinics near NAL; staff and students from NAL and Macquarie University; and through advertisements on social media and the Macquarie University campus.

Participants had to meet four inclusion criteria: (i) aged 18–70 years, (ii) proficient in English, (iii) reporting speech-in-noise intelligibility problems, and (iv) with air-conduction pure-tone audiometry within the normal range. Normal range audiograms were defined as a four-frequency average hearing loss (i.e., the mean hearing threshold at 0.5, 1, 2 and 4 kHz) lower or equal to 25 dB hearing level in both ears. This level is defined as normal hearing by Spankovich et al. (2018) and defined as normal hearing for speech by the National Institute on Deafness and Other Communication Disorders, from the U.S. Department of Health and Human Services (Gates & Hoffman, 2011).

Twenty-seven adults (17 females, aged 19–68 years, mean ± std = 42.7 ± 11.9 years) out of 40 participants recruited met the inclusion criteria, volunteered to participate, completed the study, and were reimbursed for their time at the end of the study via Prezzee Pty Ltd (Sydney, Australia) – a trusted provider of digital gift cards.

Air-conduction pure-tone audiometry was conducted using the AC40 clinical audiometer (Interacoustic A/S, Middelfart, Denmark) in a soundproof booth. Figure 1 presents the quartile distributions of the participants' pure-tone hearing thresholds in both ears. The individual audiometric results are in supplemental Appendix A. The mean ± std [min–max] of the four-frequency average hearing loss was 8.3 ± 4.8 [0–16.3] dB for the left ear and 9.8 ± 5.0 [1.3–21.3] dB for the right ear. Figure 1 also shows large individual variability in extended high frequencies, i.e., from 9 to 12.5 kHz.

All participants responded '*Yes*' to the question '*Do you have difficulty hearing speech in noisy environments?*'. The most common venues in which participants reported hearing difficulties included restaurants, bars, and cafés; reverberant venues with hard surfaces such as shopping centres and lecture rooms; and places with multiple cross-conversations like a noisy office or a conference poster session.

2.3. Study design

Figure 2 presents a diagram of the study design, which involved two phases and three lab visits. Phase I aimed to characterise the participants' speech-in-noise intelligibility difficulties when they were unaided; and phase II evaluated the value of hearing aids in improving the hearing experience in noisy situations.

2.3.1. Phase I

The participants' hearing-in-noise difficulties were characterised via (i) a standardised questionnaire and (ii) real-world surveys using EMA.

Hearing-difficulties questionnaire. In the first appointment, participants' self-perceived hearing-in-noise difficulties were measured via the *speech* subscale of the Speech, Spatial and Qualities of Hearing Scale (SSQ; Noble et al., 2013), i.e., the first five items of the SSQ questionnaire that assess speech understanding in noisy environments, in situations with multiple talkers, and in scenarios that require participants to constantly switch their attention.

NEMA-Phase I. During the six weeks following the first appointment, participants were encouraged to attend different noisy venues where they usually experience hearing difficulties. Each participant had the NAL-EMA (NEMA) app installed on their mobile phone, and at each venue, they were asked to complete a brief survey rating different dimensions of their hearing experience in the environment. Surveys asked about (i) the type of venue (e.g., bar/pub, café/restaurant, party or gathering at home, shopping centre, etc.), (ii) perceived level of background noise, (iii) mental effort required to follow the conversation, (iv) level of understanding, (v) level of participation, (vi) level of frustration, and (vii) overall mood. The NEMA app recorded acoustic features of the environment while participants were completing the surveys, including the A-weighted background noise level.

2.3.2. Phase II

Participants categorization. This phase aimed to evaluate the value of hearing aids in ameliorating hearing-in-noise difficulties. To achieve this aim, participants were randomly assigned either to the *control* group–in which participants were binaurally fitted with an acoustically transparent device (i.e., a hearing aid that did not apply any gain), or the *experimental* group–in which the hearing aids provided a mild gain, which enabled access to advanced directionality and noise reduction features. Thus, the hypothesis was that participants fitted with mild-gain hearing aids would experience a greater acoustic benefit in noisy environments than those with acoustically transparent devices. Both participants and researchers were blind to participants' categorisation to avoid any possible bias (Misra, 2012). In total there were 14 participants in the control group (9 females, 19–63 years, mean ± std = 40.8 ± 12.6 years) and 13 participants in the experimental group (8 females, 31–68 years, mean ± std = 44.8 ± 11.1 years).

Devices fitting. In the second NAL appointment, all participants were fitted with Phonak Marvel Audéo M50-312 hearing aids (Sonova Australia Pty Ltd, Baulkham Hills, Australia) using open domes. For the control group, the real ear insertion gain applied to the hearings aids was approximately equal to 0 dB, and the '*Speech in noise*' program was configured by disabling NoiseBlock (i.e., a noise reduction algorithm) and setting microphones in omni-directional mode. For the experimental group, a real ear insertion gain of 6 dB was targeted to provide a good balance between amplification and loudness perception, such that listeners have access to the hearing aid advanced processing while having minimum objections to the level of sounds in everyday listening conditions. Other hearing aid parameters included '*Speech in noise*' program which was set with NoiseBlock on *weak* (level 8) and microphones were set to *UltraZoom & SNR-Boost* (level 20)–an adaptive monaural beamformer aimed at improving speech understanding in situations with background noise (Ricketts & Henry, 2002; Wouters, vanden Berghe, & Maj, 2002). Figure 3 shows that the gain provided across the 1 to 8 kHz frequency range was, on average across participants, around 0 dB for the control group and 6 dB for the experimental group. In both groups, the fitting formula was NAL-NL2 (Keidser, Dillon, Flax, Ching, & Brewer, 2011) and linear compression was applied. The different hearing aid programs were automatically selected via *AutoSense OS* in order to ensure the hearing aids would self-adjust in different acoustic scenarios, switching to the '*Speech in noise*' program in environments above 67 dB SPL (Watson, 2015). Figure 3 presents the averaged real ear insertion gain (REIG) across the two groups for the International Speech Test Signal (Holube, Fredelake, Vlaming, & Kollmeier, 2010; i.e., a signal that combines speech segments in six different languages) presented from a front speaker situated 1 meter from the participant at 65 dB SPL. Real ear measures were recorded using the Aurical Freefit (Natus Medical Inc., Middleton, WI), with probe tubes inserted in the participants' left and right ear canal.

NEMA-Phase II. Following the second NAL appointment, participants were again asked to trial the hearing aids in various acoustically challenging places where they usually struggle communicating with others, and evaluate their hearing experience via NEMA. In addition to questions about the type of venue, level of noise, number of people in the conversation, mental demand, level of understanding, frustration and mood, the NEMA surveys in this second phase also included questions about the appropriateness of the device's sound quality, perceived acoustic benefit, and overall satisfaction. The A-weighted background noise level was again measured during these surveys.

Hearing difficulties questionnaires. In the third and final appointment, participants evaluated their hearing experience with hearing aids via three questionnaires: (i) the *speech* subscale of the SSQ (Noble et al., 2013); (ii) the Satisfaction with Daily Amplification (SADL; Cox and Alexander (1999) – a standardized questionnaire that assesses positive effects, negative features, service/cost, and personal image; and (iii) an end-of-study questionnaire that asked open-ended questions about the positives and negatives of hearing aids, whether the participant would use hearing aids in similar situations in the future, and whether they would purchase a pair of hearing aids considering the cost is AUD 5000 (approximately USD 3250).

Laboratory tests. In addition, participants in this appointment completed two laboratory tests. The tests evaluated how well the hearing aids improved hearing-in-noise performance and reduced the mental effort required to understand speech in noise. These tests are presented below.

ECO-SIN. Speech-in-noise hearing performance was measured via the ECO-SiN (Miles et al., 2020) in aided and unaided conditions. The main advantage of this test is that it uses both naturally produced speech extracted from real conversations and realistic background noise, which provides the opportunity for capturing the hearing aid benefit in a real-world environment. This test was administered in the anechoic chamber of the Australian Hearing Hub (Sydney, Australia). The background noise was actual *dinner restaurant* noise obtained from the Ambisonics Recordings of Typical Environments (ARTE) database (Weisser et al., 2019), presented at 73 dB SPL from an array of 41 speakers spherically distributed in five rows. The target sentences were presented from a speaker situated in front of the participant–level started at 78 dB SPL (i.e., at a signal to noise ratio (SNR) of +5 dB) and varied according to the staircase method until the 50% speech reception threshold (SRT-50, i.e., the SNR corresponding to 50% intelligibility) was estimated. Speech intelligibility was measured at morpheme level–the smallest unit of meaning within a word. The order of the two test conditions (*aided* and *unaided*) was randomised across participants to compensate for any possible learning effect.

Dual task. The mental effort required to understand speech in noise was measured via a dual-task paradigm developed at NAL (Valderrama et al., 2022). The primary task consisted of repeating a sentence presented in background noise. Background noise was actual *cafeteria* noise obtained from the ARTE database (Weisser et al., 2019) and presented from an array of 41 speakers at 65 dB SPL. Target speech was the Australian version of the Matrix test (Kelly et al., 2017), in which a closed set of words from the categories *Name* + *Verb* + *Number* + *Adjective* + *Object* are combined randomly to form sentences (e.g. *Rachel wins three dark chairs*). Target speech was delivered from a speaker situated in front of the participant, and level was adjusted for each participant, corresponding to their SRT-50. These SRTs were obtained from a psychometric function fitted to intelligibility scores obtained in different SNRs – which varied from −16 dB to +12 dB in steps of 4 dB. Intelligibility was measured in terms of percentage of words correct. The secondary task was a visual task driven by the auditory stimulus of the primary task. Two large vertical rectangles situated in the horizontal plane were presented on an acoustically transparent screen situated in front of the participants. At the onset of each sentence, a circle appeared in the middle of one of the rectangles, and participants were instructed to use a keyboard and press one of two arrows depending on the subject of the sentence. If the subject (i.e., the first word) was a male name, the participant pressed the arrow pointing towards the circle. If it was a female name, the participant was instructed to press the arrow pointing away from the circle. Mental effort was measured in terms of the reaction time (Gagné, Besser, & Lemke, 2017) from sentence onset to the button press. This test was administered in trials of 10 sentences, in which participants were also asked to rate their self-perceived mental effort on a 7-point scale ranging from “No effort” to “Extreme effort” (Bernarding, Strauss, Hannemann, Seidler, & Corona-Strauss, 2017; Desjardins, 2016) after the presentation of five sentences. Two trials were presented in the *unaided* condition and two additional trials in the *aided* condition – thus, a total of four trials were presented per participant. The presentation order of these trials was randomised to distribute any possible learning effect equally between the two conditions.

2.4. Data analysis

Data analysis was conducted using custom scripts developed in MATLAB (version R2022a, The Mathworks Inc., Natick, MA), using functions from the 'Signal Processing', 'Optimization', and 'Statistics and Machine Learning' toolboxes.

2.4.1. Questionnaires

Self-reported hearing-in-noise difficulties were compared between the control and experimental groups in the unaided and aided conditions. via a series of linear mixed-effects models. The speech subscale of the SSQ scores was the predictor variable. The question number, group and test condition as independent variables, and participants were random variables. For the SADL questionnaire, a global score was obtained from each participant as described in Cox and Alexander (1999), and the scores from the control and experimental groups were statistically compared via an independent-samples two-sided *t*-test. In the end-of-study questionnaire, the open-ended questions were analyzed qualitatively to extract the main themes. The raw data from the questionnaires can be found in supplemental Appendix A.

2.4.2. Real-world assessment

The control group recorded 92 EMA entries in the two phases: the average number of recorded entries per participant was 9.2. The experimental group recorded 142 EMA entries in the two phases: the average number of recorded entries per participant was 10.5. However, concerning the specificity of location, the control group had only two responses in the Bar/Pub listening condition, whereas the experimental group had 13 responses. Therefore, this specific location was excluded from the statistical analyses, nonetheless, the recorded EMA entries were plotted in the figures to indicate the diversity of everyday listening conditions that participants attended to while evaluating the devices.

The first step in NEMA data analysis was to reduce the dimensionality of the dataset via a factor analysis using *equamax* rotation. This step was motivated by several survey items that correlated with each other. This process reduced the dimensionality of the NEMA dataset to three factors (DFE = 18, $\chi^2 = 77.97$, p-value < 0.001). Table 1 presents the Pearson's correlation coefficient between each survey item and the three factor scores. This table shows that there was a high correlation between the first factor and survey items associated with the *hearing experience* of the participants, the second factor correlated with survey items that addressed different dimensions of the *device performance*, and the third factor correlated with survey items that characterised participants' self-perceived level of *acoustic challenge* of the sound environment. A single score for each factor was obtained by reversing the score of survey items in which lower score was associated with a positive outcome (i.e., noisiness, mental effort, and frustration), and averaging the score of the survey items that highly correlated with each factor (highlighted in grey in Table 2). This approach was taken to preserve interpretability of the data, as each factor score maintained the original 5-level scale from 1 (the worst outcome) to 5 (the best outcome).

Factor scores from the control and experimental groups were compared via a series of Tukey's honestly significant difference tests (Tukey's HSD) – a non-parametric statistical test appropriate for data not normally distributed, which accounts for multiple comparisons. Factors 1 and 3 were evaluated in the unaided and aided conditions; and Factor 2 was evaluated only in the aided condition. Statistical significance in these analyses was achieved considering type I error $\alpha = 0.05$. Additionally, factor scores were analysed in terms of the location in which the survey had been conducted. The NEMA raw data is available in supplemental Appendix B.

2.4.3. Speech-in-noise hearing performance

Individual performance in the ECO-SiN test was measured in terms of SRT-50, i.e., the SNR at which participants attained 50% intelligibility. The effect of wearing hearing aids on intelligibility was assessed via a paired Student's *t*-test comparing the difference between the SRT-50 scores in the aided and unaided conditions.

2.4.4. Mental demand

In the dual-task test paradigm, the hearing aid effect was measured by evaluating the statistical significance of the difference between the unaided and aided conditions for intelligibility, reaction time, and self-reported effort measures, both for control and experimental participants. Statistical significance in these tests was assessed via paired Student's *t*-tests.

2.4.5. Available data

A portion of the recruited participants withdrew at different stages of the study due to COVID-19 restrictions, relocating to another city, and other personal reasons. Further, some participants' phones did not support the NEMA app, which prevented them from conducting this part of the study. This situation led to an imbalanced dataset across some tests (see Table 2). Available data from participants who had withdrawn from the study were generally used and analysed, with the exception of NEMA analysis, which used only data from participants who had completed the two phases.

3. Results

3.1. Questionnaires

Figure 4.A presents the mean score of the speech subscale of the SSQ for control (top scale) and experimental (bottom scale) participants, in the unaided (red) and aided (blue) conditions. This

figure shows that, at group level, the participants reported a moderate degree of hearing-in-noise difficulties, since the mean score was 5.4 and 5.1 for control and experimental participants, respectively. The difference in the scores between control and experimental participants was not statistically significant (estimate ± standard error (SE) = -0.34 ± 0.59; 95% confidence interval (CI) = [-1.51,0.84]; t(128) = -0.57; p-value = 0.57) in the unaided condition, which is consistent with the randomised allocation of participants across the two groups. Further, Figure 4.A shows a different effect of hearing aids in the two groups. On average, control participants trialling *placebo* hearing aids did not experience a substantial amelioration of their self-reported hearing-in-noise difficulties, as observed by close mean scores (i.e., 5.4 and 5.8 in the unaided and aided conditions, respectively) and this difference was not statistically significant (estimate ± SE = -0.36 ± 0.32; 95% CI = [-0.99,0.27]; t(134) = -1.12; p-value = 0.26). However, experimental participants reported statistically significant lower levels of hearing-in-noise difficulties when they were fitted with mild-gain hearing aids, i.e., their mean score shifted from 5.1 unaided to 6.4 aided (estimate ± SE = -1.33 ± 0.40; 95% CI = [-2.12,-0.54]; t(113) = -3.35; p-value = 0.001). The raw data of the SSQ questionnaire is in supplemental Appendix A.

Figure 4.B presents the mean SADL score for control (red) and experimental (blue) participants, along with normative data obtained from a large cohort of regular hearing aid users (i.e., 4.9 ± 1.0; Cox & Alexander, 1999). Results show that participants fitted with mild-gain hearing aids reported significantly higher satisfaction with the devices than those fitted with acoustically transparent hearing aids (control: 3.7 vs experimental: 4.8, 95% CI = [-1.67,-0.38]; t(23) = -3.30; p-value = 0.003). Results also showed that the experimental group's satisfaction levels were consistent with those reported by individuals with hearing loss; and satisfaction with zero-gain hearing aids fell below the norms. The raw data of the SADL questionnaire is available in supplemental Appendix A.

In the end-of-study questionnaire, participants were asked about the most positive features of the study hearing aids. Both control and experimental participants highlighted that they are discreet and comfortable to wear. Regarding the performance of the hearing aids as an assistive listening device, only four control participants (i.e., 29%) reported hearing better with them. In contrast, 10 of the 11 experimental participants (i.e., 91%) said that the hearing aids had improved their hearing ability in noisy environments. For example, one experimental participant wrote, "[I had] *increased clarity of speech from people in front of me, with a reduction in majority of background noise. When in a quiet environment, I felt like I had 'super hearing' so had to put less effort in to listening for speech.*"

When asked about the negative features, both control and experimental participants reported comfort issues and distorted hearing due to unnatural amplification of sounds. For example, one participant said, "*There are individual noises from the background that can be distracting when they are amplified. An example was a cafe situation that had an indoor fountain, and the sound of flowing water was louder than without the device*." Two participants reported feeling a certain degree of stigma when wearing hearing aids. Further, four control participants flagged the limited hearing benefit that they gained from the hearing aids. However, this concern was not mentioned by any experimental participant. For example, one control participant said, "*They don't make a change to my hearing"*, and another mentioned "*I sometimes do not know if they are working or not*." These responses from the control group were not unexpected, considering that they were fit with zero-gain hearing aids.

The end-of-study questionnaire also asked participants whether they would continue using the hearing aids in similar situations in the future. Most control participants (9/14, 64%) responded '*No*', while most experimental participants (8/11, 73%) responded '*Yes*'. For example, one control participant said, "*No, they don't really help"*, and one experimental participant said, "*The benefit that I gain from them is too great to not use them*." However, when they were asked if they would purchase the hearing aids at their cost of AUD 5000, all control and experimental participants responded that they would not buy this technology for that price. For example, one control participant said, "*No, I am not convinced they helped me"*, and one experimental participant said, "*My hearing issues are not that restrictive to justify the cost.*"

3.2. Real-world assessment

Table 3 shows the median score and the statistical evaluation of group differences in all the experimental conditions in the three factors. Results showed no statistically significant differences in Factors 1 and 3 between the control and experimental groups in the unaided condition. This was expected since participants were randomised across the two groups. Results also showed that the experimental group reported a significantly better hearing experience (25% median increase) when they were aided, relative to the unaided condition. However, this effect was not statistically significant in the control group. This analysis also revealed that the experimental group reported higher levels of device performance (25% median increase) than the control group. Further, Factor 3 group comparisons showed that when participants were aided, the control group were neutral about how challenging they found different sound environments, while the experimental group reported significantly lower scores than the control group.

Figure 5 presents a visual representation of the sound level (panel A) and factor scores (panels B to D) in different environments. Panel A of this figure shows that 'Bar / Pub' was the loudest scenario, with an averaged mean level of 85.5 dB SPL; followed by 'Café / Restaurant' and 'Shopping centre'. The 'Other' and 'Gathering at home' sound environments presented a large variability, with mean levels of 66.5 and 62.3 dB SPL, respectively. Panels B to D present the factor scores per sound environment for the control and experimental groups in the aided condition. The statistical analysis of Hearing experience (Factor 1) per sound environment showed that, relative to the control group, experimental participants reported higher levels of hearing experience in 'Other' (p-value = 0.042; 95% CI = [0.21,6.33]) and 'Shopping centre' (p-value = 0.028; 95% CI = [0.61,10.44]) sound environments. For device (Factor 2) experimental participants reported higher levels of perceived device performance in ‘Café / Restaurant’ (p-value = 0.005; 95% CI = [3.14,10.67]), ‘Other’ (p-value = 0.007; 95% CI = [2.29,14.91]), and ‘Gathering at home’ (p-value = 0.030; 95% CI = [0.54,10.80]). Listening effort (Factor 3) analysis showed that experimental participants found the ‘Café / Restaurant’ sound environment more challenging than control participants (p-value = 0.0006; 95% CI = [-21.07,-5.80]). However, Listening effort scores were better in the experimental group than in the control group in the ‘Other’ environment (p-value = 0.022; 95% CI = [1.00,13.25]). Although the absence of statistics from the control group in ‘Bar / Pub’ may have prevented a fair comparison between the two groups in the loudest scenario, panels B to D show that (i) participants’ hearing experience was poorer in that scenario compared to other venues, (ii) device performance was in the range of neutrality, and (iii) participants perceived those environments as highly challenging.

3.3. Speech-in-noise hearing performance

Figure 6 shows the SRT-50 from the ECO-SiN test for control and experimental participants, in the unaided and aided conditions. The SRT-50 mean ± standard deviation for the control group was −2.25 ± 2.55 dB unaided, and -2.09 ±1.75 dB aided; and for the experimental group was −2.14 ±2.22 dB unaided, and -1.41 ± 1.87 dB aided. Both control and experimental participants showed non-significant differences between the unaided and aided conditions (control: 95% CI = [-1.64,1.95], $t(8) = 0.20$, p-value = 0.845; experimental: 95% CI = [-0.77,2.23], $t(8) = 1.13$, p-value = 0.293).

3.4. Mental effort

The SNR at which the dual task was administered was adjusted to each participant´s SRT-50, i.e., the SNR at which unaided participants presented 50% intelligibility. The mean ± standard deviation of the SRT-50 across participants was 0.53 ± 1.89 dB.

Figure 7 presents the dual task outcomes for control and experimental participants in the unaided and aided conditions. Panel A shows that both control and experimental participants presented statistically comparable intelligibility scores when they were unaided and aided. The intelligibility mean ± standard deviation for the control group was 71.1 ± 11.5% unaided, and 68.7 ± 11.2% aided (difference: 95% CI = [-3.40,8.29]; t(8) = 0.96; p-value = 0.363); and for the experimental group, it was 73.8 ± 10.6% unaided, and 69.9 ± 9.8% aided (difference: 95% CI = [-1.76,9.54]; t(8) = 1.59; p-value = 0.151). Consistent with this result, behavioural (panel B) and self-reported (panel C) measures of the mental effort required to perform the dual task showed statistically similar outcomes between the aided and unaided conditions, both for control and experimental participants. The mean ± standard deviation of participants' reaction time for the control group was 1604.9 ± 657.9 ms unaided, and 1510.8 ± 564.8 ms aided (difference: 95% CI = [-60.24,248.28]; t(8) = 1.41; p-value = 0.197); and for the experimental group was 1301.9 ± 210.7 ms unaided, and 1314.2 ± 306.7 ms aided (difference: 95% CI = [-140.39,115.70]; t(8) = -0.22; p-value = 0.830). The mean ± standard deviation of self-reported effort for control participants was 5.4 ± 1.0 unaided, and 5.5 ± 0.9 aided (difference: 95% CI = [-0.54,0.32]; t(8) = -0.59; p-value = 0.568), and for the experimental group it was 5.6 ± 0.8 unaided, and 5.5 ± 0.9 aided (difference: 95% CI = [-0.30,0.58]; t(8) = 0.73; p-value = 0.488).

4. Discussion

The aim of the current study was to assess mild-gain hearing aids as potential interventions to assist people with a normal audiogram who have difficulties hearing speech in noise hear and process speech better in complex listening environments. Listening was assessed with and without a device in an experimental group wearing mild-gain hearing aids and a placebo control group using retrospective questionnaires, EMA, speech-in-noise testing, and mental effort testing.

Regarding the retrospective questionnaires, experimental participants reported significantly lower levels of hearing-in-noise difficulties when they were fitted with mild-gain hearing aids compared to no device, whereas the placebo control group showed no difference. Additionally, experimental participants reported significantly higher satisfaction with the devices than those in the placebo control group. This demonstrates that mild-gain hearing aids reduce self-reported hearing-in-noise difficulties and present similar satisfaction levels to those reported by individuals with hearing loss (Cox & Alexander, 1999). These findings are consistent with Roup et al. (2018) who found that participants with hearing difficulties reported significantly less hearing handicap when they were wearing mild-gain hearing aids compared to no device. These results are also consistent with Singh and Doherty (2020) who found that participants with self-reported difficulty hearing in background noise had significantly lower hearing handicap scores after trialling a hearing aid for two weeks.

In regard to the real-world EMA, only the experimental group reported a significantly better hearing experience when they were aided compared to unaided. They also reported higher levels of device performance compared to the placebo control group. However, the experimental group reported greater acoustic challenge than the placebo control group. These results show that mild-gain hearing aids can provide a better hearing experience (i.e., improved understanding, participation, and mood). However, issues such as background noise and mental fatigue can cause problems for wearers. This is the first study using EMA to asses the effect of mild-gain hearing aids on participants with normal audiograms but difficulty hearing speech in noise so there are no prior studies assessing EMA results in this population to compare these results to.

Regarding the laboratory tests of speech-in-noise perception and mental effort, no significant differences were found for the experimental or placebo control participants between the unaided and aided conditions. This shows that mild-gain hearing aids do not provide a significant benefit for speech-in-noise perception or mental effort when measured by laboratory tests.

These findings are different to the findings of Roup et al. (2018) who found that speech recognition in noise performance on a traditional test was significantly better in the aided compared to unaided condition for both high-predictability and low-predictability sentences. This is the first study use a mental effort task to assess the effect of mild-gain hearing aids on participants with normal audiograms but difficulty hearing speech in noise so there are no prior studies assessing EMA results in this population to compare these results to.`

Overall, these results suggest that mild-gain hearing aids can assist people with a normal audiogram and speech-in-noise difficulties to have a better self-reported hearing experience in noisy environments. However, these perceived benefits are not observed in the laboratory tests of speech-in-noise and mental effort. Possible reasons for this difference are that the speech-in-noise and mental effort tests assess very specific processes, whereas the questionnaires and EMA provide a more holistic assessment of different factors that may affect the hearing and device experience rather than just speech intelligibility and effort. The factors included in the questionnaires and EMA include the listener's ability to follow conversations, ability to participate in conversations, mood, and personal image, all of which are important for an individual's experience in addition to speech intelligibility and effort. Additionally, it may be that the laboratory tests are not realistic enough representations of real-world listening. However, the ECO-SiN test was chosen as it uses stimuli drawn from natural conversations presented in real-world recorded background noise and is likely to give a more accurate picture of a person's listening abilities than traditional tests. In any event, both the ECO-SiN test and the mental effort tests are auditory-only tests so do not include visual cues that may aid speech intelligibility (Sueyoshi & Hardison, 2005). Another possibility is that people *perceived* that their hearing experience was better with hearing aids even though they were not understanding more words or demonstrating reduced listening effort. Though if this were the case, we would expect to see this better perceived experience for both the experimental and placebo control groups, which was not evident, hence a placebo effect can be ruled out.

Interestingly, despite the positive self-reports about mild-gain hearing aids in the experimental group (91% highlighted that the hearing aids had improved their hearing in noisy environments), no participants said that they would purchase the hearing aids at the recommended retail price of $5,000. This is similar to the findings of Roup et al. (2018) where only three of the 17 participants purchased the hearing aid after the trial. This low uptake was also found by Singh and Doherty (2020) where only two out of the ten participants with hearing difficulties would consider purchasing the hearing aids after the trial. This raises the questions of whether their hearing problem is worth the cost of the hearing aids. It may be that the participants do not visit situations that they have difficulty in often enough to warrant the cost of the hearing aids or that the impact on their quality of life is not large enough. This was also suggested by Singh and Doherty (2020). This opens up the possibility of hearables as a device for this population as they are a less expensive alternative to hearing aids. While 33% of clients in the earlier Mealings et al. (2020) study reported being ready and willing to try a personalised hearing aid, 62% of respondents were ready and willing to try a hearable. A hearable is a device that "fits in or on an ear that contains a wireless link, whether that's for audio, or remote control of audio augmentation" (Hunn, 2016). Hearables represent a morphing of hearing aids and consumer electronic devices to create customizable off-the-shelf devices (Taylor, 2015).

Hearables have several benefits over hearing aids (in addition to being less expensive). The stigma of traditional hearing aids can often be a barrier to their uptake (Wallhagen, 2010), whereas hearables are likely to be less stigmatized as they are associated with leisure activities such as Bluetooth music streaming (Taylor, 2015). The ability to make adjustments also gives a greater sense of autonomy and empowerment to the client (Maidment, Ali, & Ferguson, 2019).

Despite these benefits, the efficacy of hearables in people with a normal audiogram who also have speech-in-noise difficulties has only been assessed in one study (Valderrama, Mejiaa, Wong, Chong-White, & Edwards, 2023). The authors found that AirPods Pro provided a SNR advantage of +5.36 dB and speech intelligibility performance in the laboratory improved by

11.8%. Participants reported a hearing experience in noisy environments that was a bit better with the devices compared to without. This study also had a slightly higher uptake rate compared to the hearing aid studies with five out of 17 participants (29%) reporting that they would continue using AirPods Pro in the future.. Interestingly, however, SADL scores were much higher in the current study with hearing aids than with the AirPods Pro in the study by Valderrama et al. (2023). Therefore, hearing aids are more expensive, but seem to provide a better hearing experience compared to the AirPods Pro.

Personal sound amplification products have also been evaluated in a population with hearing loss by Cho et al. (2019) who found that speech perception, sound quality, listening effort, and user preference of personal sound amplification products versus basic and premium hearing aids did not differ for people with mild and moderate hearing loss. However, if the hearing loss was more severe, premium hearing aids performed better. Direct-to-consumer hearing devices (including hearables) have also been evaluated in term of capabilities, costs, and cosmetics (Almufarrij, Munro, Dawes, Stone, & Dillon, 2019). The authors found that higher purchase prices of the devices were generally associated with a better performance and willingness-to-wear, but noted that many products have poor sound quality and can produce uncomfortably loud sounds. They suggest that the challenge for manufacturers is “to develop low-cost products with cosmetic appeal and appropriate electroacoustic characteristics”. Therefore, hearables designed to meet this challenge may be a viable and affordable option for people with speech-in-noise difficulties but a normal audiogram.

4.1 Strengths of the study

This study has several strengths. First, it extends previous research by including a control group with a placebo hearing aid. This enabled us to rule out a placebo effect in our results which previous studies have been unable to do (Roup et al., 2018; Singh & Doherty, 2020). Second, the study uses a more realistic speech-in-noise test, the ECO-SiN test, with natural conversations presented in real-world recorded background noise rather than contrived speech stimuli in babble or speech-shaped noise that is found in many other tests (Miles et al., 2020). This helps to bridge the gap between laboratory testing and real-world experiences. Third, the study not only assessed speech intelligibility accuracy, but also mental effort, as listening to speech in noise can increase cognitive load and result in listeners taking longer to respond (Peelle, 2018). Fourth, in addition to traditional retrospective questionnaires, we also used EMA to give live real-world data of participants' hearing experience and reduce memory bias.

4.2 Limitations of the study

This study has several limitations which could be addressed in future research studies. First, the use of AutoSense OS both in laboratory and in real-life measures may have compromised the assessment of the hearing aids. We had no control over the activation of the '*Speech in noise'* program, in either the real-world or laboratory environments. If this program was not activated this could explain the non-significant differences between control and experimental groups in the speech-in-noise and mental effort tests.

Another limitation of the study is associated with COVID-19 restrictions, which has substantially affected the planned timing of the study. Seventeen participants commenced NEMA phase in February and March 2020, but due to COVID-19 restrictions, the remaining participants could only start this phase in July and August of that year. Due to COVID-19 restrictions, NEMA phase II could only start in October 2020. Moreover, since COVID-19 restrictions had substantially reduced the opportunities of the participants to socialise and attend noisy venues, a decision was made to extend the trialling period from the planned 6 weeks to 10 months in phase I (most of this time occurred while COVID-19 lock-downs) and from 6 weeks to 12 weeks in phase II. This decision enabled participants to (1) characterise their hearing difficulties in phase I, and (2) characterise their hearing experience with hearing aids in phase II. It shall be acknowledged that these factors influencing data collection may have introduced some bias. For example, given that about half of the participants completed phase I in summer and the other half in winter, lifestyle associated with season and weather could have influenced the participants' responses.

In addition, it shall also be noted that the number of completed surveys varies significantly amongst participants. Future studies characterising real-world hearing experience via NEMA shall benefit from setting clear targets to their participants as well as to follow them up along the trialling period in order to track progress and to assist them in any question they may have.

Further, it is noteworthy mentioning that the speech-in-noise and mental effort tests were audio-only. While this does represent some listening situations (e.g., listening to the radio or a podcast), many communication interactions occur when the listener can see the speaker's face. It would be beneficial for future research to use audio-visual stimuli as access to visual cues and gestures supports speech intelligibility (Sueyoshi & Hardison, 2005).

Finally, the question "Would you purchase the hearing aids (cost $5000)" could be better formulated and/or followed by as "How much would you spend?" to gain an idea of whether less expensive hearable devices are a more viable option or whether there is an association with the individual differences in the perceived value of the technology.

5. Conclusions

The results of this study suggest that mild-gain hearing aids can assist people with a normal audiogram who have difficulties hearing speech in noise to have a better self-reported hearing experience in noisy environments, even though differences are not observed in the laboratory. However, the price of these hearing aids is a barrier to their uptake. It would therefore be beneficial for future research to assess the efficacy of hearables in this population as a more affordable option to improve their hearing experience.

References

Almufarrij, I., Munro, K. J., Dawes, P., Stone, M. A., & Dillon, H. (2019). Direct-to-Consumer Hearing Devices: Capabilities, Costs, and Cosmetics. *Trends in Hearing*, *23*, 1–18. doi:10.1177/2331216519858301

Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare*, *9*, 211–217. doi:10.2147/JMDH.S104807

Bench, J., Kowal, A., & Bamford, J. (1979). The Bkb (Bamford-Kowal-ench) sentence lists for partially-hearing children'. *British Journal of Audiology*, *13*(3), 108–112. doi:10.3109/03005367909078884

Bernarding, C., Strauss, D. J., Hannemann, R., Seidler, H., & Corona-Strauss, F. I. (2017). Neurodynamic evaluation of hearing aid features using EEG correlates of listening effort. *Cognitive Neurodynamics*, *11*(3), 203–215. doi:10.1007/s11571-017-9425-5

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, *236*(4798), 157–161. doi:10.1126/science.3563494

Cho, Y. S., Park, S. Y., Seol, H. Y., Lim, J. H., Cho, Y. S., Hong, S. H., & Moon, I. J. (2019). Clinical performance evaluation of a personal sound amplification product vs a basic hearing aid and a premium hearing aid. *JAMA Otolaryngology - Head and Neck Surgery*, *145*(6), 516–522. doi:10.1001/jamaoto.2019.0667

Cox, R. M., & Alexander, G. C. (1999). Measuring Satisfaction with Amplification in Daily Life: The SADL Scale. *Ear and Hearing*, *20*(4), 306–320. doi:10.1097/00003446-199908000-00004

Desjardins, J. L. (2016). The effects of hearing aid directional microphone and noise reduction processing on listening effort in older adults with hearing loss. *Journal of the American*

Academy of Audiology, *27*(01), 029–041. doi:10.3766/jaaa.15030

Gagné, J.-P., Besser, J., & Lemke, U. (2017). Behavioural assessment of listening effort using a dual-task paradigm: A review. *Trends in Hearing*, *21*, 233121651668728. doi:10.1177/2331216516687287

Galvez, G., Turbin, M. B., Thielman, E. J., Istvan, J. A., Andrews, J. A., & Henry, J. A. (2012). Feasibility of ecological momentary assessment of hearing difficulties encountered by hearing aid users. *Ear and Hearing*, *33*(4), 497–507. doi:10.1097/AUD.0b013e3182498c41

Gates, G. A., & Hoffman, H. (2011). *What the numbers mean: An epidemiological perspective on hearing*. *National Institute on Deafness and Other Communication Disorders, from the U.S. Department of Health and Human Services*. Retrieved from https://www.nidcd.nih.gov/health/statistics/what-numbers-mean-epidemiological-perspective-hearing

Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scandinavian Audiology*, *11*(2), 79–87. doi:10.3109/01050398209076203

Hasan, S. S., Chipara, O., Wu, Y. H., & Aksan, N. (2014). Evaluating auditory contexts and their impacts on hearing aid outcomes with mobile phones. *Proceedings - PERVASIVEHEALTH 2014: 8th International Conference on Pervasive Computing Technologies for Healthcare*, (July), 126–133. doi:10.4108/icst.pervasivehealth.2014.254952

Henry, J. A., Galvez, G., Turbin, M. B., Thielman, E. J., McMillan, G. P., & Istvan, J. A. (2012). Pilot study to evaluate ecological momentary assessment of tinnitus. *Ear and Hearing*, *33*(2), 179–290. doi:10.1097/AUD.0b013e31822f6740

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology*, *49*(12), 891–903. doi:10.3109/14992027.2010.506889

Holube, I., von Gablenz, P., & Bitzer, J. (2020). Ecological momentary assessment in hearing research: current state, challenges, and future directions. *Ear and Hearing*, *41*, 79S-90S. doi:10.1097/AUD.0000000000000934

Hunn, N. (2016). *The Market for Hearable Devices 2016-2020*. *WiFore Wireless Consulting*. Retrieved from http://bit.ly/hearables2020

Keidser, G., Dillon, H., Flax, M., Ching, T., & Brewer, S. (2011). The NAL-NL2 prescription procedure. *Audiology Research*, *1*(1), e24. doi:10.4081/audiores.2011.e24

Kelly, H., Lin, G., Sankaran, N., Xia, J., Kalluri, S., & Carlile, S. (2017). Development and evaluation of a mixed gender, multi-talker matrix sentence test in Australian English. *International Journal of Audiology*, *56*(2), 85–91. doi:10.1080/14992027.2016.1236415

Kumar, G., Amen, F., & Roy, D. (2007). Normal hearing tests: is a further appointment really necessary? *Journal of the Royal Society of Medicine*, *100*, 66.

Maidment, D. W., Ali, Y. H. K., & Ferguson, M. A. (2019). Applying the COM-B model to assess the usability of smartphone-connected listening devices in adults with hearing loss. *Journal of the American Academy of Audiology*, *30*(5), 417–430. doi:10.3766/jaaa.18061

Mealings, K., Yeend, I., Valderrama, J. T., Gilliver, M., Pang, J., Heeris, J., & Jackson, P. (2020). Discovering the unmet needs of people with difficulties understanding speech in noise and a normal or near-normal audiogram. *American Journal of Audiology*, *29*(3), 329–355. doi:10.1044/2020_AJA-19-00093

Miles, K. M., Keidser, G., Freeston, K., Beechey, T., Best, V., & Buchholz, J. M. (2020). Development of the Everyday Conversational Sentences in Noise test. *The Journal of the Acoustical Society of America*, *147*(3), 1562–1576. doi:10.1121/10.0000780

Misra, S. (2012). Randomized double blind placebo control studies, the “Gold Standard” in intervention based studies. *Indian Journal of Sexually Transmitted Diseases and AIDS*, *33*(2), 131. doi:10.4103/0253-7184.102130

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*, *95*(2), 1085–1099. doi:10.1121/1.408469

Noble, W., Jensen, N. S. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. *International Journal of Audiology*, *52*(6), 409–412. doi:10.3109/14992027.2013.781278

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing*, *39*(2), 204–214. doi:10.1097/AUD.0000000000000494

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., … Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37*, 5S-27S. doi:10.1097/AUD.0000000000000312

Ricketts, T., & Henry, P. (2002). Evaluation of an adaptive, directional-microphone hearing aid: Evaluación de un auxiliar auditivo de micrófono direccional adaptable. *International Journal of Audiology*, *41*(2), 100–112. doi:10.3109/14992020209090400

Roup, C. M., Post, E., & Lewis, J. (2018). Mild-gain hearing aids as a treatment for adults with self-reported hearing difficulties. *Journal of the American Academy of Audiology*, *29*(6), 477–494. doi:10.3766/jaaa.16111

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*(1), 1–32. doi:10.1146/annurev.clinpsy.3.022806.091415

Singh, J., & Doherty, K. A. (2020). Use of a mild-gain hearing aid by middle-age normal-hearing adults who do and do not self-report trouble hearing in background noise. *American Journal of Audiology*, 1–10. doi:10.1044/2020_AJA-19-00051

Spankovich, C., Gonzalez, V. B., Su, D., & Bishop, C. E. (2018). Self reported hearing difficulty, tinnitus, and normal audiometric thresholds, the National Health and Nutrition

Examination Survey 1999–2002. *Hearing Research*, *358*, 30–36. doi:10.1016/j.heares.2017.12.001

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661–699. doi:10.1111/j.0023-8333.2005.00320.x

Timmer, B. H. B., Hickson, L., & Launer, S. (2017). Ecological momentary assessment: Feasibility, construct validity, and future applications. *American Journal of Audiology*, *26*(3S), 436–442. doi:10.1044/2017_AJA-16-0126

Tremblay, K. L., Pinto, A., Fischer, M. E., Klein, B. E. K., Klein, R., Levy, S., … Cruickshanks, K. J. (2015). Self-reported hearing difficulties among adults with normal audiograms: The Beaver Dam offspring study. *Ear & Hearing*, *36*(6), 290–299.

Valderrama, J. T., Mejiaa, J., Wong, A., Chong-White, N., & Edwards, B. (2023). The value of Headphone Accommodations in Apple AirPods Pro for managing speech-in-noise hearing difficulties of individuals with normal audiograms. *International Journal of Audiology*.

Valderrama, J., Wong, A., Galloway, J., Mejia, J., Herbert, N., & Edwards, B. (2022). The use of binaural beamforming to reduce listening effort. In *45th Association for Research in Otolaryngology (ARO) Annual Midwinter Meeting 2022*. San Jose, CA, CA. Retrieved from https://pmg.dcatalog.com/v/4/

Wallhagen, M. I. (2010). The stigma of hearing loss. *Gerontologist*, *50*(1), 66–75. doi:10.1093/geront/gnp107

Watson, S. (2015). Phonak AutoSense OS: The evolution of automation technology. *AudiologyOnline*, *13646*.

Weisser, A., Buchholz, J. M., Oreinos, C., Badajoz-Davila, J., Galloway, J., Beechey, T., & Keidser, G. (2019). The Ambisonic Recordings of Typical Environments (ARTE) Database. *Acta Acustica United with Acustica*, *105*(4), 695–713. doi:10.3813/AAA.919349

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. doi:10.1038/sdata.2016.18

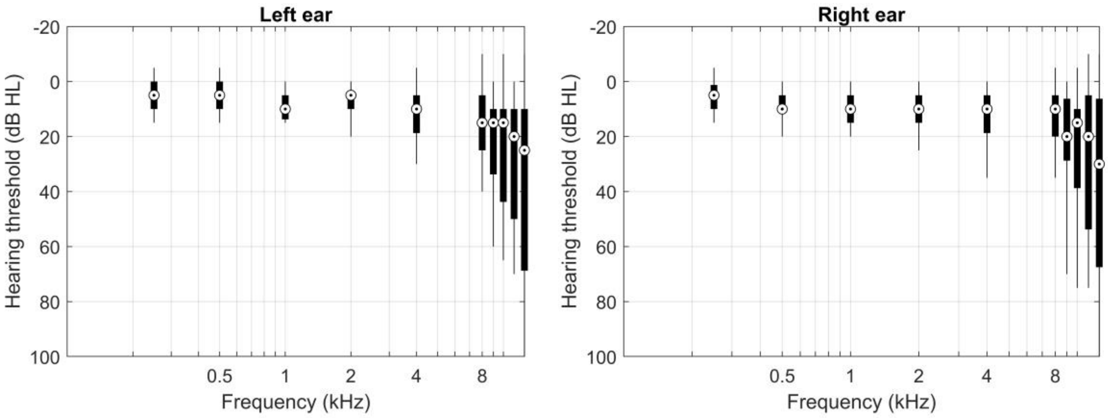
Wouters, J., vanden Berghe, J., & Maj, J.-B. (2002). Adaptive noise suppression for a dual-microphone hearing aid: Supresión adaptativa del ruido para un auxiliar auditivo con micrófono dual. *International Journal of Audiology*, *41*(7), 401–407. doi:10.3109/14992020209090417

Wu, Y.-H., Stangl, E., Zhang, X., & Bentler, R. A. (2015). Construct validity of the ecological momentary assessment in audiology research. *Journal of the American Academy of Audiology*, *26*(10), 872–884. doi:10.3766/jaaa.15034

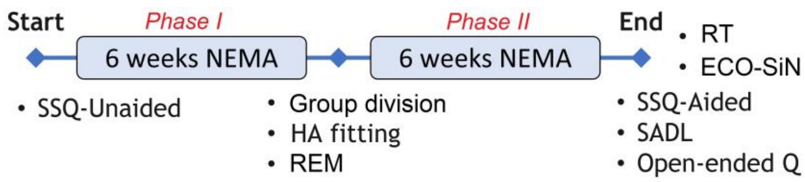
Zhao, F., & Stephens, D. (2007). A critical review of King-Kopetzky syndrome: Hearing difficulties, but normal hearing? *Audiological Medicine*, *5*(2), 119–124. doi:10.1080/16513860701296421

Figures

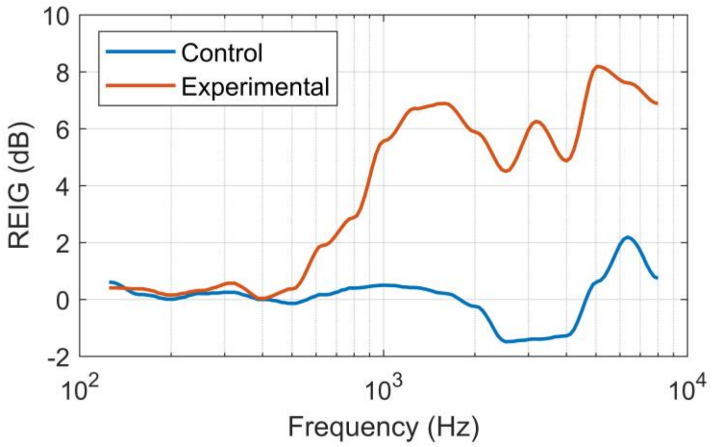
- Figure 1: Pure-tone hearing thresholds from 0.25 to 12.5 kHz in left and right ears. The central mark represents the median, the box edges are the 25th and 75th percentiles, and the whiskers are the maximum and minimum values.



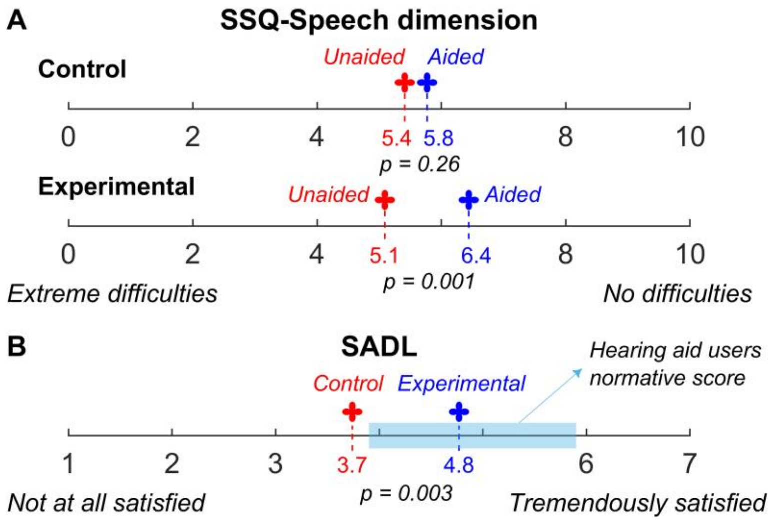
- Figure 2, single column, caption: The study design involved two phases and three appointments at the NAL (represented with diamonds). SSQ – Speech, Spatial and Qualities of Hearing Scale (Noble et al., 2013). NEMA – National Acoustic Laboratories-Ecologically Momentary Assessment. HA – hearing aid. REM – real ear measurements. RT – reaction time test to measure mental demand. ECO-SiN – Everyday Conversational Sentences in Noise (Miles et al., 2020) to measure speech-in-noise hearing performance. SADL – Satisfaction with Amplification in Daily Life (Cox & Alexander, 1999). Open-ended Q – end-of-study questionnaire based on open-ended questions.



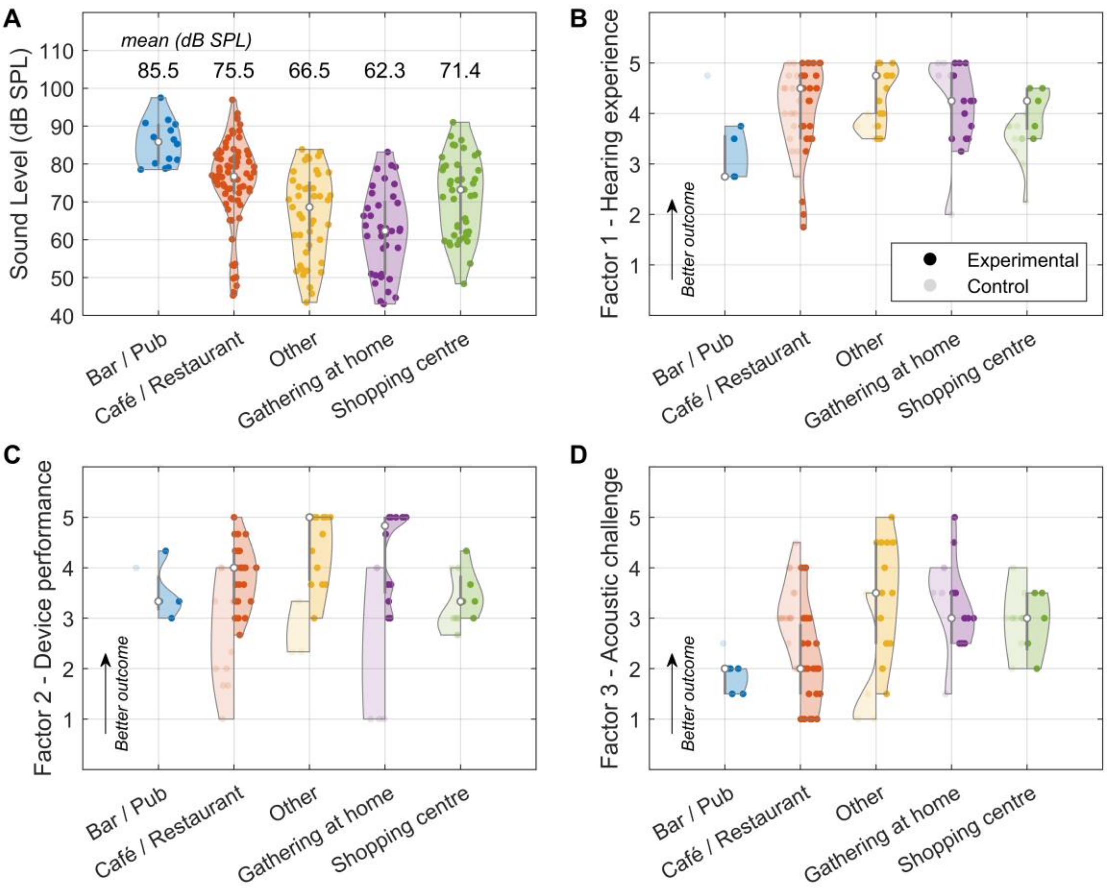
- Figure 3: Averaged real ear insertion gain (REIG) across participants for the control and experimental groups.



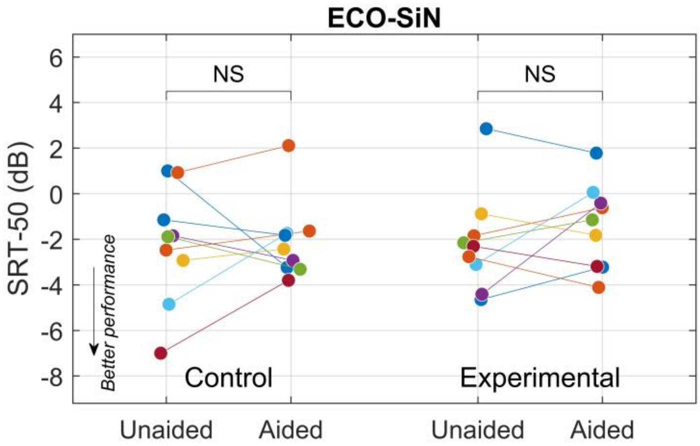
- Figure 4: [A] Mean score of the speech dimension of the Speech, Spatial and Qualities (SSQ) questionnaire (Noble et al., 2013) for control (top) and experimental (bottom) participants in the unaided (red) and aided (blue) conditions. [B] Mean score of the Satisfaction with Amplification in Daily Life (SADL; Cox & Alexander, 1999) for control (red) and experimental (blue) participants. Normative score obtained from a large cohort of hearing aid users is shown in a semitransparent blue rectangle (Cox & Alexander, 1999).



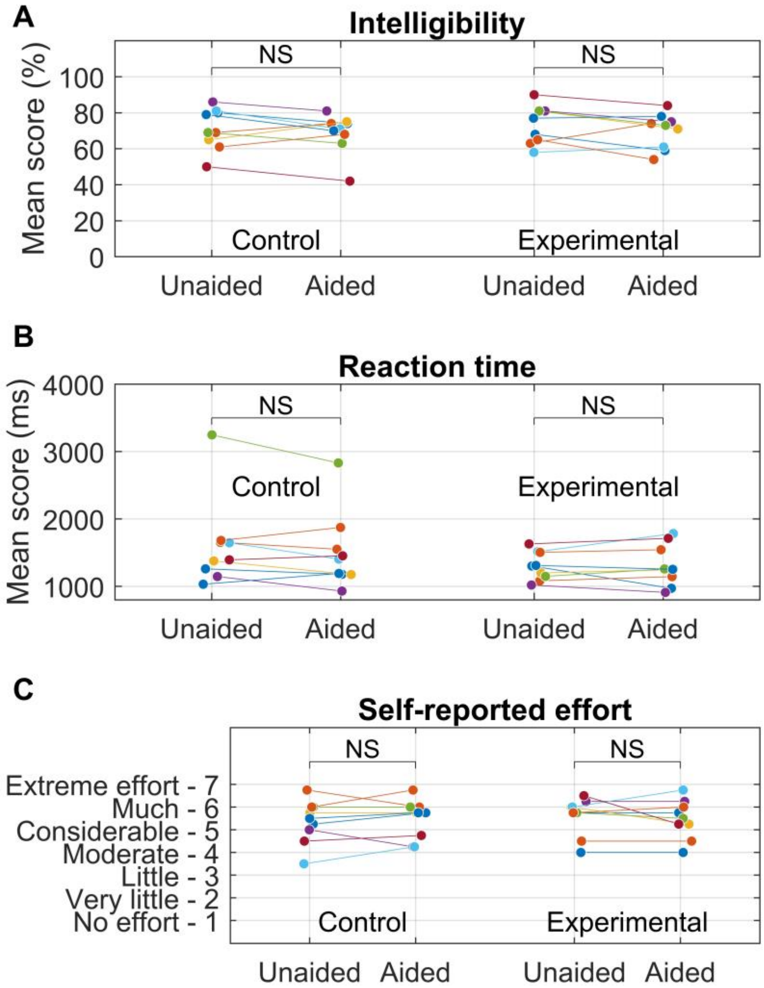
- Figure 5: The figures show the distribution of EMA entries separated by the location where the survey was conducted. The contour lines described the fitted distribution of the data, projected on the vertical axis. Each location is separated into experimental (left) and control (right) group responses. [A] Sound level distributions per sound environment. Mean level in dB SPL is shown at the top, and the median of each distribution is presented as a white circle. [B–D] Factor score distributions per sound environment for the control (left, translucid) and experimental (right, opaque) groups. Higher factor scores represent better outcomes, and a score equal to 3 represents neutrality.



- Figure 6: Signal to noise ratio at which participants presented 50% intelligibility (i.e. 50% speech reception threshold, SRT-50) in the Everyday Conversational Sentences in Noise test (ECO-SiN) for control (left) and experimental (right) participants, in *unaided* and *aided* conditions. Different colours represent different participants. NS – non statistically-significant difference.



- Figure 7: Dual task outcomes by control and experimental participants in the unaided and aided conditions, including mean scores of intelligibility (panel A), reaction time (panel B), and self-reported effort (panel C).



Tables

- Table 1. Pearson's correlation coefficient between NEMA surveyed items and the three factors resulting from a factor analysis aimed at reducing the dimensionality of the dataset.

Survey item	Factor 1. Hearing experience	Factor 2. Device	Factor 3. Listening effort
Noisiness	0.03	0.01	0.92
Mental demand	0.37	0.05	0.78
Participation	0.78	0.10	0.16
Frustration	0.82	-0.03	0.20
Understand	0.77	0.03	0.25
Mood	0.83	0.10	0.07
Benefit	-0.06	0.94	0.12
Quality	0.16	0.74	0.02
Satisfaction	0.21	0.87	0.02

- Table 2. Participants' characteristics of available data per test. C – Control. E – Experimental. SSQ–Speech, Spatial and Qualities of Hearing Scale (Noble et al., 2013). NEMA – National Acoustic Laboratories-Ecologically Momentary Assessment. SADL – Satisfaction with Amplification in Daily Life (Cox & Alexander, 1999). ECO-SiN – Everyday Conversational Sentences in Noise (Miles et al., 2020). RT – reaction time test to measure mental demand.

	Group	Sample size	Females	Age mean [range]
Hearing tests, SSQ questionnaire, Real ear measures	C	14	9	40.8 [19-63]
	E	13	8	44.8 [31-68]
NEMA-Phase I	C	10	7	45.3 [29-63]
	E	11	6	44.0 [31-57]
NEMA-Phase II	C	7	4	45.1 [31-56]
	E	11	6	44.0 [31-57]
End-of-study questionnaires (SSQ, SADL, Open-ended Q)	C	14	9	40.8 [19-63]
	E	11	6	43.5 [31-57]
Laboratory measures (ECO-SiN, RT)	C	9	6	40.9 [31-63]
	E	9	4	43.0 [31-57]

Table 3. Median score and statistical evaluation of groups differences within the three factors conducted via Tukey's honestly significant difference tests (Tukey's HSD) for multiple comparisons. Higher scores in the three factors represent better outcomes, and a score equal to 3 represents neutrality. SE – Standard error. q – Tukey's critical value. $q_{0.05}$ – Tukey's studentized range statistic for α = 0.05.]

Groups comparison	**Median**	**Difference**	**SE**	**q**	**$q_{0.05}$**	**p-value**
Factor 1. Hearing experience						
EXP-Unaided vs EXP-Aided	[3.25 vs 4.25]	20.73	3.344	6.201	3.633	<0.0001
CTR-Unaided vs CTR-Aided	[3.50 vs 3.75]	8.93	4.282	2.085	3.633	0.0858
CTR-Unaided vs EXP-Unaided	[3.50 vs 3.25]	8.15	3.508	2.322	3.633	0.6109
CTR-Aided vs EXP-Aided	[3.75 vs 4.25]	3.66	4.148	0.882	3.633	1.0000
Factor 2. Device performance						
CTR-Aided vs EXP-Aided	[3.00 vs 4.00]	27.23	4.426	6.152	2.772	0.0001
Factor 3. Acoustic challenge						
EXP-Unaided vs EXP-Aided	[3.00 vs 2.50]	3.39	3.344	1.013	3.633	1.0000
CTR-Unaided vs CTR-Aided	[3.00 vs 3.00]	3.38	4.282	0.788	3.633	1.0000
CTR-Unaided vs EXP-Unaided	[3.00 vs 3.00]	9.58	3.508	2.73	3.633	0.2095
CTR-Aided vs EXP-Aided	[3.00 vs 2.50]	16.34	4.148	3.939	3.633	0.0224

Supplementary material

- Appendix A: Raw data of participant’s audiometric results and questionnaire responses.
- Appendix B: NEMA data.