



Measuring and modeling context-dependent preferences for hearing aid settings

Alessandro Pasta^{1,4} · Michael Kai Petersen^{2,3} · Kasper Juul Jensen⁴ ·
Niels Henrik Pontoppidan³ · Jakob Eg Larsen¹ ·
Jeppe Høy Christensen³

Received: 19 June 2020 / Accepted in revised form: 4 March 2022 / Published online: 12 May 2022
© The Author(s) 2022

Abstract

Despite having individual perceptual preferences toward sounds, hearing aid users often end up with default hearing aid settings that have no contextual awareness. However, the introduction of smartphone-connected hearing aids has enabled a rethinking of hearing aids as user-adaptive systems considering both individual and contextual differences. In this study, we aimed to investigate the feasibility of such context-aware system for providing hearing aid users with a number of relevant hearing aid settings to choose from. During normal real-world hearing aid usage, we applied a smartphone-based method for capturing participants' listening experience and audiological preference for different intervention levels of three audiological parameters (Noise Reduction, Brightness, Soft Gain). Concurrently, we collected contextual data as both self-reports (listening environment and listening intention) and continuous data logging of the acoustic environment (sound pressure level, signal-to-noise ratio). First, we found that having access to different intervention levels of the Brightness and Soft Gain parameters affected listening satisfaction. Second, for all three audiological parameters, the perceived usefulness of having access to different intervention levels was significantly modulated by context. Third, contextual data improved the prediction of both explicit and implicit intervention level preferences. Our findings highlight that context has a significant impact on hearing aid preferences across participants and that contextual data logging can help reduce the space of potential interventions in a user-adaptive system so that the most useful and preferred settings can be offered.

✉ Alessandro Pasta
alpas@dtu.dk

¹ Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark

² Centre for Technology Entrepreneurship, Technical University of Denmark, Kongens Lyngby, Denmark

³ Eriksholm Research Centre, Oticon A/S, Snekkersten, Denmark

⁴ Demant A/S, Smørum, Denmark

Moreover, the proposed mixed-effects model is suitable for capturing predictions on an individual level and could also be expanded to predictions on a group level by including relevant user features.

Keywords Contextual awareness · User-adaptive system · User preferences · User satisfaction · Hearing healthcare · Hearing aids

1 Introduction

1.1 Background

Hearing aid users with similar hearing loss perceive sounds in highly individual ways, exhibiting differences in the ability to understand speech in noisy environments (Killion 2002), in the loudness perception (Oetting et al. 2018), and in the perception of sounds close to their hearing threshold (Marozeau and Florentine 2007). Despite that, the prescription of hearing aid amplification is primarily based on pure-tone audiometry, a test that measures the hearing thresholds for tonal stimuli at typically eight different frequencies (Walker et al. 2013). Pure-tone audiometry is a threshold test of signal detection but does not adequately represent real-world hearing abilities (Killion 2002; Baguley et al. 2016), because it does not convey information about central auditory processing, nor the auditory processing of real-world signals (Musiek et al. 2017). For these reasons, the initial prescription is considered a starting point rather than the optimal solution to treat a hearing loss (Abrams et al. 2011). A subsequent fine-tuning of the hearing aid might be performed in follow-up visits, during which the hearing care professional modifies the hearing aid settings based on users' recollections of past listening experiences (Kochkin et al. 2010). However, the success of fine-tuning depends on the hearing care professional's ability to interpret and translate users' recollections (Elberling and Hansen 1999; Arlinger et al. 2017). Moreover, this is a time-consuming procedure, often requiring multiple visits to obtain a satisfactory configuration (Abrams et al. 2011), and it does not guarantee a significant advantage over a default initial prescription (Cunningham et al. 2001; Shi et al. 2007). Crucially, hearing aid users are often hesitant to seek help from their hearing care professional, which highlights the importance of promoting user empowerment and self-management through new technology (Bennett et al. 2019). All in all, alternative user-driven ways of personalizing hearing aids are warranted.

Furthermore, hearing aid users report listening difficulties in different real-life situations ranging from face-to-face conversations to social interactions (Galvez et al. 2012). To cope with the different situations, hearing aid users seem to prefer switching between highly contrasting hearing aid settings depending on contextual variables such as sound environment and listening intention (Johansen et al. 2017; Korzepa et al. 2018). This emphasizes the importance of everyday context on users' listening experience and on their preferences toward specific hearing aid settings. To address a need for contextual adaptation, hearing aid users can currently be provided with different pre-configured programs. Such programs are aimed at improving the listening experience in specific contexts [e.g., speech in noise, music (Hockley et al. 2010)] by setting

predefined levels for different audiological parameters. Other hearing aid programs dynamically change their level of intervention according to the sensed environment. For instance, the level of noise reduction can be adjusted based on the ambient sound intensity levels (Schum 2003). However, such programs are based on the average hearing aid user and disregard the fact that listening preferences are highly individual (Brons et al. 2013). As a consequence, users tend not to use other programs than the default one (Nelson et al. 2006). Ultimately, this indicates that it is crucial to account both for individual preferences and for the impact of real-world context on users' listening experience and preferences when prescribing or fine-tuning hearing aid settings.

1.2 Related work

Several studies have documented that user-driven adjustments of hearing aids are feasible and potentially beneficial when investigated under controlled laboratory conditions (Yoon et al. 2017; Boothroyd and Mackersie 2017; Nelson et al. 2018; Jensen et al. 2019). Some studies have also reported on the benefits of training a personalized hearing aid program with a combination of sensor data and subjective preference feedback obtained from users in different real-life situations. However, results are mixed. Keidser and Alamudi (2013) trained the hearing aid settings using the SoundLearning algorithm (Chalupper et al. 2009), which learns and adjusts the amplification gain independently in four frequency bands and according to six different sound environments classified by the hearing aids (i.e., speech, speech and noise, quiet, noise, music, and car noise). They reported that 8 out of 18 participants preferred their trained hearing aid prescription in an evaluation phase (8 showed no preference and 2 preferred the un-trained prescription) and they concluded that training was efficient in those participants who initially wanted a change in their prescription. In another study, Aldaz et al. (2016) used smartphone-connected hearing aids to train settings with more "contextual awareness" by having participants perform A/B comparisons between the general program and context-specific programs alternating the microphone directionality or the noise reduction on and off. In the evaluation phase, the authors concluded that 7 out of 15 participants preferred the trained setting, 1 preferred the untrained setting, and 7 showed no preference. Notably, the learned preferences for microphone directionality and noise reduction were found to be nearly uniform across the different sound environments, which suggests that training did not effectively account for context. Overall, for roughly half of the participants in the two studies above (Keidser and Alamudi 2013; Aldaz et al. 2016), training was not efficient as the trained settings were not preferred outside of the training phase.

While training a personalized hearing aid configuration shows promise for specific individuals, it poses some challenges. Firstly, it assumes that users have one and only one preference in each context. However, when choosing between two alternative hearing aid settings in a specific sound environment, some users are consistent in their reported preferences, while others are not (Walravens et al. 2020). This inconsistency might be due to the fact that a selected setting does not yield a significant improvement in user experience and therefore leads to noisy preference assessments. Thus,

understanding when a preferred setting is perceived to truly improve the listening experience would help focusing on the relevant audiological parameters and contexts. Alternatively, the inconsistency might be due to an incomplete notion of context, with two situations classified under the same context resulting in two different preferences. For example, different listening intentions might require different hearing aid settings even though the sound environment does not change. Previous studies attempting to adjust hearing aid programs to contextualized and individualized preferences did not consider the listening intention.

Furthermore, learning context-dependent listening preferences requires gathering preferences on multiple audiological parameters and in several real-world contexts. Typically, listening experiences in hearing aid users are measured with experience sampling, i.e., having users explicitly provide in situ ratings of their listening experience (Shiffman et al. 2008). This method has proven successful in terms of documenting real-world benefits of hearing aid settings (Andersson et al. 2021). Despite being more reliable (Amatriain et al. 2009), explicit feedback is scarce and places a burden on the user (Jawaheer et al. 2010). Moreover, since the hearing aid signal processing acts on several parameters of the sound (e.g., frequency compression, gain amplification) with varying strengths (e.g., levels of amplification), the space of possible hearing aid settings is vast (Pasta et al. 2019). Therefore, when gathering user preferences in such a vast space and in several contexts, it is important to focus on settings and situations that cause a tangible improvement in user experience. Incidentally and importantly, data logging from modern hearing aids can provide information about implicit preferences toward hearing aid settings as well as information about the environmental context without imposing a burden on the user (Christensen et al. 2021). Previous research across different domains has shown that implicit and explicit feedback possess different characteristics and can complement each other (Jawaheer et al. 2010, 2014; Akehurst et al. 2012).

1.3 Research objective

The above-stated challenges with hearing aid prescription and personalization are addressed by investigating the feasibility of a context-aware user-adaptive system, which aims to offer a choice among relevant hearing aid settings based on collected preferences and contexts of many other users (Pasta et al. 2019). Specifically, we apply a method for capturing users' experiences and (explicit and implicit) audiological preferences for different intervention levels of three audiological parameters. The data are collected by smartphone-connected hearing aids, which enabled users to evaluate different settings during their everyday life. Concurrently, contextual data are acquired both through self-reporting and through continuous data logging. Importantly, all data collection is performed using the typical daily-life setup (i.e., a smartphone and a pair of hearing aids) of a hearing aid user.

First, we analyze if listening satisfaction is related to the perceived usefulness of an audiological parameter. A user-adaptive system should be able to offer a choice among settings that, when relevant, leads to higher user satisfaction. Thus, we gather and compare in situ ratings of listening satisfaction and of usefulness of choosing

among different intervention levels (henceforth called “choice-usefulness”) of the three audiological parameters.

Second, we analyze whether everyday contexts influence the choice-usefulness. Indeed, we hypothesized that context has a measurable and distinct impact on the explicitly reported usefulness of choosing among different intervention levels of the parameters. This would entail that a user-adaptive system can reduce the space of possible hearing aid settings by assigning context-aware usefulness to the audiological parameters.

Third, we apply statistical modeling of user preferences for different intervention levels and hypothesize that contextual predictors enable a better account of the observed preferences. If so, a user-adaptive system would benefit from contextual information when predicting the preferred levels of intervention for a specific audiological parameter. The statistical modeling is performed for both explicitly reported level preferences and for implicit preferences derived from user interactions. Indeed, if a system could rely only on implicit preferences, the training phase would be less burdensome for the user.

2 Methods

2.1 Participants

We recruited experienced hearing aid users having a hearing loss compatible with the Oticon Opn™ S1 MiniRITE hearing aids and being iOS users. Seven participants (6 men and 1 woman) with mean age 58 years ($SD = 12$ years) were recruited. Five of them were working, while two were retired. The participants all had more than 5 years of experience with hearing aid usage. All participants had a binaural hearing loss ranging from mild to moderately severe, as classified by the American Speech-Language-Hearing Association (Clark 1981). The study was approved by the Research Ethics Committees of the Capital Region of Denmark. Before the study began, all participants received written information about the study and gave their informed consent. One participant did not allow for contextual data collection and was therefore excluded from the analysis.

2.2 Apparatus

The participants were prescribed a pair of Oticon Opn™ S1 miniRITE (Oticon A/S, Smørum, Denmark) hearing aids and a frequency-specific amplification according to their hearing loss profile. All had iPhones with iOS 12 installed and additionally downloaded a custom smartphone app connected to the hearing aids via low-energy Bluetooth. Via the app, participants could control their hearing aid settings and submit in situ reports (see details in the Sect. 2.4). Furthermore, the app enabled continuous data logging of the active hearing aid settings and of the sound environment. The latter consisted of timestamped minute-based logs of the ambient acoustic environment

sensed by the hearing aid microphones (see Sect. 2.5). The app interface also included an open-ended response form for optional additional comments.

2.3 Audiological parameters

During the study, three audiological parameters were evaluated: Noise Reduction (NR), Brightness (BR), and Soft Gain (SG). Each parameter targeted a specific dimension of the sound with four levels of intervention. The Noise Reduction parameter provides varying strength of noise reduction and directionality depending on the selected level (in ascending order of intensity, from level 1 to level 4). Thus, level 1 provides the lowest level of noise reduction and directionality, which amplifies most sound sources coming from all directions. In contrast, level 4 suppresses all sounds classified as non-speech and only amplifies sounds coming from the frontal direction. The Brightness parameter adjusts the amplification gain for high frequencies (i.e., frequencies above 1.5 kHz), while the Soft Gain parameter adjusts the amplification gain for soft sounds (i.e., sounds below 50 dB SPL). Common to all parameters, levels 1 and 2 provide a lower intervention compared to the default prescription (i.e., the level that would be automatically prescribed by the fitting software), while levels 3 and 4 provide increased intervention compared to the default prescription. The three targeted audiological parameters have been shown to be particularly important for the listening experience of hearing aid users (Ng et al. 2013; Johansen et al. 2017; Wendt et al. 2017) and to be perceived differently by individuals (Killion 2002; Marozeau and Florentine 2007).

2.4 Procedure

The participants were instructed to use their hearing aids “as usual” in their everyday lives for 3 consecutive weeks and to regularly select and compare, via the supplied smartphone app, the four contrasting intervention levels. Only one audiological parameter was active in each week, while the others were temporarily set at default prescription levels. This was a deliberate design choice aimed at simplifying participant interactions (i.e., less settings to navigate) and ensuring that participants could consciously track the effects of their actions on their listening experience (Pasta et al. 2019). The order by which the parameters were evaluated was fixed (week 1: NR; week 2: BR; week 3: SG). A visualization of the study timeline that each participant went through is given in Fig. 1.

Each time the participants changed level, they had the option to submit an in situ report of their explicit preference (i.e., preferred intervention level from 1 to 4); their current listening satisfaction (Likert rating scale from 1 to 5); the usefulness of having a choice among the four contrasting levels (Likert rating scale from 1 to 5, henceforth referred to as “choice-usefulness”); and the listening intention, listening environment and the state of motion (e.g., stationary, walking) from predefined categories. Moreover, the level selections during normal hearing aid usage were logged and used to define implicit preferences (see Sect. 2.6).

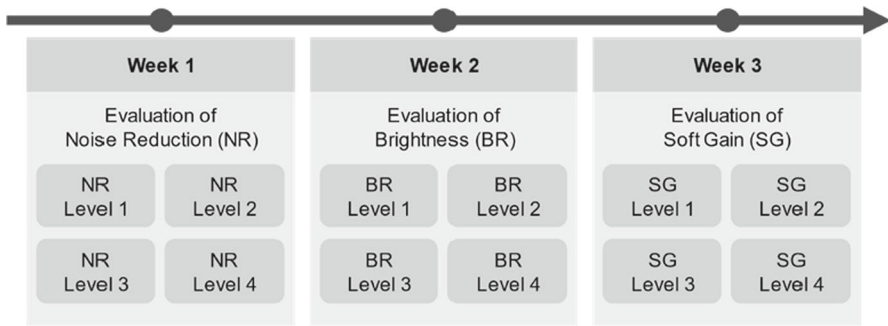


Fig. 1 Study timeline. Each parameter (Noise Reduction, Brightness, Soft Gain) was evaluated for the duration of one week. Each week, the participants were provided with four intervention levels of the parameter of the week

2.5 Contextual data

Self-reported context is represented by in situ reports of listening intention, listening environment, and motion state selected from drop-down lists with predefined categories. Note that for simplicity and due to sparse data, the motion state is not included in further analysis. In addition, due to the fairly low number of assessments received for some contexts (e.g., $n = 4$ for ‘Lecture’), categories were collapsed across similar contexts. Table 1 shows the labels for all possible listening intentions (Table 1a) and listening environments (Table 1b) before (‘Original label’) and after (‘New label’) collapsing.

Besides the self-reported context, timestamped acoustic data logged from the hearing aids measured the ambient sound pressure levels (SPLs) and signal-to-noise ratios (SNRs) in decibels across a broad frequency band (0.1–10 kHz) (Christensen et al. 2019). The SPL is the most used indicator of the sound wave strength and correlates well with human perception of loudness (Long 2014a). The SNR is the difference between the energy of a signal and the energy of any present noise, and it is the key to speech intelligibility (Long 2014b). Each in situ report and level selection were associated with acoustic data averaged across a 3-min preceding time-window.

2.6 Statistical modeling

Predictions of choice-usefulness and of explicit and implicit intervention level preferences were made using cumulative link proportional-odds mixed models. These models are ideal for multilevel modeling of longitudinal ordinal data (Hedeker 2008) and they are a class of the generalized mixed-effects modeling framework, which is popular among recommender systems (Condliff et al. 1999; Hedeker 2005; Chen et al. 2020b). Since the number of observations (in situ reports and level selections) for each participant varied, we included data from all participants into global models. The individual-level effects were modeled as random effects and estimated with partial pooling. Such random effects allow model predictions to differ among participants,

Table 1 Self-reported context (i.e., listening intention and listening environment) labels

(a) Listening intention				(b) Listening environment			
Original label	n	New label	n	Original label	n	New label	n
Ignore speech	12	Only me	128	Home	240	Inside/Quiet	282
Just me	116			Kitchen	12		
Meal	42	Lecture	4				
Music	42	Focus	170	Office	26	Outside/Noise	140
TV	86	Meeting	26				
Speech	84	Party	8				
Talk	108	Social	238	Public Indoor	54		
Meeting	8	Restaurant	22				
		Restroom	4				
		Heavy duty	4				
		Outdoor	16				
		Public outdoor	44				
		Traffic	4				
		Transport	72				

Original labels that the participants could select among, and new labels after collapsing across similar contexts

while partial pooling entails that, if a participant has fewer observations, her effect estimate will be partially based on the more abundant data from other participants. This is a good compromise between estimating an effect by completely pooling all users, which masks participant-level variation, and estimating an effect for all participants completely separately, which could give poor estimates for low-sample participants (Gelman and Hill 2006).

Prior to modeling, the continuous predictor SPL was converted into “Low intensity” and “High intensity”, while the continuous predictor SNR was converted into “Low quality” and “High quality”. This was done by using the median values for each participant as the cut-off between low and high. General recommendations for mixed-effects modeling were followed (Harrison et al. 2018). Fitting and supplementary statistics were performed in *R* using base functions and the ‘ordinal’ package (RDocumentation 2019).

For in situ reports, two separate models were applied for predicting the choice-usefulness rating and the explicitly preferred intervention level. The models were specified with both subjective and objective contextual predictors on the form:

$$\text{logit}(P(Y_i \leq j)) = \theta_j - \beta_1(\text{env}_i) - \beta_2(\text{intent}_i) - \beta_3(\text{SPL}_i) - \beta_4(\text{SNR}_i) - u(ID_i),$$

$$i = 1, \dots, n, j = 1, \dots, J - 1 \tag{1}$$

This is a model for the cumulative probability of the i th choice-usefulness rating (or preferred intervention level) falling in the j th category or below, where i indexes all observations and $j = 1, \dots, J$ indexes the response categories. In the model for choice-usefulness, $J = 5$. In the model for preferred intervention level, $J = 4$. θ_j are threshold parameters (or cut-points), which are assumed to be equidistant between the response categories. We take the participant effects (ID) to be random and assume that the effects are IID and normal: $u(ID_i) \sim N(0, \sigma_u^2)$. The self-reported listening environment (env) and listening intention (intent) are added as fixed effects predictors together with the categorical SPL and SNR.

The same model, but without the subjective contextual predictors, was applied to predict the implicit preferences (i.e., level selections) during normal hearing aid usage from user interaction event-logs. Note that only level selections that were kept for minimum three minutes were included as observations to the latter model. This was to ensure that random level selections (i.e., playing around) did not confound the outcome.

Besides inspection of coefficient magnitude and confidence intervals, likelihood ratio-tests based on the χ^2 test statistic were employed to test the significance of contextual predictors.

3 Results

3.1 Descriptive statistics: hearing aid usage and auditory ecology

Prior to assessing the main hypotheses of the study, we describe the main features of the collected data.

The number of level selections and the number of submitted assessments varied across participants (see Table 1). Overall, 8.8% ($SD = 3.0\%$) of all level selections led to an in situ report and a preference submission. This percentage did not differ markedly among the three audiological programs (Noise Reduction: $M = 9.8\%$; Brightness: $M = 10.3\%$; Soft Gain: $M = 6.8\%$), indicating a fair comparison of the programs.

The logged acoustic data documented that the participants had different exposure to different sound environments (see Table 2 and Fig. 2). However, importantly, there was agreement between the sound exposure they experienced during their normal device usage (i.e., changing levels throughout the day) and when submitting in situ self-reports (of their listening experience and level preferences). The scatter plots in Fig. 2 show the distribution of SPL (Fig. 2a) and SNR (Fig. 2b) as deciles measured either when performing in situ ratings (y-axis) or when changing levels throughout the day (x-axis). Notably, despite participant-specific offsets (e.g., participant n. 1 consistently experiences higher SPL during preference submission than during everyday device usage), the relationship between SPL deciles for preference submissions and level selections is linear with slope $\beta = 0.958$ ($F = 121.57$, $p < 0.001$).

On a group level, the relationship between SNR deciles for preference submission and level selections is also linear with slope $\beta = 0.729$ ($F = 37.00$, $p < 0.001$). However, participant n. 6 experiences, on average, much higher SNRs during everyday level selections than when performing in situ ratings, which indicates that most of the

Table 2 Participants' characteristics and data logs

Participant	Hearing loss (PTA) (dB)	In situ reports (count)	Level selections (count)	Percentage of selections leading to an in situ report (%)	Median SPL (dB)	Median SNR (dB)
1	67	36	858	4.2	56.40	19.17
2	47	21	227	9.3	49.54	14.76
3	44	19	248	7.7	56.28	10.11
4	36	131	1072	12.2	59.33	17.71
5	54	35	293	12.0	59.03	13.41
6	35	26	343	7.6	66.68	23.97
Mean	47	45	507	8.8	57.88	20.50
SD	11	43	363	3.0	5.57	6.35

PTA refers to the average of hearing threshold levels at four specified frequencies (0.5, 1, 2, 4 kHz). The median SNR and SPL are computed from all data-logs for each level selection

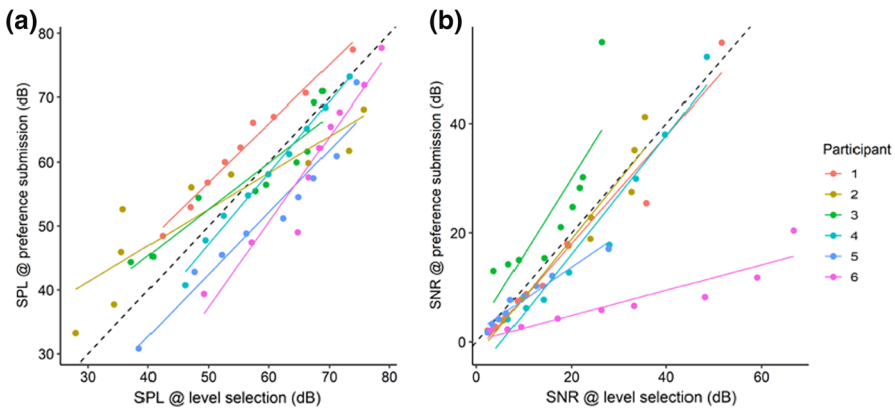


Fig. 2 Relationship between the distribution of SPL (a) and SNR (b) for in situ preference submissions (y-axis) and while selecting intervention levels during normal device usage (x-axis). Each dot represents the acoustic value at a decile (1st–9th) for one participant (colors). The dashed line indicates a slope of $y = x$. (Color figure online)

participant's ratings were performed under noisy or quiet conditions (i.e., low quality of the signal). Please note that the discrepancy might also be driven by the participant experiencing very high SNRs for some of the logged level selections.

We also assessed whether the self-reported contexts possessed different acoustic characteristics. If so, subjective self-reports conveyed more than the individual perception of auditory scenes. Figure 3 shows boxplots of each reported listening environment (Fig. 3a) and listening intention (Fig. 3b) against either SPL (top panels)

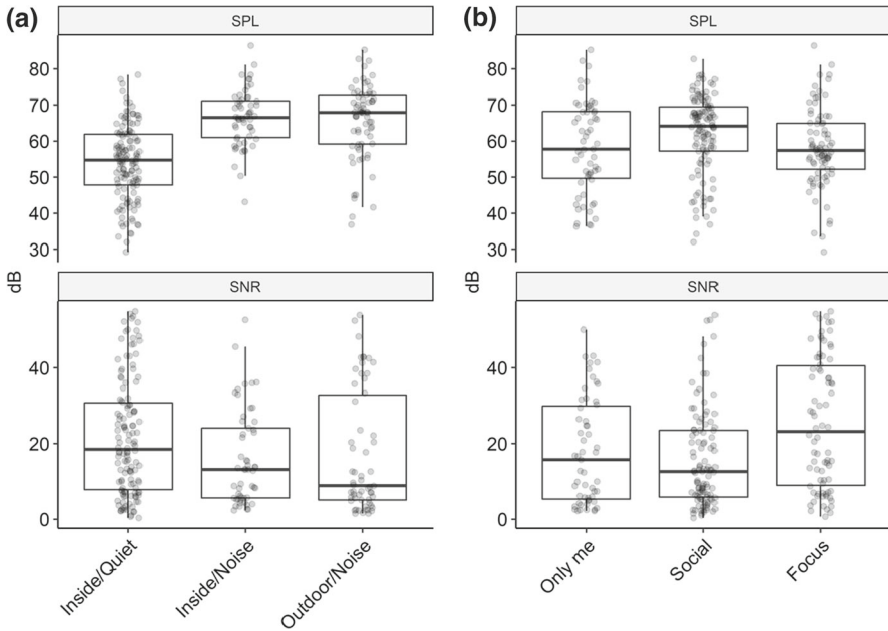


Fig. 3 SPL and SNR for self-reported listening environments (a) and listening intentions (b)

or SNR (bottom panels). Please note that the boxplots are based on pooled data among all participants.

The moderating effects of the self-reported contexts on SNR and SPL were evaluated by applying linear mixed-effects models. These models predict either SPL or SNR while controlling for time-of-day with random-effects offsets (e.g., SPL might simply be higher mid-day compared to end-of-day due to daily life activities). We observed main effects of listening environment on SPL ($F(2,221) = 42.844, p < 0.001$) and of listening intention on SNR ($F(2,229) = 3.450, p = 0.033$). For SPL, the largest effect was between listening environments “Inside/Quiet” and “Outdoor/Noise” ($\beta = 12.457, SE = 2.881, t = 4.324, p < 0.001$). For SNR, the largest effect was between listening intentions “Only me” and “Focus” ($\beta = 7.176, SE = 3.288, t = 2.182, p = 0.030$).

3.2 Relationship between choice-usefulness and listening satisfaction

Participants were asked to rate the usefulness of having four intervention levels available to choose from (i.e., choice-usefulness) and then to rate the current listening satisfaction. Thus, high ratings of choice-usefulness followed by high ratings of listening satisfaction are assumed to represent situations where audiological needs are met. To investigate how strongly a useful choice of intervention levels impacts satisfaction, we computed the correlation between the two types of ratings across all participants. Figure 4 shows contingency tables for each audiological parameter, which indicate a

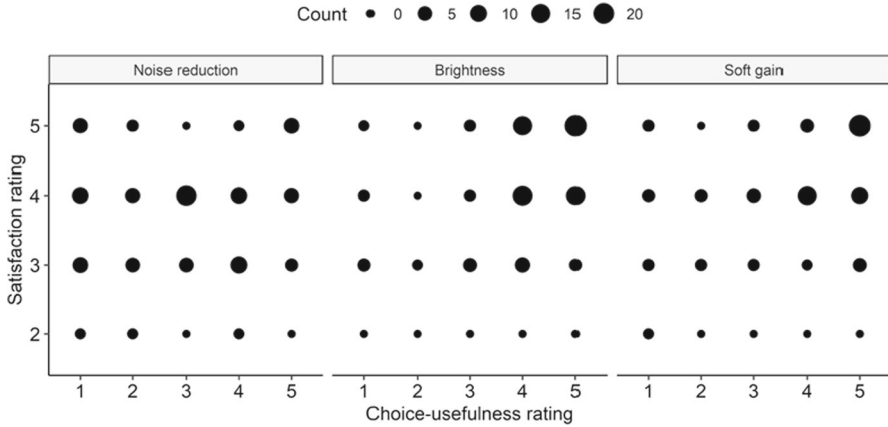


Fig. 4 Contingency tables for rating the listening satisfaction (y-axis) and the choice-usefulness (x-axis) of each audiological program. Data are pooled among all in situ assessments from all participants

stronger correlation for the Brightness and Soft Gain parameters than for the Noise Reduction parameter. Indeed, Pearson’s correlation tests revealed that satisfaction and choice-usefulness were not related when using the Noise Reduction parameter ($r = 0.064, t = 0.637, df = 96, p = 0.526$), but they were for the Brightness ($r = 0.383, t = 3.908, df = 89, p < 0.001$) and Soft Gain ($r = 0.400, t = 3.830, df = 77, p < 0.001$) parameters.

Separating the rating data by the contextual SNR revealed a higher correlation between satisfaction and choice-usefulness for Brightness in lower quality environments ($r = 0.443, t = 3.129, df = 40, p = 0.003$) compared to in higher quality environments ($r = 0.071, t = 0.423, df = 35, p = 0.675$), suggesting that having access to different levels of the Brightness parameter is more strongly associated with listening satisfaction when the quality of the listening environment is below the median. In contrast, Soft Gain exhibited higher correlation in high quality listening environments ($r = 0.572, t = 3.886, df = 31, p < 0.001$) compared to lower quality listening environments ($r = 0.366, t = 2.080, df = 28, p = 0.047$). The correlation between satisfaction and choice-usefulness for the Noise Reduction parameter was again not significant after splitting the data by SNR. In summary, the relationship between choice-usefulness and listening satisfaction is distinct among the audiological parameters and varies with the context (here, SNR).

3.3 Contextual impact on choice-usefulness and explicit level preferences

One of the main aims of the study is to investigate, for the three audiological parameters, the impact of context on the perceived choice-usefulness and on explicit level preferences. Ideally, this can lead to context-aware recommendations of audiological programs combining the most relevant parameters and intervention levels for each situation.

In this section, we investigate the contextual impact by applying mixed-modeling of the in situ ratings using both subjective and objective contextual predictors. Across the three audiological parameters, the contextual predictors (self-reported listening environment and intention, SPL, SNR) were found to significantly increase the prediction of choice-usefulness ratings (likelihood ratio test, $\chi^2(6) = 21.71, p = 0.002$) and intervention level preferences (likelihood ratio test, $\chi^2(6) = 14.418, p = 0.025$). Figure 5a, b shows the estimated coefficients when modeling data from each parameter separately with random effects offsets for participants (i.e., Eq. 1). Notably, listening intention, listening environment, and SPL modulated both the choice-usefulness and level preference.

The random effects offsets (Fig. 5c, d) indicate that participants had comparable ratings and level preferences (i.e., most falling within ± 1 SD). Nevertheless, a few outliers were observed. For instance, participant n. 4 consistently rated the Soft Gain choice-usefulness higher than 1 SD from the group mean and participant n. 6 rated it much lower than the group mean, albeit in the latter case, the large error bars indicate

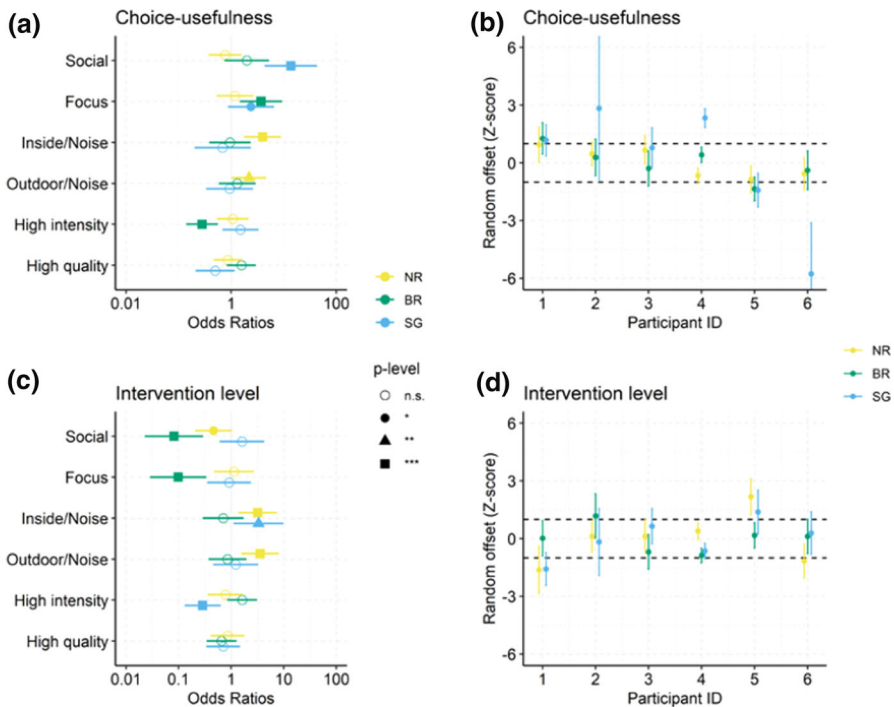


Fig. 5 In **a** and **b**, coefficients (as log odds ratios) and 95% confidence intervals for predicting choice-usefulness and explicit level preference from in situ ratings. In **c** and **d**, the corresponding random offsets due to participant effects. Note that the models were fitted separately for each audiological parameter (NR, BR, and SG). The baseline conditions for the contextual predictors were given as follows: “Only me” (for listening intention), “Quiet/Indoor” (for listening environment), “Low intensity” (for SPL), “Low quality” (for SNR)

that the estimated random offset is based on few observations. For Noise Reduction, participant n. 5 preferred significantly higher levels than the group mean.

3.4 Preference prediction from real-world usage patterns

The level preferences modeled in Figure 5 represent explicitly preferred levels. That is, levels that the participants purposefully reported as preferences. However, during normal real-world usage, participants made ~ 11 times more level selections than preference submissions (see Table 2), with automatically logged SPL and SNR associated with them. While some of these level selections were made to perform momentary comparisons, other level selections were made and used for longer periods of time. Participants made on average 377 active level selections (i.e., level selections that are set for at least three minutes), which is ~ 8 times more than preference submissions. Thus, a user-adaptive system could potentially leverage on these in case explicit feedback is not available.

We first assessed the contextual modulation of the implicit preferences by applying the statistical model in Eq. 1 to data from all participants. As was the case with the explicit preference data (Fig. 5), SPL and SNR significantly improved the model’s ability to predict intervention level ($\chi^2(8) = 17.43, p = 0.026$). The context-aware model prediction is shown in Fig. 6a as a red solid line together with both the observed preferences (dots with error bars) and the prediction from a NULL model—i.e., an intercept per program only model (blue solid line in Fig. 6a). Visually, differences in predictions between the two models are subtle. However, the contextual aware model does capture more variation in the observed preferences (Person’s correlation—NULL model: $r = 0.48, 95\% \text{ CI } [0.23\text{--}0.67], df = 46, p < 0.001$; Context-aware model: $r = 0.57, 95\% \text{ CI } [0.34\text{--}0.73], df = 46, p < 0.001$), which is evident in Fig. 6b with the

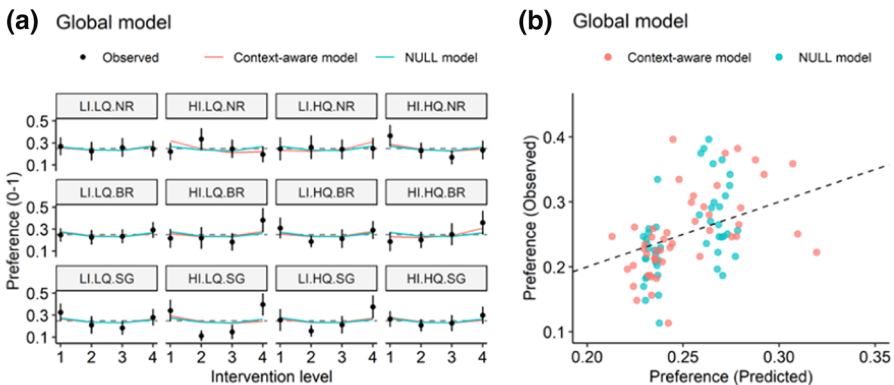


Fig. 6 Observed and predicted implicit preference for intervention level using data from all participants. In **a**, model predictions are shown together with the observed relative preference for each intervention grouped by all combinations of SPL and SNR (columns) and separated by audiological parameter (rows). LI = “Low intensity”; HI = “High intensity”; LQ = “Low quality”; HQ = “High quality”. In **b**, the difference between predicted and observed preference is shown as a scatter plot with the dashed line indicating a $y = x$ relationship

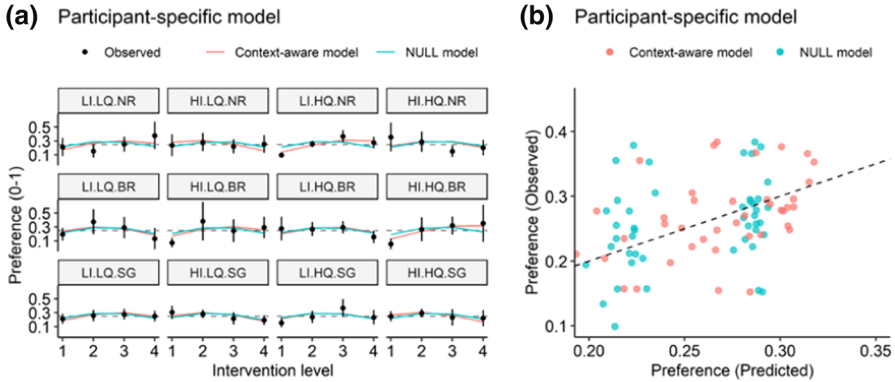


Fig. 7 Observed and predicted implicit preference for intervention level for participant n. 4. In **a**, model predictions are shown together with the observed relative preference for each intervention grouped by all combinations of SPL and SNR (columns) and separated by audiological parameter (rows). LI = “Low intensity”; HI = “High intensity”; LQ = “Low quality”; HQ = “High quality”. In **b**, the difference between predicted and observed preference is shown as a scatter plot with the dashed line indicating a $y = x$ relationship

context-aware model being able to predict a wider range of preferences. For example, for the “High intensity”/“High quality” condition with the Brightness parameter, the context-aware model is able to better fit the observed data.

In addition, we assessed how well the model performed on a single-user level by fitting it only on data gathered from participant n. 4 (the participant with most data logged, see Table 2). The predictions are shown in Fig. 7, and again SPL and SNR significantly improved the model fit ($\chi^2(8) = 19.97, p = 0.010$), and produced closer fitting predictions (Person’s correlation—NULL model: $r = 0.45, 95\% \text{ CI } [0.19-0.65], df = 46, p = 0.001$; Context-aware model: $r = 0.67, 95\% \text{ CI } [0.48-0.80], df = 46, p < 0.001$).

4 Discussion

This study applied a novel smartphone-based method for capturing real-world in situ experiences and preferences, combined with data of environmental sound logged from the hearing aids. By investigating the impact of everyday context on users’ listening experience and preferences, this study aimed to shed light on the feasibility of a context-aware user-adaptive system for providing useful audiological interventions.

Descriptive analysis of the collected data showed that the sound experienced when changing settings throughout the day was, on a group level, equal to that experienced when submitting self-reports (Fig. 2). However, specific participants (e.g., n. 6 in Fig. 2) showed a stronger deviation between the sounds experienced in the two situations. This may be either because self-reports are cognitively demanding (hence, they are completed in quiet environments only) or because reports are only submitted when problems are experienced (thus, leading to worse SNR for preference reports than for

normal level changes). Nevertheless, lack of representativeness of in situ preference and experience reports can, to some extent, be expected (Schinkel-Bielefeld et al. 2020; Ziesemer et al. 2020). Comparing the distributions of sound collected during self-report submission and everyday level changes (i.e., Fig. 2) can help validate participants' data. Consequently, more trust can be placed in the data collected from those participants who exhibit a relationship close to $\beta = 1$ between the sounds in the two situations. Moreover, we found that self-reported context both supports and differentiates the automatically logged contextual sound data (Fig. 3). The self-reported listening environments were associated with different loudness of the environment (i.e., SPL). The self-reported listening intentions were associated with different quality of the environment (i.e., SNR). In particular, "focused" listening intentions (e.g., watching TV) were associated with higher SNRs, indicating that the sound signals convey clean and relevant information. Conversely, "social" listening intentions exhibited low SNR and high SPL, suggesting that such sound environments are characterized by poor signals and loud noise. This analysis documents that self-reports not only reflect subjective perceptual evaluations of the listening scene, but also convey objective information that are relevant for a hearing aid adjustment. At the same time, objectively similar acoustic environments might imply different audiological needs according to the self-reported listening environments and intentions. This means that a more fine-grained resolution of user context can be obtained by combining self-reports with objective data logging.

We also investigated how the perceived usefulness (i.e., rated "choice-usefulness") of being offered a choice between different intervention levels of three independent audiological parameters affected the rated listening satisfaction. That is, if the intervention levels of an offered audiological parameter solve the listening needs of a user (or not), then the rated satisfaction should increase (or decrease). For the Noise Reduction parameter there was no correlation between satisfaction and choice-usefulness. This might be explained by the limited audibility of the change between the intervention levels of the parameter. Indeed, a substantial change (from 3 to 4 decibel) of acceptable noise level is required to yield a minimal clinically important and perceptual difference (Wong et al. 2018). Conversely, for the Brightness and Soft Gain parameters, significant correlations between satisfaction and choice-usefulness were observed and noted to be moderated by the quality of the sound environment (i.e., SNR). This implies that the listening experience can indeed be influenced by offering the user a choice among different intervention levels, and that the outcome depends on the parameter and on the context the user is in.

By applying statistical multi-level modeling, we examined the influence of context on the perceived usefulness of the offered intervention levels. In summary, self-reported and objective contexts have measurable and distinct impacts on the rated choice-usefulness. For instance, the Brightness parameter was significantly more useful in "Focus" listening intentions and in low intensity sound environments. This is consistent with previous studies showing that high-frequency amplification can be useful to improve speech understanding (Hornsby et al. 2011; Levy et al. 2015) and sound localization (Best et al. 2005). These findings suggest that a user-adaptive system can assign context-aware usefulness to the three audiological parameters based on previous user feedback. Contextual information would help filter the complex space

of possible hearing aid settings, by providing an indication about which parameter the user should be asked to adjust. As applied in this study, the proposed mixed-effects model (Eq. 1) accounts for individual differences by including a random term for each user, which enables user-level predictions. However, the model could be expanded by simply adding random terms for relevant user features, such as hearing loss, age, measures of auditory perception, and patterns of hearing aid use (Pasta et al. 2021). This would enable group-level predictions for users with similar features. In this way, users with sparse feedback (i.e., new users or users that do not supply explicit feedback) that share similar features with other users could benefit from the learned context-dependent preferences to alleviate the cold start problem (Chen et al. 2020a).

While choice-usefulness can help determine in which contexts a given audiological parameter should be adjusted, context-aware predictions of the preferred intervention level can help decide which levels of the parameter are the most important. Thus, the collected preferences for different intervention levels were modelled with contextual predictors. Across all participants, there was clear evidence that contextual data improve the prediction of the self-reported explicit level preferences. Coefficients from the statistical model (Fig. 5) revealed distinct predictors for the three audiological parameters: higher levels of Noise Reduction were preferred in noisier environments, but not in social situations; higher levels of Brightness were preferred when being alone; higher levels of Soft Gain were preferred in quieter environments. These effects provide an indication on the direction that should be taken (i.e., increasing or decreasing) when adjusting the level of each parameter depending on the context. This may result in more relevant levels proposed to the user, ensuring a more effective interaction and a more engaging experience. The relevance of both logged and self-reported context suggests that it is important to model context both by directly observable features (here, SPL and SNR), as well as by hidden features reflecting the user's specific intentions (e.g., enhancing speech or ignoring voices) and situational environment. Similarly, although the differences were subtle, we found that contextual information obtained from continuous data logging (SPL and SNR) improves predictions of implicit level preferences (i.e., preferences derived from usage patterns) both on a group level (Fig. 6) and user level (i.e., participant n. 4 in Fig. 7). Thus, a system aimed at autonomous preference prediction will benefit from continuously logging contextual data from the hearing aid microphones and from using device interactions (e.g., level selections) to assign a preference to the offered choices. In addition, while objective data logging cannot fully capture subjective listening intentions (Fig. 3) and implicit preferences are potentially less reliable (Amatriain et al. 2009), capitalizing on these data would help overcome the scarcity of user feedback and enable preference modeling of new or less engaged users.

5 Limitations

Real-world data have high ecological validity (Verma et al. 2017; Hicks et al. 2019) but also lack control for when, where, and how much data are logged. In that sense, a limitation of this study is that some participants collected less data than others.

However, in our statistical modeling, we specifically adjusted for effects of individual differences among participants by partial pooling (see Sect. 2).

Due to specific requirements for the participants to be included in the study (being experienced hearing aid users, having a hearing loss compatible with the Oticon Opn™ S1 MiniRITE hearing aids, being iOS users), the sample size of the study is rather small. However, the aim of the study was to evaluate the impact of everyday context on hearing aid users with similar features. Thus, the repeated-measures design (i.e., continuous data logging and repeated in situ reports) ensured that a high number of observations were acquired for statistical modeling of data representing everyday hearing aid usage. This helped compensating for the rather small sample size. A limitation remains in disentangling specific user features from the results as this would require a larger sample. Moreover, while the statistical modeling (Eq. 1) accounts for individual differences among the participants in terms of choice-usefulness and level preference random offsets, it does not account for individual effects of context on individual preferences. However, the inclusion of more participants with more repeated measures could enable an expansion of the statistical model to also account for participant-specific random slopes. That would enable modeling participant-specific sensitivity toward contextual predictors (Harrison et al. 2018; Gao et al. 2019).

The three audiological parameters were evaluated in chronological order. Thus, temporal effects (e.g., hearing aid acclimatization (Wright and Gagné 2021) or study fatigue) cannot be disentangled from the main results. Future research could assign random order of the parameters to each participant to investigate detailed differences among them without confounds from temporal effects.

6 Conclusions

Rethinking hearing aids as user-adaptive systems can provide a context-aware and personalized alternative to hearing aids with predefined settings. Our results show that participants' listening experience can effectively be influenced by providing a choice among different intervention levels of specific audiological parameters. Moreover, contextual data significantly improved predictions of how useful the offered choice among intervention levels were perceived to be. Additionally, contextual data significantly improved the prediction of both explicit and implicit level preferences.

We conclude that, when rethinking hearing aids as context-aware user-adaptive systems, both objective (i.e., SNR and SPL) and subjective (i.e., self-reported listening intention and environment) contextual data should be taken into consideration to optimize recommendations of the most relevant parameters and intervention levels. We propose training a proportional-odds mixed-effects model on preference and level selections data from experienced hearing aid users to provide context-aware recommendations to new users.

Acknowledgements We would like to thank Eriksholm Research Centre and Oticon A/S for providing hardware, access to test subjects, clinical approval, and clinical resources. We would also like to thank the test participants for their contribution and Rikke Rossing for the clinical support.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abrams, H., Edwards, B., Valentine, S., Fitz, K.: A patient-adjusted fine-tuning approach for optimizing hearing aid response. *Hear. Rev.* **18**, 18–27 (2011)
- Akehurst, J., Koprinska, I., Yacef, K., et al.: Explicit and implicit user preferences in online dating. In: Cao, L., Huang, J.Z., Bailey, J., et al. (eds.) *New Frontiers in Applied Data Mining. PAKDD 2011 International Workshop*. Springer, Berlin, Heidelberg, pp. 15–27. (2012)
- Aldaz, G., Puria, S., Leifer, L.J.: Smartphone-based system for learning and inferring hearing aid settings. *J. Am. Acad. Audiol.* **27**, 732–749 (2016). <https://doi.org/10.3766/jaaa.15099>
- Amatriain, X., Pujol, J.M., Oliver, N.: I Like It... I Like It Not: evaluating user ratings noise in recommender systems. In: Houben, G., McCalla, G., Pianesi, F., Zancanaro, M. (eds.) *User Modeling, Adaptation, and Personalization. UMAP 2009. Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, pp. 247–258. (2009)
- Andersson, K.E., Andersen, L.S., Christensen, J.H., Neher, T.: Assessing real-life benefit from hearing-aid noise management: SSQ12 questionnaire versus ecological momentary assessment with acoustic data-logging. *Am. J. Audiol.* **30**, 93–104 (2021). https://doi.org/10.1044/2020_AJA-20-00042
- Arlinger, S., Nordqvist, P., Öberg, M.: International outcome inventory for hearing aids: data from a large Swedish quality register database. *Am. J. Audiol.* **26**, 443–450 (2017). https://doi.org/10.1044/2017_AJA-16-0123
- Baguley, D.M., Cope, T.E., McFerran, D.J.: Functional auditory disorders. In: Hallett, M., Stone, J., Carson, A. (eds.) *Handbook of Clinical Neurology*, pp. 367–378. Elsevier (2016)
- Bennett, R.J., Laplante-Lévesque, A., Eikelboom, R.H.: How do hearing aid owners respond to hearing aid problems? *Ear Hear.* **40**, 77–87 (2019). <https://doi.org/10.1097/AUD.0000000000000595>
- Best, V., Carlile, S., Jin, C., van Schaik, A.: The role of high frequencies in speech localization. *J. Acoust. Soc. Am.* **118**, 353–363 (2005). <https://doi.org/10.1121/1.1926107>
- Boothroyd, A., Mackersie, C.: A “Goldilocks” approach to hearing-aid self-fitting: user interactions. *Am. J. Audiol.* **26**, 430–435 (2017). https://doi.org/10.1044/2017_AJA-16-0125
- Brons, I., Houben, R., Dreschler, W.A.: Perceptual effects of noise reduction with respect to personal preference, speech intelligibility, and listening effort. *Ear Hear.* **34**, 29–41 (2013). <https://doi.org/10.1097/AUD.0b013e31825f299f>
- Chalupper, J., Junius, D., Powers, T.: Algorithm lets users train aid to optimize compression, frequency shape, and gain. *Hear. J.* **62**, 26–33 (2009). <https://doi.org/10.1097/01.HJ.0000359131.58356.8f>
- Chen, W.-H., Hsu, C.-C., Lai, Y.-A., et al.: Attribute-aware recommender system based on collaborative filtering: survey and classification. *Front. Big Data* **2**, 1–30 (2020a). <https://doi.org/10.3389/fdata.2019.00049>
- Chen, Z., Zhu, S., Niu, Q., Zuo, T.: Knowledge discovery and recommendation with linear mixed model. *IEEE Access* **8**, 38304–38317 (2020b). <https://doi.org/10.1109/ACCESS.2020.2973170>
- Christensen, J.H., Pontoppidan, N.H., Rossing, R., et al.: Fully synthetic longitudinal real-world data from hearing aid wearers for public health policy modeling. *Front. Neurosci.* **13**, 1–5 (2019). <https://doi.org/10.3389/fnins.2019.00850>
- Christensen, J.H., Saunders, G.H., Porsbo, M., Pontoppidan, N.H.: The everyday acoustic environment and its association with human heart rate: evidence from real-world data logging with hearing aids and wearables. *R. Soc. Open Sci.* **8**, 201345 (2021). <https://doi.org/10.1098/rsos.201345>
- Clark, J.G.: Uses and abuses of hearing loss classification. *ASHA* **23**, 493–500 (1981)

- Condliff, M.K., Lewis, D.D., Madigan, D., Posse, C.: Bayesian mixed-effects models for recommender systems. In: ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Evaluation. (1999)
- Cunningham, D.R., Williams, K.J., Goldsmith, L.J.: Effects of providing and withholding postfitting fine-tuning adjustments on outcome measures in novice hearing aid users. *Am. J. Audiol.* **10**, 13–23 (2001). [https://doi.org/10.1044/1059-0889\(2001\)001](https://doi.org/10.1044/1059-0889(2001)001)
- Elberling, C., Hansen, V.K.: Hearing instruments: interaction with user preference. In: Rasmussen, A.N., Osterhammel, P.A., Andersen, T., Poulsen, T. (eds.) *Auditory Models and Non-linear Hearing Instruments*. Proceedings of the 18th Danavox Symposium, pp. 341–357. Copenhagen (1999)
- Galvez, G., Turbin, M.B., Thielman, E.J., et al.: Feasibility of ecological momentary assessment of hearing difficulties encountered by hearing aid users. *Ear Hear.* **33**, 497–507 (2012). <https://doi.org/10.1097/AUD.0b013e3182498c41>
- Gao, B., Zhan, G., Wang, H., et al.: Learning with linear mixed model for group recommendation systems. In: Proceedings of the 2019 11th International Conference on Machine Learning and Computing—ICMLC '19, pp. 81–85. ACM Press, New York (2019)
- Gelman, A., Hill, J.: *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge (2006)
- Harrison, X.A., Donaldson, L., Correa-Cano, M.E., et al.: A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ* **6**, e4794 (2018). <https://doi.org/10.7717/peerj.4794>
- Hedeker, D.: Generalized linear mixed models. In: Everitt, B.S., Howell, D.C. (eds.) *Encyclopedia of Statistics in Behavioral Science*, pp. 729–738. Wiley, Chichester (2005)
- Hedeker, D.: Multilevel models for ordinal and nominal variables. In: De Leeuw, J., Meijer, E. (eds.) *Handbook of Multilevel Analysis*, pp. 237–274. Springer, New York (2008)
- Hicks, J.L., Althoff, T., Sosic, R., et al.: Best practices for analyzing large-scale health data from wearables and smartphone apps. *npj Digit. Med.* **2**, 45 (2019). <https://doi.org/10.1038/s41746-019-0121-1>
- Hockley, N.S., Bahlmann, F., Chasin, M.: Programming hearing instruments to make live music more enjoyable. *Hear. J.* **63**, 30 (2010). <https://doi.org/10.1097/01.HJ.0000388538.27405.93>
- Hornsby, B.W.Y., Johnson, E.E., Picou, E.: Effects of degree and configuration of hearing loss on the contribution of high- and low-frequency speech information to bilateral speech understanding. *Ear Hear.* **32**, 543–555 (2011). <https://doi.org/10.1097/AUD.0b013e31820e5028>
- Jawaheer, G., Szomszor, M., Kostkova, P.: Comparison of implicit and explicit feedback from an online music recommendation service. In: Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems, HetRec 2010, Held at the 4th ACM Conference on Recommender Systems, RecSys 2010, pp. 47–51. (2010)
- Jawaheer, G., Weller, P., Kostkova, P.: Modeling user preferences in recommender systems: a classification framework for explicit and implicit user feedback. *ACM Trans. Interact. Intell. Syst.* (2014). <https://doi.org/10.1145/2512208>
- Jensen, N.S., Hau, O., Nielsen, J.B.B., et al.: Perceptual effects of adjusting hearing-aid gain by means of a machine-learning approach based on individual user preference. *Trends Hear.* **23**, 1–23 (2019). <https://doi.org/10.1177/2331216519847413>
- Johansen, B., Petersen, M., Korzepa, M., et al.: Personalizing the fitting of hearing aids by learning contextual preferences from internet of things data. *Computers* **7**, 1 (2017). <https://doi.org/10.3390/computers7010001>
- Keidser, G., Alamudi, K.: Real-life efficacy and reliability of training a hearing aid. *Ear Hear.* **34**, 619–629 (2013). <https://doi.org/10.1097/AUD.0b013e31828d269a>
- Killion, M.C.: New thinking on hearing in noise: a generalized articulation index. *Semin. Hear.* **23**, 057–076 (2002). <https://doi.org/10.1055/s-2002-24976>
- Kochkin, S., Beck, D.L., Christensen Laurel, A., et al.: MarkeTrak VIII: the impact of the hearing healthcare professional on hearing aid user success. *Hear. Rev.* **17**, 12–34 (2010)
- Korzepa, M.J., Johansen, B., Petersen, M.K., et al.: Learning preferences and soundscapes for augmented hearing. In: Proceedings of Intelligent User Interfaces. Association for Computing Machinery (2018)
- Levy, S.C., Freed, D.J., Nilsson, M., et al.: Extended high-frequency bandwidth improves reception of speech in spatially separated masking speech. *Ear Hear.* **36**, 214–224 (2015). <https://doi.org/10.1097/AUD.000000000000161>. Extended
- Long, M.: Fundamentals of acoustics. In: *Architectural Acoustics*, Second Edi. Elsevier, Boston, 39–79 (2014a). <https://doi.org/10.1016/B978-0-12-398258-2.00002-7>

- Long, M.: Human perception and reaction to sound. In: *Architectural Acoustics*, Second Edi. Elsevier, Boston, 81–127 (2014b). <https://doi.org/10.1016/B978-0-12-398258-2.00003-9>
- Marozeau, J., Florentine, M.: Loudness growth in individual listeners with hearing losses: a review. *J. Acoust. Soc. Am.* **122**, EL81–EL87 (2007). <https://doi.org/10.1121/1.2761924>
- Musiek, F.E., Shinn, J., Chermak, G.D., Bamiou, D.-E.: Perspectives on the pure-tone audiogram. *J. Am. Acad. Audiol.* **28**, 655–671 (2017). <https://doi.org/10.3766/jaaa.16061>
- Nelson, J., Kiessling, J., Dyrlund, O., Groth, J.: Integrating hearing instrument datalogging into the clinic. In: *American Academy of Audiology Convention*. Minneapolis. (2006)
- Nelson, P.B., Perry, T.T., Gregan, M., VanTasell, D.: Self-adjusted amplification parameters produce large between-subject variability and preserve speech intelligibility. *Trends Hear.* **22**, 1–13 (2018). <https://doi.org/10.1177/2331216518798264>
- Ng, E.H.N., Rudner, M., Lunner, T., et al.: Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.* **52**, 433–441 (2013). <https://doi.org/10.3109/14992027.2013.776181>
- Oetting, D., Hohmann, V., Appell, J.-E., et al.: Restoring perceived loudness for listeners with hearing loss. *Ear Hear.* **39**, 664–678 (2018). <https://doi.org/10.1097/AUD.0000000000000521>
- Pasta, A., Petersen, M.K., Jensen, K.J., Larsen, J.E.: Rethinking hearing aids as recommender systems. In: *The 4th Workshop on Health Recommender Systems Co-located with ACM RecSys'19*. Copenhagen, Denmark, pp. 11–17. (2019)
- Pasta, A., Szatmari, T.-I., Christensen, J.H., et al.: Clustering users based on hearing aid use: an exploratory analysis of real-world data. *Front. Digit. Health* **3**, 109 (2021). <https://doi.org/10.3389/fgdh.2021.725130>
- RDocumentation. Regression models for ordinal data via cumulative link (mixed) models: ordinal package. In: *RDocumentation*. <https://www.rdocumentation.org/packages/ordinal/versions/2019.12-10/topics/ordinal-package> (2019). Accessed 12 Apr 2021
- Schinkel-Bielefeld, N., Kunz, P., Zutz, A., Buder, B.: Evaluation of hearing aids in everyday life using ecological momentary assessment: what situations are we missing? *Am. J. Audiol.* **29**, 591–609 (2020). https://doi.org/10.1044/2020_AJA-19-00075
- Schum, D.J.: Noise-reduction circuitry in hearing aids. *Hear. J.* **56**, 32 (2003). <https://doi.org/10.1097/01.HJ.0000294049.75608.1c>
- Shi, L.F., Doherty, K.A., Kordas, T.M., Pellegrino, J.T.: Short-term and long-term hearing aid benefit and user satisfaction: a comparison between two fitting protocols. *J. Am. Acad. Audiol.* **18**, 482–495 (2007). <https://doi.org/10.3766/jaaa.18.6.3>
- Shiffman, S., Stone, A.A., Hufford, M.R.: Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* **4**, 1–32 (2008). <https://doi.org/10.1146/annurev.clinpsy.3.022806.091415>
- Verma, L., Sanju, H.K., Scaria, B., et al.: A comparative study on hearing aid benefits of digital hearing aid use (BTE) from six months to two years. *Int. Arch. Otorhinolaryngol.* **21**, 224–231 (2017). <https://doi.org/10.1055/s-0036-1592117>
- Walker, J.J., Cleveland, L.M., Davis, J.L., Seales, J.S.: Audiometry screening and interpretation. *Am. Fam. Physician* **87**, 41–47 (2013)
- Walravens, E., Keidser, G., Hickson, L.: Consistency of hearing aid setting preference in simulated real-world environments: implications for trainable hearing aids. *Trends Hear.* (2020). <https://doi.org/10.1177/2331216520933392>
- Wendt, D., Hietkamp, R.K., Lunner, T.: Impact of noise and noise reduction on processing effort. *Ear Hear.* **38**, 690–700 (2017). <https://doi.org/10.1097/aud.0000000000000454>
- Wong, L.L.N., Chen, Y., Wang, Q., Kuehnel, V.: Efficacy of a hearing aid noise reduction function. *Trends Hear.* **22**, 1–14 (2018). <https://doi.org/10.1177/2331216518782839>
- Wright, D., Gagné, J.-P.: Acclimatization to hearing aids by older adults. *Ear Hear.* **42**, 193–205 (2021). <https://doi.org/10.1097/AUD.0000000000000913>
- Yoon, S.H., Nam, K.W., Yook, S., et al.: A trainable hearing aid algorithm reflecting individual preferences for degree of noise-suppression, input sound level, and listening situation. *Clin. Exp. Otorhinolaryngol.* **10**, 56–65 (2017). <https://doi.org/10.21053/ceo.2015.01690>
- Ziesemer, K., Maria König, L., Jo Boushey, C., et al.: Occurrence of and reasons for “missing events” in mobile dietary assessments: results from three event-based ecological momentary assessment studies. *JMIR mHealth uHealth* (2020). <https://doi.org/10.2196/15430>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Alessandro Pasta is a Ph.D. student at the Technical University of Denmark (DTU) in the area of Computer Science and works as Senior Data Scientist at Demant A/S. He previously earned a M.Sc. degree in Economics and Management of Innovation and Technology from Bocconi University. His research aims to personalize hearing aid technologies by deepening the understanding of hearing aid users' preferences and behavior. His research leverages data collected in real-world environments through smartphone-connected hearing aids.

Michael Kai Petersen is an Industry Fellow at the Centre for Technology Entrepreneurship at the Technical University of Denmark (DTU) and a co-founder of augmentedhearing.io, focused on extending true wireless earbuds with AI to improve speech understanding in challenging acoustics. Specifically, he is working on enhancing voices by applying AI to not only remove ambient noise but also dynamically adapt amplification based on preferences learned across multiple users in real-life speech scenarios.

Kasper Juul Jensen is head of the Customer Insights Department at Demant A/S. His focus is to ensure data-driven decision-making based on state-of-the-art methodologies within the fields of behavioral and data sciences. In fact, in his work, he goes a long way to merge the fields as he believes that one always benefits greatly from the other. His research focuses on the possibilities of hearing aids which are more aware of context and user intent than what we see to today and what improvement to hearing care this might entail.

Niels Henrik Pontoppidan manages research on data-driven audiology at Eriksholm Research Centre, part of Oticon A/S. His research focuses on exploring the boundaries of personalized hearing technology and expanding on the benefits from current state-of-the-art technology. From 2016 to 2020, he coordinated the Horizon 2020 project EVOTION investigating public hearing health policies with big data. His research has also demonstrated how deep learning enhances segregation in cocktail party situations.

Jakob Eg Larsen, PhD is Associate Professor in Cognitive Systems at the Technical University of Denmark (DTU), at the Department of Applied Mathematics and Computer Science (DTU Compute) where he heads the Mobile Informatics and Personal Data Laboratory (MILAB). His research interests include human-computer interaction (HCI) and personal informatics. Specifically, he is interested in methods and instruments to support humans in getting value from personal data.

Jeppe Høy Christensen is a Research Engineer at Eriksholm Research Centre, part of Oticon A/S. Jeppe has a M.Sc. in Biomedical Engineering and a Ph.D. in Cognitive Psychology. His research vision is to personalize hearing care technology and help hearing aid users to remedy hearing fatigue. His methods include the use of unobtrusive data logging technology, wearable sensors, experience sampling, and statistical modeling to investigate how behavioral and physiological signals are related to real-world hearing difficulties and how this is modulated by different degrees of hearing loss.