SOUNDSCAPE PERSONALISATION AT WORK: DESIGNING AI-ENABLED SOUND TECHNOLOGIES FOR THE WORKPLACE

Thomas DEACON (t.deacon@surrey.ac.uk) [1], **Arshdeep SINGH**[1], **Gabriel BIBBÓ**[1], and **Mark D. PLUMBLEY**[1]

[1]*Center for Vision, Speech and Signal Processing*, **University of Surrey**, UK

ABSTRACT

Poor workplace soundscapes can negatively impact productivity and employee satisfaction. While current regulations and physical acoustic treatments are beneficial, the potential of AI sound systems to enhance worker wellbeing is not fully explored. This paper investigates the use of AI-enabled sound technologies in workplaces, aiming to boost wellbeing and productivity through a *soundscape approach* while addressing user concerns. To evaluate these systems, we used *scenario-based design* and focus groups with knowledge workers from open-plan offices and those working remotely. Participants were presented with initial design concepts for AI sound analysis and control systems. This paper outlines user requirements and recommendations gathered from these focus groups, with a specific emphasis on soundscape personalisation and the creation of relevant datasets.

1. INTRODUCTION

New Artificial Intelligence (AI) technologies offer promising opportunities to improve workplaces by monitoring and controlling environmental sound. Specifically we refer to the AI technology of machine listening that can automatically categorise everyday sounds [1]. Such technology can be used to monitor workplaces to develop new strategies for improving quality of life. Inspired by the *soundscape approach* - where environmental noise researchers and practioners consider the overall acoustic environment including its potential for positive and restorative effects on human health and wellbeing [2] - we research how innovations in machine listening technology could enable the exploration into how context affects people, bringing us closer to understanding how sounds feel, rather than just basing systems on objective measures like loudness. But AI technologies do not come without ethical risks and significant issues around adoption. If such obstacles can be navigated, new soundscape technologies can provide a means for organisations to support wellbeing and improved work environments.

The design area we are addressing in this research is the creation of AI sound sensing and soundscape control systems for the workplace (both offices and at home). We use a parallel design process of deep-learning research [3], co-design [4], and prototyping [5]. So far, we have gathered multiple concepts for the design of AI sound sensing interventions in the workplace. Next, we look to uncover potential adoption issues and further understand workers needs through concept evaluation.

This paper investigates the perception of AI-enabled soundscape technologies for enhancing workers' wellbeing and productivity using scenario-based design [6] and reflexive thematic analysis [7]. We used fictional technology scenarios with both positive and negative viewpoints on the same concepts in structured focus groups. This work borrows from previous design fiction methods used to understand future technologies [8–10]. Our participant recruitment targeted adult knowledge workers, including programmers, researchers, and accountants, across two work contexts: open-plan offices (OPOs) and work-from-home (WFH). The paper's findings present how workers perceive design concepts, highlighting needs and concerns about proposed soundscape systems in the two work contexts. Our study contributes novel findings to Sound and Music Computing (SMC) in the areas of applied audio AI and sonic experience design for work settings.

2. RELATED WORK

2.1 Knowledge Workers, Noise, and Wellbeing

Our work focuses on knowledge/information workers whose type of work often involves "reflective work" such as problem-solving, strategising, and creative development [11]. An issue for reflective work is that it can be influenced by the environment, particularly sound. For example, open plan office (OPO) spaces are often filled with distractions from other people and sounds [12]. Persistent issues with distraction can significantly impact productivity and overall workplace satisfaction [13, 14]. When working from home, sound can also become a distraction [15]. Many home workspaces also lack the common sound-masking capabilities that offices have, which makes workers more susceptible to noise from family, pets, and the outside environment. This can be especially challenging for those who require a quiet space to focus on reflective work, with a common coping strategy being to use headphones to create a sense of privacy [16,17]. Therefore, it is beneficial to create an environment that allows individuals in any environment to concentrate on their tasks without being disrupted by external noise, while also improving the soundscape in order to minimise physiological harms [18].

2.2 AI Soundscape Analysis and Control Systems

AI sound technologies could alleviate the issues of workplace sound to enhance the capacities of workers to engage in reflective work. For instance, evidence suggests that a moderate level of background noise enhances creativity compared to high-, low-, and no-noise conditions [19]. In order to effectively apply this discovery in the workplace, an AI audio service is required to handle several tasks: capturing and analysing sounds on-site, collecting and processing personalisation features, producing contextually relevant audio, and broadcasting the audio within the work environment. Below, we discuss current research related to these areas.

2.2.1 Computational Analysis of Soundscapes

Within computational analysis of soundscapes, Sound Event Detection (SED) and Acoustic Scene Classification (ASC) technologies allow a machine to sense sound in a way comparable to aspects of human hearing, where we perceive sources and objects instead of just signals and mathematical features [20]. Automatic SED is an audio machine learning approach to spot "what is happening in an audio signal and when it is happening" [1], while ASC is the process of labelling an audio recorded in a specific environment e.g. park, large room [21]. Combined, these systems allow contextual understanding of sound, while providing human readable outputs like scene and event classifications. When linked to other supporting systems and technical developments in AI and SMC, such classification technologies provide a basis to explore personalised soundscape systems.

2.2.2 Soundscape Personalisation

A key design aim in our project is to understand the value of individuated soundscapes at work. Individualisation can be attained through a combination of personalisation and customisation [22]. Personalisation tailors content, structure, and presentation to an individual automatically using *adaptivity* [22]. Customisation offers *adaptability* [22], allowing individuals to personally modify content presentation or structure. In this way, personalisation is system-initiated and system-driven, while customisation is user-initiated and user-driven process. An example of an adaptive system is when AI algorithms analyse biometric and behavioural data to create personalised soundscapes that promote focus or relaxation [23]. Already, companies like *Endel* have developed an AI product that generates personalised soundscapes using circadian rhythm and heart-rate to optimise for modes like focus and relaxation. [24]. For Deaf [1] and hard of hearing individuals sonic customisation refers to visualisations based on sound recognition tools built to suit the individual needs and preferences of users, as pre-trained sound recognition models may not meet their diverse requirements [25]. Central to our idea of soundscape personalisation is making systems that can use real-time data and sonic preferences, while understanding user contexts related to specific goals and settings.

2.2.3 Personalised Audio Systems for Work

In the area of workplace noise abatement, personal audio systems exist which can improve speech privacy in shared spaces by focusing speech towards the target listener and masking surrounding sounds, without causing annoyance to nearby individuals [26]. Also, a personalised audio masking signal can significantly enhance speech privacy in office spaces without significantly increasing annoyance when compared to white noise [27]. Previously, a real-time natural soundscape generation system was proposed for workspace voice-masking [28]. This system, based on current weather conditions, can improve speech privacy and reduce distractions, while maintaining a pleasing and informative environment. *Sound Bubbles* are a way of creating a personalised auditory environment by generation and modification of sound sequences around a worker, that are in line with the physical environment, enhancing their ability to concentrate [29–31].

Within audio engineering, extensive research has been conducted into *Sound Zones* as a way to create physically limited areas within a room where sounds can be controlled, allowing for multiple personal listening experiences within the same space without the need for headphones or disturbing others [32–35]. Since that foundational work, Sound Zones have been developed in domestic audio interaction design settings [36–41]. Sound Zones have recently been adapted to hospitals where the value of individualised sound environments in shared spaces is proposed and evaluated [42].

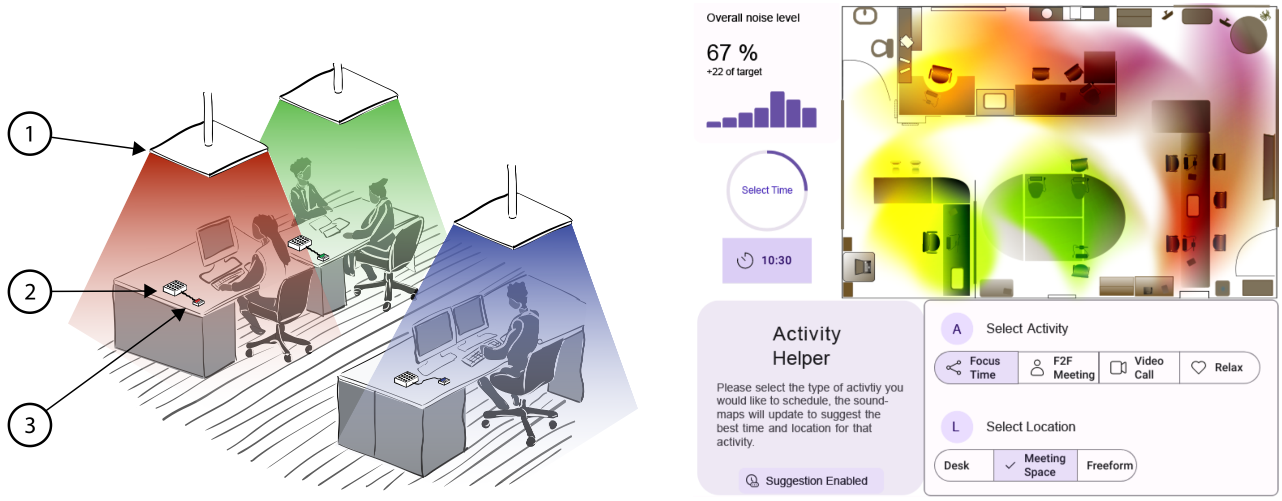
3. DESIGN CONCEPTS FOR WORKPLACE AI SOUND TECHNOLOGIES

Based on requirements developed in our previous co-design work [4], our workplace AI sound technology design concepts are based on technological possibility within the state-of the art in machine listening [43], sound engineering [40], workplace soundscape design [44], and workplace wellbeing AI [45]. The following design proposals include the technologies located in soundscape understanding, personalised sound control, and novel audio reproduction. These are imaginary conceptual prototypes: we did not build physical systems for this study. Conceptual prototypes provide a quick method of iterating a design space focused on human-centred design.

3.1 Sound Analysis Concepts

Sound Analyser: A sound analyser is combination of microphones and a sound labelling device, that feeds an AI model, that once trained can estimate whether sounds are pleasant and annoying in the workplace, as well as report sound levels. The purpose of Sound Analyser is to gather information about the stress and pleasure caused by

[1] We use Deaf with a capital D to refer to people who have been deaf all their lives, or since before they started to learn to talk. See `https://www.diversitystyleguide.com/glossary/deaf-deaf/`



(a) Concept prototype of AI sound system that creates personalised Sound Zones that integrate preferences. Element numbered are 1: Sound Zones via Highly Directional Speakers, 2: Sound Analyser Microphone Array, 3: Sound Analyser Preference Input.

(b) Concept prototype for AI Sound Map system that indicates suitable places to do certain types of work at different times of the day. Colours indicate suitability for a given task, green-yellow indicate good, red-purple bad.

Figure 1. Concept prototypes

sounds. It can provide data for decision making and help a variety of stakeholders understand types of sounds in a space. A simple rendering of a sound analyser box and preference input system can be found in Fig. 1a. A key user action of this system is to annotate sounds in some way: each scenario offers different means of doing this.

Sound Mapping of Spaces: Sound Maps are a visual way to represent noise levels and sound wellbeing impacts within certain contexts. For example, highlighting whether a given time and place are suitable for focused deep work given the user's preference profile. This idea is represented in Fig. 1b, which is a mock-up of a UI (the system is fictional). Preference profiles would include user specific recommendations based on computation of expected sound levels and timbre profile (e.g. a lot of speech being distracting). A further use of a Sound Map is to measure office space appropriateness and support improved sound etiquette in workplaces.

3.2 Sound Control Concepts

Adaptive Sound Masking and Artificial Soundscapes: Adaptive Sound Masking involves generating noise to "mask" other sounds, using personalised profiles and situational audio analysis. Its purpose is to improve productivity by reducing distractions in certain areas, protecting speech privacy, and generally improving office soundscapes. Artificial Soundscapes involve relaying or generating sounds from other locations, such as a beach or a rainforest. The purpose of Artificial Soundscapes is to improve productivity/wellbeing by creating a positive and calming ambience, signalling time of day, and reducing distraction from noise.

Personalised Audio via Highly Directional Sound (HDS): This concept involves the use of highly directional speakers, mounted on the ceiling or desk, to create Sound Zones such as those pictured in Fig. 1a. Sound Zones are designed to provide sound that can only be heard by one person in mixed-use spaces [42]. Primarily, they serve as reproduction technology for Adaptive Sound Masking and Artificial Soundscapes.

4. METHODOLOGY

4.1 Study Aim

The aim of our study was to gain worker insights on audio AI in open-plan office (OPO) and work-from-home (WFH) settings, to understand needs for AI-driven soundscape wellbeing and productivity systems.

4.2 Exploring Future Technologies with Stories

We use scenario-based design (SBD) to avoid fixating on solutions and ensure a better fit with the problem space. SBD engages knowledge workers in the design process [6]. While SBD has been applied to explore home autonomous systems [9], there is a lack of evaluation using SBD for AI audio technology in smart buildings, making our research a novel contribution to applied sound sensing AI. Our SBD evaluation involves stories, illustrations, and a structured feedback workshop. Inspired by the *contravision* approach [8], we present positive and negative stories to uncover concerns related to the design and acceptance of futuristic personal technologies. This approach is well-suited for understanding sensitive issues in pervasive system design [8]. Similar provocative design methods have recently been used for soundscape research in the SMC field [10].

In *contravision* scenarios, participants' preference for a specific rendering is not crucial. Instead, we focus on gathering information about broader concepts that influence engagement and understanding of technology [8]. *Contravision* scenarios provoke discussion and allow participants to reflect on their workplace expectations [9]. The ability to articulate the implications, provide comments on acceptability, and envision the consequences of such systems, is more important than a preference for a particular technology [46]. This facilitates discussions on values, fears, desires, and understanding, of technology; providing insights for design proposals, research challenges, use-cases, and deployment concerns.

4.3 Scenarios and Narrative Context

The scenarios are stories about AI sound systems in a spatial setting (OPO/WFH). Each scenario set (OPO/WFH) has a different protagonist who represents a typical knowledge worker. Their identity is established before the scenario presentations. The protagonist's preferences and environmental context are provided to help the reader engage with the story setting.

Both the OPO and WFH scenarios resembled each other in terms of core content. This is done by setting out the narrative "claims" that the stories reflect [6, 8]. Table 1 highlights the structural elements of the scenario sets. However, scenarios do not have to have the same ordering of these elements. The wording of the scenarios also varied based on work context (OPO/WFH). Scenario set A explores Sound Analysis concepts, while Scenario set B covers Sound Control concepts. All scenarios are available online [2] .

4.4 Participants Screening and Selection

Twenty-two knowledge worker participants were recruited using University mailing lists, social media, and word of mouth. Participants for focus groups have two groupings: OPO (N=10); WFH (N=12). The OPO group was an in-person event while the WFH took place remotely on MS Teams. General screening criteria included being an adult (over 18) and their job, or work experience, involves problem-solving, research, or creative ideation (to qualify as knowledge worker). The OPO group were open-plan office workers within 15 miles of University of Surrey campus. The WFH group were remote workers residing within 3 hours of the UK time zone. Participants were selected via purposive sampling based on job type, age breadth, and work context. Demographics of each group were: *OPO* - Sex: Male (N=3), Female (N=7); Age: 25-39 (N=5), 40-60 (N=4), 60+ (N=1). *WFH* - Sex: Male (N=5), Female (N=7); Age: 18-24 (N=3), 25-39 (N=7), 40-60 (N=2). All participants self-reported having functional hearing, ranging from poor to excellent. All participants provided informed consent and received a £40GBP voucher for their time. The study received favourable ethical approval, ID: FEPS 22-23 015 EGA Amend 1.

4.5 Focus Group Process

Focus groups lasted for two hours. First we introduced the concepts of AI sound technology and set up the story context. We then used a structured feedback process for each scenario set (A & B). This process included 2 phases of *Scenario Read and Reflect* each followed by a *Group Discussion*, then concluded with an *Open Discussion*. In the following sub-sections, we describe these phases in detail.

Scenario Read and Reflect: Participants were given scenarios to read individually in silence. After reading, participants reflected individually on the scenarios using the following questions:

1. What are your thoughts on working in an environment like this?
2. What are your thoughts on the believability of this scenario?
3. How would you use the system in this scenario?
4. Scenario Set Specific Question:
    - **Set A Only**: What do you think about the Sound Map in this scenario?
    - **Set B Only**: What are your thoughts on the sound systems in this scenario?

Group Discussions: These were conducted with smaller sub-groups of 3-5 people to encourage active participation and gather diverse information. Each sub-group scribe (a research team member) reported back to the whole group, and facilitated group discussion.

Open Discussions: Before the open discussion, key feedback from each sub-group is introduced by group scribes. After that, the open discussion was allowed to evolve in an open-ended manner supported by the facilitator.

4.6 Data Collection

Data collection for the focus groups included audio and video recordings, transcriptions, field-notes, and booklet responses. In-person groups used physical booklets for scenarios and question answers, and discussions were video recorded. Remote groups used an online form for scenarios and question answers, and discussions were recorded in MS Teams rooms. Booklets, notes, and memos were scanned and summarised in Word files. All the data was consolidated for analysis in MAXQDA24, a software program used for qualitative data analysis [47].

4.7 Analysis Methods

The data from booklets and discussions was analysed using reflexive thematic analysis (RTA) [7] to identify the reasoning and topics put forward by the participants. RTA is a method for identifying and analysing themes in qualitative data related to various research questions. RTA can investigate individual users' concerns while also addressing themes throughout the research process. Following best practice [7], we do not report numbered counts when reporting themes. A theme is considered important in relation to an analysis question, and it may only occur once to be significant. Instead, the relevance of the finding to the research focus is emphasised. For this paper, we report only themes related to the design concepts proposed in Sec. 3 and the design goal of soundscape personalisation introduced in Sec. 2. This choice is based

[2] https://adjuvant.github.io/SAW-FG-SMC24-Stories/

Table 1. Scenario Claims for Design Concepts

Scenario Element	*Positive Story Claims (Utopia)*	*Negative Story Claims (Dystopia)*
Set A Claims: Sound Analysis Concepts Scenario Elements. Sound Analyser (SA), Sound Map (SM).		
SA Devices	Handheld device	Desktop box, web app with notifications
SM Devices	Installed microphone array, room visualisation	Installed microphone array, room visualiser, LED alarms
SA Purpose	Build SM preference profile of what is pleasant or annoying	
SM Purpose	Informing sound etiquette policy	
Receiving the SA	Human Resources wellbeing initiative, everyone does tagging	HR wellbeing initiative, specific people given a tagging task
Using the SA	Easy, just point at something and label it (control, agency), engages you in moving around the office	Forced to figure out placement and system settings. Real-time tagging, system requires you to go through sounds and label them seemingly at random (human-in-the-loop).
Experience of tagging sounds	Strange at first but satisfying, new appreciation for sounds is mindful	Notice mostly annoying sounds, feels strange to label positive ones in the moment.
Impact of SM	Shows good and bad places/times for work activities. Improved office vibe	Feeling of being judged/judging others
Set B Claims: Sound Control Concepts Scenario Elements		
How the system is installed	Personalised device, engineer installation, software calibration, control app, desk-based	Workspace-wide array of overhead speakers
System description	Requires training via annotations, subscription service. System can relay sounds.	System installed without prior knowledge/consent, capturing multi-modal data seemingly without permission
Soundscape change	Gentle augmentation of soundscape. Benefits to focus.	Unforeseen adverse effects of the AI sound sensing technology, the system experiments on you to determine preferences (fully adaptive)
Reflection	Satisfaction and no reality dissonance	Discomfort, anger, social confrontation, assertion that technology can go too far
Novel elements	Doubt at the system initially, system has a set of moods that can be extended, social brainstorming presents a problem for different preferences	Chatbot service, profile config, tech support need to amend the system when consent status changed.

on the page limit and relevance to SMC community. Broader thematic analysis findings will be reported in future publications.

5. THEMATIC ANALYSIS: WORKPLACE SOUNDSCAPE PERSONALISATION

The following research questions guided the RTA process:

1. How do participants understand the purpose of current prototypes/design concepts?
2. What concerns do participants have about such technology in the OPO workplace or WFH?
3. How are AI personalisation and customisation technologies received?

In the following analysis, themes are denoted using sans-serif font, italics and bold case, and capitalised words e.g. ***Theme Title***, and we provide references to the sub-themes on which top-level themes are based. When a theme is related to a specific scenario set (A: Sound Annotator; B: Sound Controller) it is indicated in square brackets in the theme title e.g. [A] or [B]. For the presentation of any data, rather than use verbatim, we edited many extracts to improve readability, however we did not alter keyword choice or intention.

5.1 Open-plan Office (OPO)

Personalisation Issues Participants highlighted issues with personalisation in OPO settings. The theme involves concerns about the reliability of sound tagging, level of detail in personalisation, delivery of personalised sounds via networked speaker systems in shared spaces. We highlight some of the contributing sub-themes below.

Scepticism Of Personalisation Claims It was suggested that people have different preferences and perceptions of sounds which felt incongruous with the claims around personalisation: *"Different people perceive sounds differently"* (P2). Also, there was the belief that systems would not be granular enough to handle conflicting noise preferences in office spaces: *"in a communal office there would be too much contradicting noise based of numerous peoples preferences."* (P4). Scenarios around sound control were not accepted by some as possible as it was not understood *"how different people's soundscapes, played through speakers could co-exist in the same environment"* (P1). Overall, the benefits of systems felt unclear and many participants did not find proposals believable.

Limited Usefulness And Appeal Of AI Participants suggest that some AI sound systems, such as Sound Maps and personalised speaker systems, may not be worth using due to triviality of applications, lack of applicability in certain spaces like meeting rooms or shared public areas, and preferences for personal control through headphones. For example: *"I think it is useful if there is enough spaces; only if you actually have many zones with different purposes for them, because I feel like in most offices, every human can create a sound map in their head: open space is noisy, meeting rooms for private conversations and video calls"* (P1). Some participants propose simpler solutions, such as physical signage for indicating quiet areas, rather than relying solely on AI sound analysis and visualisations *"to tell people the expectations of certain areas"* (P4). Personalised HDS systems were considered too expensive for most workplaces. Also it was suggested that systems *"will work on headphone instead of speaker"* (P10), highlighting simpler solutions for public areas not using networks of speakers that also provide more control, *"I can use my own headphones to listen to sounds I like when I want, rather than being forced to hear something all the time"* (P6). Overall, participants express concerns about the limited usefulness and appeal of AI sound systems for controlling and analysing sounds in work situations, and propose simpler solutions as more effective alternatives.

5.2 Work-from-home (WFH)

Productivity Benefits This theme explores the potential benefits of AI audio systems in the home workplace. Participants express their belief in the adoption potential of these systems, highlighting positive impacts on personal productivity and the creation of a comfortable work environment. For example, P18 responds to Sound Control systems questions: *"In an environment where I need more of attention and focus I would consider it"*. Participants believed sound analysis reduces distractions and increases efficiency by optimising task timing and location, potentially enhancing comfort and productivity. Some participants believe sound mapping could optimize home office environments. For example, P17 stated their preference for *"staying in different environments based on which sounds are most conducive and productive"*.

In terms of Sound Control Systems, participants recognise the positive impact of a well-designed system on comfort and wellbeing. However, they express concerns about disruption and distraction caused by poorly designed systems. Reliability and accuracy are identified as essential features, ensuring consistent and effective sound control. Participants also emphasise the need to optimise sound environments for both productivity and overall wellbeing.

Desire For Personalisation And Customisation This theme highlights behaviours and needs around personalisation. Participants find personalised sound environments innovative and helpful, with one participant noting that *"it's quite conducive and relaxing since I have control of the sound around me"* (P16). The personalised aspect is appealing to participants, who think they may feel less stressed when using AI sound tools: *"I like that this environment is centred around the person. It feels like an ambient and relaxing environment"* (P17). Participants illustrate a desire for personalisation to create a good work environment to *"feel comfortable and welcomed"* (P13) as systems *"allow me to control all forms of noise"* (P13).

The *"seamless nature"* (P17) of personalised sound control is seen as an evolution of existing technologies, such as YouTube and lo-fi music, and allows individuals to have more control over their work environment by having *"sounds rise and fall at different times"* (P17). Participants appreciate the ability to experiment with different sounds and customise their auditory environment to meet their individual preferences and needs: *"The sounds could be adjusted to reflect my behaviour and work style"* (P22). Participants voiced a preference for systems that could be easily personalised to their individual needs, and where they could maintain a level of control over operations (customisation), with one participant noting that *"the mood enhancement is really the best thing here"* (P12).

Some worry about stress and dissatisfaction if the system doesn't cater to their specific preferences, such as blending preferences of other household members' likes and dislikes. Furthermore, participants express a desire for personalised sound settings, including saving and labelling preferred sounds, setting up individual preferences, analysing sounds around them, and adjusting background noises to specifications or moods. The ability to input and incorporate one's own sounds into the system is seen as appealing, potentially increasing self-awareness of sounds and personal behaviour. This theme highlights the role of individual user preference control in influencing acceptance and emotional responses to the system.

5.3 Shared Perspectives

Annotation Problematic Both groups found manual processes for labelling to be strange choices, for a variety of reasons. Part of this was finding the suggested methods for tagging sounds as annoying, while another aspect was that some did not think methods of annotation would be that useful and might create bad data. OPO participants question the accuracy of datasets and practicality proposed analysis devices to tag sounds. This was due to concerns about forgetting to use sound analysers/taggers and the subjectivity of data and potential biases. Some participants express scepticism that OPO devices and workers can accurately capture sound in a work environment due to the fleeting nature of sounds and the potential for individuals to focus on their tasks rather than tagging sound events.

Another concern was impact on neurodivergent individuals if datasets are too biased towards normative sensory experiences. Example contributions: *"For someone like me, I might forget about the sound analyser when I'm focused on work. Therefore, it doesn't provide an accurate sound analysis of my space. In a week, regardless of how many times I use it, I sometimes forget to take it with me."* (P11), *"I'm not sure anyone would enjoy tagging things, there is a novelty value at first but this would soon wear off."* (P9), and *"I would be concerned that people would purposely listen out for sounds to tag but whether these sounds would be distracting for their work is another matter."* (P9). Overall, participants express concerns about the practicality and effectiveness of manual sound event annotation to make data of sufficient quality.

Distraction & Annoyance This theme is closely linked with the previous ***Annotation Problematic*** theme but includes aspects of all scenarios. Unsurprisingly, most participants preferred utopian scenarios for their less intrusive nature, highlighting the distracting and unpleasant nature of systems in sound control dystopia scenarios. For both analysis and control technologies, there were concerns about the practicality of implementing such systems in communal workspaces and the potential for conflicting noise preferences, which was perceived as annoying. Disruption caused by notifications, flashing box/lights, and need to tag on demand were a key distractions. Related to the previous theme on annotation, it was suggested that tagging was a annoying task interruption: *"The sound analyser was quite disruptive to my workflow because I constantly had to hold the device to tag and label sounds instead of focusing on my work."* (P11). Overall, the participants felt that the systems could be helpful but the methods of customisation offered and the ways the systems behaved could be frustrating.

Spatial Challenges This theme addresses the obstacles of implementing AI audio in various workplace environments, including shared spaces, the usability of spatial policy systems, and personal space in WFH.

OPO Spatial Policy Usability This theme highlights the practicality of implementing sound maps in different working environments where the function is to inform new policy or adjust behaviour. Concerns are raised about the size and layout of a workspaces (e.g. *"it makes sense only if the space actually offers high variety of zones with different purposes"* (P1)), how space is associated teams and activities, and the system only really applies to workplaces that use hot-desking. Combined, the theme questions the benefits derived from integrating the systems, as it requires everyone to work together and may not leave room for ad-hoc activities. An implication of this is that systems could create cliques and build intolerance.

OPO Impact Of Sound Control On Shared Spaces For OPO, customising soundscapes in shared spaces introduces both challenges and potential conflicts. While some participants see the potential for personalised soundscapes, *"maybe it could work in some public space but not in a packed office"* (P10), it is suggested that conflicts could arise due to varying preferences and spatial setting, *"we have allocated work space and we are sitting right next to each other would blend preferences between users"* (P2). Also it is inferred that accommodating diverse preferences in shared spaces will be a challenge. Some suggest seeking input from aurally diverse groups *"how does it impact on hard of hearing person?"* (P5), and that sound control systems have potential for over-stimulation of neuro-divergent individuals. The generally applicability of systems is questioned, given the perceived difficulty of accommodating varying preferences *"can be difficult in public spaces"* (P7). Additionally, there is concern about creating more noise and annoyance in shared spaces.

WFH Living With AI The use of AI in a WFH setup raises concerns about privacy and work-life balance. While people are comfortable with personal assistants like Alexa, a work-only system may be a barrier to adoption. The potential pitfalls include collecting extraneous data from others and invading personal space. Some suggest restricting the system to only function in the work area and register only data from the worker and not other domestic residents. Overall, the system feels invasive and raises privacy concerns so simple configuration controls are needed for any such analysis or control systems.

6. IMPLICATIONS FOR DESIGN

Many participants see benefits in workplace AI audio systems, yet also identified challenges that need addressing in further research. OPO participants were quite sceptical of proposed systems in utopian and dystopian scenarios, highlighting the need to consider the social implications of AI audio systems. WFH participants showed a strong interest in personalising their auditory environment for mood enhancement, comfort, and the elimination of unwanted sounds. There were noted challenges in implementing systems in shared space, particularly due to conflicting preferences. This suggests a need for solutions that can manage shared spaces' complexities while still offering individual personalisation options.

User preferences and customisation were highlighted as important factors in AI audio system design. The preference for traditional sound management methods, like headphones and personal control over the auditory experience, indicates a potential direction for integrating personalisation into existing practices. Yet, it also emphasises the need to revisit the experience design of personalisation in the workplace more generally. Hence, we recommend that research on AI soundscape personalisation in workplace should focus on diverse data collection, individual control, and navigate the unique challenges of shared environments.

6.1 How can we iteratively train AI and not frustrate users?

The use of tagging and inputting preference data in scenario systems may be distracting and intrusive, leading users to believe that the proposed systems are too much work. Additionally, participants believe that they will use analysis and annotation functionalities in unreliable ways, which can skew the datasets used to train AI systems. This calls for the research community to develop methods for calibrating datasets for specific use cases and to provide convenient ways for users to input their own data and adjust preferences. When it comes to audio AI, we must address the issue that sound is a fleeting medium, and work technologies should not distract people from their core purpose while at work. Also as AI systems require large amounts of data, we must think about methods to scale data collection. We offer two opportunities below.

6.1.1 Gathering Detailed Information on Workers' Indoor Soundscape Needs

Datasets need to accommodate different people's sound preferences and aural diversity. Here are two suggestions:

Explore more methods of soundscape descriptor elicitation. There is a lack of research into soundscape descriptors for specific task contexts. Additionally, existing indices for indoor soundscapes may not be comprehensive enough for application in workplace settings.

Engage in participatory design for future systems. By setting specific goals and contexts, such as a WFH productivity system or a public space wellbeing soundscape generator, requirements can be determined through co-design and elicitation methods.

6.1.2 Use Prediction Models

Given the opinion of many participants was that in-the-moment tagging would be unreliable for behavioural reasons, an alternative design solution for is to leverage recent research on AI prediction of soundscape annoyance [48]. Here a baseline annoyance model could be tuned in context utilising continuous learning methods [49]. Currently the problem with this is that the annoyance prediction model is designed for outdoor soundscape research, not task-oriented indoor settings. Also a prediction model approach does not furnish us with ways of understanding work context to specify what is annoying when, but it may provide more immediate usability for a given set of general work modes e.g. a focus mode [50].

6.2 *Contraindications* for Design of Sound Control Systems in the Workplace

As highlighted in the ***Distraction & Annoyance*** theme, the claims of systems presented in dystopian sound control scenarios were rejected by most participants in both groups. This provides us with some clear counter-productive directions for future design, namely:

- Do not built expensive infrastructure e.g. networked HDS speaker arrays, when more behaviourally compatible and cost-effective solutions may suffice.
- Make sure wellbeing and productivity interventions are discussed with workers before deployment. We emphasise participatory design to bridge this gap.
- Think about what data is required to make a system function (i.e. is face tracking needed), and ask, can we use less?
- Think about the level of interactivity that is really needed for soundscape augmentation systems: do we really want fully automated, personalised systems constantly collecting data?
- Make sure designed systems have appropriate support for user understanding, control, and resolution of complaints/technical issues.

We hope these contraindications show the need for careful consideration when designing AI systems for peoples' lives, and also demonstrate that design fiction can offer insights into potential pitfalls that need attention.

7. CAVEATS AND LIMITATIONS

Our scenarios were designed to be provocative and elicit extreme value reflections from participants. While this form of human-centered design is still a valid exercise, it is important to balance it in the design space where we aim to improve people's work lives. To do this, follow-up exercises will provide more detail on the forms of desired support in work environments, with or without AI.

Our work only addressed OPO and WFH as separate environments. We did not consider how to frame hybrid work within the context of our scenarios. Additionally, we only conducted two English-language focus groups in a Global North, University-educated setting. As such, the positionality of our work should not be considered universal.

8. CONCLUSION

This paper presents an exploration of AI-enabled sound technologies in the workplace, focusing on how the field of SMC can enhance worker wellbeing and productivity through personalised soundscapes. Through focus groups involving open-plan office (OPO) and work-from-home (WFH) workers, the study uncovers insights into user requirements, challenges, and the potential of AI to create conducive or distracting work environments. Key findings indicate a strong desire for WFH participants for personalised sound environments to improve mood, comfort, and focus by mitigating unwanted noise. However, challenges such as the feasibility and cost of implementing highly directional sound systems in OPO settings, and the need for nuanced data collection and predictive models for effective personalisation, are highlighted. The paper promotes user control and customisation in sound control systems, considering the diverse needs across different work settings. By addressing the limitations and offering design recommendations, this research opens up new possibilities for enhancing workplace wellbeing and productivity through AI-enabled soundscapes.

Acknowledgments

This research was supported by Engineering and Physical Sciences Research Council (EPSRC) Grant EP/T019751/1 "AI for Sound".

9. REFERENCES

[1] A. Mesaros, T. Heittola, T. Virtanen, and M. D. Plumbley, "Sound Event Detection: A tutorial," *IEEE Signal Processing Magazine*, vol. 38, no. 5, pp. 67–83, Sep. 2021.

[2] Ö. Axelsson, M. E. Nilsson, and B. Berglund, "A principal components model of soundscape perception," *The Journal of the Acoustical Society of America*, vol. 128, no. 5, pp. 2836–2846, Nov. 2010.

[3] A. Singh and M. D. Plumbley, "A Passive Similarity based CNN Filter Pruning for Efficient Acoustic Scene Classification," in *Interspeech 2022*, 2022.

[4] E. Corrigan-Kavanagh, A. Fernandez, and M. Plumbley, "Envisioning Sound Sensing Technology for Enhancing Urban Living," in *Environments By Design*. Architecture Media Politics Society, Jan. 2022, p. 11.

[5] G. Bibbó, A. Singh, and M. D. Plumbley, "Recognise and Notify Sound Events using a Raspberry PI based Standalone Device [Demo]," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2023)*, New Paltz, NY, USA, Oct. 2023.

[6] M. B. Rosson and J. M. Carroll, "Scenario-Based Design," in *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, J. Jack and A. Sears, Eds. Lawrence Erlbaum Associates, 2002, pp. 1032–1050.

[7] V. Braun and V. Clarke, "Reflecting on reflexive thematic analysis," *Qualitative Research in Sport, Exercise and Health*, vol. 11, no. 4, pp. 589–597, Aug. 2019.

[8] C. Mancini, Y. Rogers, A. K. Bandara, T. Coe, L. Jedrzejczyk, A. N. Joinson, B. A. Price, K. Thomas, and B. Nuseibeh, "Contravision: Exploring users' reactions to futuristic technology," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10. New York, NY, USA: Association for Computing Machinery, Apr. 2010, pp. 153–162.

[9] T. Nilsson, A. Crabtree, J. Fischer, and B. Koleva, "Breaching the future: Understanding human challenges of autonomous systems for the home," *Personal and Ubiquitous Computing*, vol. 23, no. 2, pp. 287–307, Apr. 2019.

[10] R. Atienza, H. Lindetorp, and K. Falkenberg, "Playing the Design: Creating Soundscapes through Playful Interaction," in *20th Sound and Music Computing Conference (SMC 2023)*. Stockholm, Sweden: Sound and Music Computing Network, Jun. 2023.

[11] W. Reinhardt, B. Schmidt, P. Sloep, and H. Drachsler, "Knowledge Worker Roles and Actions—Results of Two Empirical Studies," *Knowledge and Process Management*, vol. 18, no. 3, pp. 150–174, 2011.

[12] A. Haapakangas, V. Hongisto, J. Hyönä, J. Kokko, and J. Keränen, "Effects of unattended speech on performance and subjective distraction: The role of acoustic design in open-plan offices," *Applied Acoustics*, vol. 86, pp. 1–16, 2014.

[13] H. Jahncke, S. Hygge, N. Halin, A. M. Green, and K. Dimberg, "Open-plan office noise: Cognitive performance and restoration," *Journal of Environmental Psychology*, vol. 31, no. 4, pp. 373–382, Dec. 2011.

[14] Y. Al Horr, M. Arif, A. Kaushik, A. Mazroei, M. Katafygiotou, and E. Elsarrag, "Occupant productivity and office indoor environment quality: A review of the literature," *Building and Environment*, vol. 105, pp. 369–389, Aug. 2016.

[15] G. E. Puglisi, S. Di Blasio, L. Shtrepi, and A. Astolfi, "Remote Working in the COVID-19 Pandemic: Results From a Questionnaire on the Perceived Noise Annoyance," *Frontiers in Built Environment*, vol. 7, p. 688484, 2021.

[16] S. Torresin, R. Albatici, F. Aletta, F. Babich, T. Oberman, A. E. Stawinoga, and J. Kang, "Indoor soundscapes at home during the COVID-19 lockdown in London – Part I: Associations between the perception of the acoustic environment, occupantś activity and well-being," *Applied Acoustics*, vol. 183, p. 108305, Dec. 2021.

[17] S. Torresin, E. Ratcliffe, F. Aletta, R. Albatici, F. Babich, T. Oberman, and J. Kang, "The actual and ideal indoor soundscape for work, relaxation, physical and sexual activity at home: A case study during the COVID-19 lockdown in London," *Frontiers in Psychology*, vol. 13, p. 1038303, Dec. 2022.

[18] T. Münzel, T. Gori, W. Babisch, and M. Basner, "Cardiovascular effects of environmental noise exposure," *European Heart Journal*, vol. 35, no. 13, pp. 829–836, Apr. 2014.

[19] R. Mehta, R. Zhu, and A. Cheema, "Is noise always bad? Exploring the effects of ambient noise on creative cognition," *Journal of Consumer Research*, vol. 39, no. 4, pp. 784–799, 2012.

[20] T. Virtanen, M. D. Plumbley, and D. Ellis, *Computational Analysis of Sound Scenes and Events*. Berlin, Germany: Springer International Publishing, 2018.

[21] B. Ding, T. Zhang, C. Wang, G. Liu, J. Liang, R. Hu, Y. Wu, and D. Guo, "Acoustic scene classification: A comprehensive survey," *Expert Systems with Applications*, vol. 238, p. 121902, Mar. 2024.

[22] H. Treiblmaier, M. Madlberger, N. Knotzer, and I. Pollach, "Evaluating personalization and customization from an ethical point of view: An

empirical study," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of The*, Jan. 2004.

[23] G. Loudon, D. Zampelis, and G. Deininger, "Using Real-time Biofeedback of Heart Rate Variability Measures to Track and Help Improve Levels of Attention and Relaxation," in *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. Singapore Singapore: ACM, Jun. 2017, pp. 348–355.

[24] T. Hayes, "The science behind Endel's AI-powered soundscapes," https://www.amazon.science/latest-news/the-science-behind-endels-ai-powered-soundscapes, Nov. 2020.

[25] S. M. Goodman, P. Liu, D. Jain, E. J. McDonnell, J. E. Froehlich, and L. Findlater, "Toward User-Driven Sound Recognizer Personalization with People Who Are d/Deaf or Hard of Hearing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 2, Jun. 2021.

[26] D. Wallace and J. Cheer, "Design and evaluation of personal audio systems based on speech privacy constraints," *The Journal of the Acoustical Society of America*, vol. 147, no. 4, pp. 2271–2282, Apr. 2020.

[27] A. Krasnov, E. R. Green, B. Engels, and B. Corden, "Enhanced speech privacy in office spaces," *Building Acoustics*, vol. 26, no. 1, pp. 57–66, Mar. 2019.

[28] C. Riggs and J. Braasch, "Real-time natural soundscape generation based on current weather conditions for workspace voice-masking," *The Journal of the Acoustical Society of America*, vol. 140, no. 4_Supplement, p. 3041, Oct. 2016.

[29] M. L. Eriksson and L. Pareto, "Sound Bubbles for Productive Office Work," in *Nordic Contributions in IS Research*, ser. Lecture Notes in Business Information Processing, U. Lundh Snis, Ed. Cham: Springer International Publishing, 2016, pp. 29–42.

[30] M. L. Eriksson, L. Pareto, R. Atienza, and K. F. Hansen, "My Sound Space: An attentional shield for immersive redirection," in *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion*, ser. AM '18. New York, NY, USA: Association for Computing Machinery, Sep. 2018.

[31] M. Ljungdahl Eriksson, L. Pareto, and R. Atienza, "The Sound Bubble : An Aesthetic Additive Design Approach to Actively Enhance Acoustic Office Environments," in *13th Conference on Sound and Music Computing*. Zentrum für Mikrotonale Musik und Multimediale Komposition (ZM4), Hochschule für Musik und Theater, 2019, pp. 253–260.

[32] J. Francombe, P. Coleman, M. Olik, K. Baykaner, P. J. B. Jackson, R. Mason, M. Dewhirst, S. Bech, and J. A. Pederson, "Perceptually Optimized Loudspeaker Selection for the Creation of Personal Sound Zones," in *52nd International Conference: Sound Field Control - Engineering and Perception*. Audio Engineering Society, Sep. 2013.

[33] P. Coleman, P. Jackson, M. Olik, and J. A. Pedersen, "Optimizing the Planarity of Sound Zones," in *52nd International Conference: Sound Field Control - Engineering and Perception*. Audio Engineering Society, Sep. 2013.

[34] M. Olik, J. Francombe, P. Coleman, P. J. B. Jackson, M. Olsen, M. Møller, R. Mason, and S. Bech, "A Comparative Performance Study of Sound Zoning Methods in a Reflective Environment," in *52nd International Conference: Sound Field Control - Engineering and Perception*. Audio Engineering Society, Sep. 2013.

[35] P. Coleman and P. Jackson, "Planarity-Based Sound Field Optimization for Multi-Listener Spatial Audio," in *2016 AES International Conference on Sound Field Control*. Audio Engineering Society, Jul. 2016.

[36] S. S. Johansen and P. A. Nielsen, "Personalised Soundscapes in Homes," in *Proceedings of the 2019 on Designing Interactive Systems Conference*, ser. DIS '19. New York, NY, USA: Association for Computing Machinery, Jun. 2019, pp. 813–822.

[37] S. S. Johansen, P. A. Nielsen, and J. Kjeldskov, "Interaction Design for Domestic Sound Zones," in *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound*, ser. AM'19. New York, NY, USA: Association for Computing Machinery, Sep. 2019, pp. 248–251.

[38] S. S. Lundgaard, P. A. Nielsen, and J. Kjeldskov, "Designing for domestic sound zone interaction," *Personal and Ubiquitous Computing*, vol. 26, no. 4, pp. 1225–1236, Mar. 2020.

[39] S. S. Johansen, P. A. Nielsen, K. Stec, and J. Kjeldskov, "Experiences of Personal Sound Technologies," in *Human-Computer Interaction – INTERACT 2021*, ser. Lecture Notes in Computer Science, C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, and K. Inkpen, Eds. Cham: Springer International Publishing, 2021, pp. 523–541.

[40] R. M. Jacobsen, K. Fangel Skov, S. S. Johansen, M. B. Skov, and J. Kjeldskov, "Living with Sound Zones: A Long-term Field Study of Dynamic Sound Zones in a Domestic Context," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Hamburg Germany: ACM, Apr. 2023, pp. 1–14.

[41] R. M. Jacobsen, K. F. Skov, M. B. Skov, and J. Kjeldskov, "Barriers for Domestic Sound Zone Systems: Insights from a Four-Week Field Study," in *Proceedings of the 18th International Audio Mostly Conference*. Edinburgh United Kingdom: ACM, Aug. 2023, pp. 185–192.

[42] K. Fangel Skov, P. Axel Nielsen, and J. Kjeldskov, "Tuning Shared Hospital Spaces: Sound Zones in Healthcare," in *Proceedings of the 18th International Audio Mostly Conference*. Edinburgh United Kingdom: ACM, Aug. 2023, pp. 63–70.

[43] Q. Kong, Y. Cao, T. Iqbal, Y. Wang, W. Wang, and M. D. Plumbley, "PANNs: Large-Scale Pretrained Audio Neural Networks for Audio Pattern Recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2880–2894, 2020.

[44] S. Felipe Contin de Oliveira, F. Aletta, and J. Kang, "Self-rated health implications of noise for open-plan office workers: An overview of the literature," *Building Acoustics*, p. 1351010X231152841, Feb. 2023.

[45] P. Mantello and M.-T. Ho, "Emotional AI and the future of wellbeing in the post-pandemic workplace," *AI & SOCIETY*, Feb. 2023.

[46] A. Berger, A. H. Ambe, A. Soro, D. De Roeck, and M. Brereton, "The Stories People Tell About The Home Through IoT Toolkits," in *Proceedings of the 2019 on Designing Interactive Systems Conference*, ser. DIS '19. New York, NY, USA: Association for Computing Machinery, Jun. 2019, pp. 7–19.

[47] "MAXQDA 2024," VERBI Software, Berlin, Germany, Nov. 2023.

[48] Y. Hou, Q. Ren, H. Zhang, A. Mitchell, F. Aletta, J. Kang, and D. Botteldooren, "AI-based soundscape analysis: Jointly identifying sound sources and predicting annoyance," *The Journal of the Acoustical Society of America*, vol. 154, no. 5, pp. 3145–3157, Nov. 2023.

[49] L. Jia, Z. Zhou, F. Xu, and H. Jin, "Cost-Efficient Continuous Edge Learning for Artificial Intelligence of Things," *IEEE Internet of Things Journal*, vol. 9, no. 10, pp. 7325–7337, May 2022.

[50] K. Saha and S. T. Iqbal, "Focus Time: Effectiveness of Computer Assisted Protected Time for Wellbeing and Work Engagement of Information Workers," in *Proceedings of the 2nd Annual Meeting of the Symposium on Human-Computer Interaction for Work*, ser. CHIWORK '23. New York, NY, USA: Association for Computing Machinery, Sep. 2023.