frontiers | Research Topics

New advances and novel applications of music technologies for health, well-being, and inclusion

Edited by
Emma Margareta Frid, Kjetil Falkenberg and Kat Agres

Coordinated by
Alex Michael Lucas

Published in
Frontiers in Computer Science
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-4415-0
DOI 10.3389/978-2-8325-4415-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

New advances and novel applications of music technologies for health, well-being, and inclusion

Topic editors
Emma Margareta Frid — Media Technology and Interaction Design, Royal Institute of Technology, Sweden
Kjetil Falkenberg — Royal Institute of Technology, Sweden
Kat Agres — National University of Singapore, Singapore

Topic Coordinator
Alex Michael Lucas — Queen's University Belfast, United Kingdom

Citation
Frid, E. M., Falkenberg, K., Agres, K., Lucas, A. M., eds. (2024). *New advances and novel applications of music technologies for health, well-being, and inclusion*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4415-0

Table of contents

frontiers | Frontiers in Computer Science

TYPE Editorial
PUBLISHED 19 January 2024
DOI 10.3389/fcomp.2024.1358454

OPEN ACCESS

EDITED AND REVIEWED BY
Roberto Therón,
University of Salamanca, Spain

*CORRESPONDENCE
Emma Frid
emmafrid@kth.se

RECEIVED 19 December 2023
ACCEPTED 08 January 2024
PUBLISHED 19 January 2024

CITATION
Frid E, Falkenberg K, Agres K and Lucas A (2024) Editorial: New advances and novel applications of music technologies for health, well-being, and inclusion.
Front. Comput. Sci. 6:1358454.
doi: 10.3389/fcomp.2024.1358454

Editorial: New advances and novel applications of music technologies for health, well-being, and inclusion

Emma Frid[1,2]*, Kjetil Falkenberg[1], Kat Agres[3] and Alex Lucas[4]

[1]Media Technology and Interaction Design, KTH Royal Institute of Technology, Stockholm, Sweden, [2]UMR9912 Sciences et Technologies de la Musique et du Son (STMS), Institut de Recherche et Coordination Acoustique Musique (IRCAM), Paris, France, [3]Centre for Music and Health, Yong Siew Toh Conservatory of Music, National University of Singapore, Singapore, Singapore, [4]Sonic Arts Research Centre, Queen's University Belfast, Belfast, United Kingdom

KEYWORDS

Digital Musical Instruments, music therapy, accessibility, multimodal interaction, music technology, sound and music, healthcare, wellbeing

Editorial on the Research Topic
New advances and novel applications of music technologies for health, well-being, and inclusion

The field of research dedicated to the design, creation, use, and evaluation of new sound and music technologies supporting health, wellbeing, and inclusion is rapidly expanding. Numerous research efforts are taking place at the intersection of areas such as universal design, accessibility, music therapy, music technology, sonic interaction design, and human-computer interaction (HCI). This Research Topic explores such intersections in music technology research aimed at promoting health, wellbeing, and inclusion, investigating how new methods, technologies, interfaces, and applications can enable everyone to enjoy the benefits of sound and music.

Previous Special Issues exploring similar topics include, for example, work by Magee (2011), Rickard (2014), and Schroeder and Michon (2021). A summary of the state of the art in music technology applied in various health scenarios was presented by Agres et al. (2021). Falkenberg and Frid (2021) also published an overview of methodological considerations for designing and reporting research on sound design and music for health. In addition, an online network called *Musical Care International Network*[1] was recently launched to bring people together to discuss and advocate for *musical care*, defined as the role of music listening and music-making in supporting any aspect of people's developmental or health needs (Spiro et al., 2023).

The papers published here reflect the diversity in practices and methods in this interdisciplinary research field. Common themes that reoccur are interfaces focused on accessibility of musical expression—in particular, the design and development of Accessible Digital Musical Instruments (ADMIs; see Frid, 2019)—and sound and music listening applications aimed at improving health and wellbeing. The papers demonstrate a richness in the methodological approaches employed, from third-wave HCI practices (Bødker, 2006) and action research (Reason and Bradbury, 2001) with emphasis on understanding the importance of the sociocultural context of musical interactions (see Waters, 2021) and the lived experience of the persons interacting with ecological artifacts (Bødker, 2006), to quantitative analysis of more traditional listening experiments.

1 https://musicalcareresearch.com/musical-care-network/

The paper by Ward presents the *Modular Accessible Musical Instrument Technology Toolkit (MAMI Tech Toolkit)*, developed over 5 years. It describes each of the tools in the toolkit, the functionality they offer, as well as the accessibility issues they address. In the paper by Lindetorp et al., the authors discuss a novel system consisting of commercially available accessible instruments from *Funki*[2] which enables students with Profound and Multiple Learning Disabilities (PMLD) to play music together with their assistants and a professional musician. Findings highlight how a system of networked and synchronized ADMIs could be conceptualized to include assistants more actively in collaborative music-making. Moreover, design considerations that support the assistants' roles as facilitators in such a context are discussed. In the work presented by Duarte et al., the authors explore the impact of different disability models in the process of designing inclusive music technology, with a case study focused on categorizing eleven ADMIs targeting d/Deaf people based on the medical, social, and cultural[3] disability models. The authors identify a lack of participatory approaches and an overall tendency toward the medical model of disability for the surveyed ADMIs. Finally, McMillan and Morreale describe an autoethnographic study of the lived experience of the first author, a disabled musician, proposing to use conceptual metaphors (Waters, 2021) and cultural probes (Gaver et al., 1999) as tools to account for the subjective experience, ecology, and specificity that contribute to a successful musician-instrument relationship.

The Research Topic also includes two papers in review format (Paisa et al., Lenzi et al.). Paisa et al. reviewed systems for tactile augmentations that provide inclusive musical experiences for persons who are deaf or hard of hearing. Results indicate that the research field is in an early phase, characterized by an exploratory approach and preliminary results, with most of the studies conducted in laboratory settings with small sample sizes, and sometimes low validity, for example due to evaluations not being performed with the intended user group (e.g., system designs intended for persons with cochlear implants evaluated with persons without hearing loss). These findings resonate with what was concluded regarding studies based on the medical model of disability, discussed in the paper by Duarte et al. The second review paper, published by Lenzi et al., presented a semi-systematic review of scientific literature on sound assessment studies in Neonatal Intensive Care Units. The authors emphasize the need for a sound quality assessment solution for indoor environments.

In the paper by Krause and Fletcher, the authors explore—through semi-structured interviews with radio personnel and focus groups with older adults—how radio personnel and listeners regard the purpose of radio, and how engaging with radio is perceived to influence wellbeing. Findings indicate that there are implicit and explicit ways in which radio facilitates wellbeing. The authors conclude that these findings may have implications for both broadcasting practices and future work on how radio might be used as an accessible tool for promoting quality of life in seniors.

The paper by Ramadas et al. also focused on a user group of older adults. The authors describe a pilot experiment using Carnatic Classical music as an intervention for improving auditory processing and cognitive skills. The intervention protocol resulted in improvements in auditory temporal processing skills and better speech perception in noise.

The paper by Agres et al. introduces *AffectMachine-Classical*, a music generation system capable of generating affective Classical music in real-time. *AffectMachine-Classical* was designed to be embedded into interactive interfaces or biofeedback systems (such as brain-computer-interfaces) to help users become aware of, and ultimately mediate, their own dynamic affective states, in support of real-time emotion self-regulation. Findings from a listening experiment indicated that the system is effective in communicating various levels of arousal via music, and is also quite convincing in terms of valence.

The editors are excited to share this Research Topic with the wider research community, and in doing so contribute to the growing body of knowledge surrounding music technology for health, wellbeing, and inclusion.

Author contributions

EF: Writing – original draft, Writing – review & editing. KF: Writing – review & editing. KA: Writing – review & editing. AL: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was partly funded by the Swedish Research Council (Vetenskapsrådet), grant number: 2020-00343, and the RIE2020 Advanced Manufacturing and Engineering (AME) Programmatic Fund (grant number: A20G8b0102), Singapore.

Acknowledgments

The guest editors would like to thank all authors and reviewers who have contributed to this Research Topic.

Conflict of interest

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

Publisher's note

2 https://www.funki.se/

3 Also referred to as the *social identity* or *cultural affiliation* model.

References

Agres, K. R., Schaefer, R. S., Volk, A., van Hooren, S., Holzapfel, A., Dalla Bella, S., et al. (2021). Music, computing, and health: A roadmap for the current and future roles of music technology for health care and well-being. *Music Sci.* 4:2059204321997709. doi: 10.1177/2059204321997709

Bødker, S. (2006). "When second wave HCI meets third wave challenges," in *Proceedings of the Nordic Conference on Human-Computer Interaction: Changing Roles, NordiCHI '06* (New York, NY: Association for Computing Machinery), 1–8.

Falkenberg, K., and Frid, E. (2021). "Designing and reporting research on sound design and music for health: Methods and frameworks for impact," in *Doing Research in Sound Design* (Waltham, MA: Focal Press), 125–150.

Frid, E. (2019). Accessible Digital Musical Instruments–A review of musical interfaces in inclusive music practice. *Multi. Technol. Interact.* 3, 57. doi: 10.3390/mti3030057

Gaver, B., Dunne, T., and Pacenti, E. (1999). Design: Cultural probes. *Interactions* 6, 21–29.

Magee, W. L. (2011). Music technology for health and well-being: The bridge between the arts and science. *Music Med.* 3, 131–133. doi: 10.1177/1943862111411719

Reason, P., and Bradbury, H. (2001). *Handbook of Action Research: Participative Inquiry and Practice*. Newcastle upon Tyne: Sage.

Rickard, N. S. (2014). Editorial: "Music and well-being" Special issue of psychology of well-being. *Psychol Well-Being* 4, 1–3. doi: 10.1186/s13612-014-0026-3

Schroeder, F., and Michon, R. (2021). Editorial: Accessibility of musical expression: Co-edited journal with selected papers of NIME2020. A special issue commissioned by MIT. *Comput. Music J.* 44, 180.

Spiro, N., Sanfilippo, K. R. M., McConnell, B. B., Pike-Rowney, G., Bonini Baraldi, F., Brabec, B., et al. (2023). Perspectives on musical care throughout the life course: Introducing the musical care international network. *Music Sci.* 6. doi: 10.1177/20592043231200553

Waters, S. (2021). The entanglements which make instruments musical: Rediscovering sociality. *J. N. Music Res.* 50, 133–146. doi: 10.1080/09298215.2021.1899247

frontiers | Frontiers in Computer Science

TYPE Systematic Review
PUBLISHED 27 April 2023
DOI 10.3389/fcomp.2023.1085539

Check for updates

Tactile displays for auditory augmentation–A scoping review and reflections on music applications for hearing impaired users

Razvan Paisa*, Niels Christian Nilsson and Stefania Serafin

Multisensory Experience Lab, Department for Architecture, Design and Media Technology, Aalborg University, Copenhagen, Denmark

OPEN ACCESS

EDITED BY
Kjetil Falkenberg,
Royal Institute of Technology, Sweden

REVIEWED BY
Rumi Hiraga,
National University Corporation Tsukuba University of Technology, Japan
Frank A. Russo,
Toronto Metropolitan University, Canada

*CORRESPONDENCE
Razvan Paisa
✉ rpa@create.aau.dk

SPECIALTY SECTION
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Computer Science

RECEIVED 01 November 2022
ACCEPTED 23 March 2023
PUBLISHED 27 April 2023

CITATION
Paisa R, Nilsson NC and Serafin S (2023) Tactile displays for auditory augmentation–A scoping review and reflections on music applications for hearing impaired users. *Front. Comput. Sci.* 5:1085539.
doi: 10.3389/fcomp.2023.1085539

The field of tactile augmentation has progressed greatly over the past 27 years and currently constitutes an emerging area of research, bridging topics ranging from neuroscience to robotics. One particular area of interest is studying the usage of tactile augmentation to provide inclusive musical experiences for deaf or hard-of-hearing individuals. This article details a scoping review that investigates and organizes tactile displays used for the augmentation of music from the field of hearing assistive devices, documented in 63 scientific publications. The focus is on the hardware, software, mapping, and evaluation of these displays, to identify established methods and techniques, as well as potential gaps in the literature. To achieve this purpose, a catalog of devices was created from the available literature indexed in the Scopus® database. We set up a list of 12 descriptors belonging to physical, auditory, perceptual, purpose and evaluation domains; each tactile display identified was categorized based on those. The frequency of use among these descriptors was analyzed and as well as the eventual relationship between them. Results indicate that the field is relatively new, with 80% of the literature indexed being published after 2009. Moreover, most of the research is conducted in laboratories, with limited industry reach. Most of the studies have low reliability due to small sample sizes, and sometimes low validity due to limited access to the targeted population (e.g., evaluating systems designed for cochlear implant users, on normal hearing individuals). When it comes to the tactile displays, the results show that the hand area is targeted by the majority of the systems, probably due to the higher sensitivity afforded by it, and that there are only a couple of popular mapping systems used by the majority of researchers. Additional aspects of the displays were investigated, including the historical distribution of various characteristics (e.g., number of actuators, or actuators type) as well as the sonic material used as input. Finally, a discussion of the current state of the tactile augmentation of music is presented, as well as suggestions for potential future research.

KEYWORDS

vibrotactile displays, vibrotactile music, hearing assistive devices, cochlear implant music, multisensory augmentation

1. Background

This section presents the rationale and objectives that form the basis for the scoping review detailed in the current paper. Subsequently, the field of vibrotactile augmentation is introduced along with a collection of relevant definitions.

1.1. Objectives

This article presents a scoping review of vibrotactile displays in the field of hearing assistive devices documented in 63 scientific publications. The main goal is to present a window into the relatively new research field of *Tactile Music Augmentation*—a research area that has particular applications for deaf or hard-of-hearing (D/HOH) individuals (Sorgini et al., 2018). The articles included are described in terms of hardware, software, mapping, and evaluation of the displays, to provide pointers toward common methods and techniques, as well as potential shortcomings and opportunities. While the primary motivation for the current work is to aid hearing impaired individuals, the majority of the technology analyzed is designed for average users. However, there is no evidence of a difference between the haptic sense of populations with hearing loss and those with normal hearing, as the most significant differences (if any) are typically thought to be perceptual, as demonstrated by Rouger et al. (2007). Thus, we believe that the devices intended for the general public have the potential to help D/HOH individuals just as well, and we have included them in our analysis. The *Scopus*®[1] database was queried and the eligible articles were dissected to create Table 1.

To provide an overview of the research field of vibrotactile augmentation of sound, we sought to meet the following research objectives:

- Identify the state-of-the-art and understand how research efforts have changed over the years.
- Understand the most successful and promising strategies for augmenting music with tactile stimulation
- Identify gaps in current research
- Provide a starting point with a strong foundation for designers, researchers, and practitioners in the field of vibrotactile augmentation.

By reaching these goals, this review will address the following research questions: "What are the most successful applications of vibrotactile augmentation?," "What is the historical distribution?" and "What are the most popular actuators, processing techniques, and mappings?"

More specifically, this review will summarize research findings published over 27 years, in order to learn which type of actuators, body areas, mappings, processing techniques, and evaluation practices are most common, and how these factors have evolved over the years. In addition to this, such an examination could identify potential relations between different system components (e.g., type of actuators and type of signal processing used). This would not imply that such correlations are the most successful, but it will suggest starting points for new vibrotactile augmentation applications.

This article accompanies previous reviews in the field of vibrotactile augmentation, and should be seen as complementary. The review covering a similar sample of literature was documented by Remache-Vinueza et al. (2021), who examined the methods and technologies used in the tactile rendering of music. Their work focuses on music through the touch modality in the general sense and encompasses literature covering use cases extending past vibrotactile augmentation of music, and thus the authors analyze works from a different perspective than this article. Other similar work includes a review of haptic wearables (Shull and Damian, 2015), a review of wearable haptic systems for the hands (Pacchierotti et al., 2017), as well as a review of clinically validated devices that use haptic stimulation to enhance auditory perception in listeners with hearing impairment (Fletcher, 2021). In addition to these publications, in Papetti and Saitis (2018), the field of *Musical Haptics* is discussed extensively. Any researcher or designer interested in the field of vibrotactile augmentation is strongly encouraged to study these publications as well.

1.2. Vibrotactile music augmentation–A short overview

Research on cutaneous augmentation has been conducted since the beginning of the twentieth century, focusing on thresholds of sensitivity by using tuning forks (Rydel and Seiffer, 1903). This primitive approach was abandoned in the follow-up work in favor of electronic transducers, and by 1935, it was known that the peak skin sensitivity is somewhere in the range of 200–250 Hz (Setzepfand, 1935). In 1954, A. Wilska mapped the threshold for 35 areas spread across the entire body (Wilska, 1954).

Music is complex and has been written and performed with respect to the hearing capabilities of humans. While vibrotactile augmentation can manipulate percepts, it is important to outline the musical dimensions that can be perceived through tactile stimulation alone, to better understand and evaluate the effect of multisensory integration. Rhythm—the temporal relationship of events in a musical context—is arguably the first aspect to be discussed, as it is fairly well transmitted through tactile channels (Jiam and Limb, 2019). Substantial research has been dedicated to understanding the vibrotactile rhythm, investigating its impact on music aesthetics (Swerdfeger et al., 2009; Hove et al., 2019), and in regards to D/HOH individuals, its enhancement properties (Gilmore and Russo, 2020; Aker et al., 2022) as well as its interaction with the auditory counterpart (Lauzon et al., 2020). When it comes to pitch—the perceived (vibrotactile) frequency—the bandwidth is very limited compared to the ear, but this does not mean that humans cannot perceive pitch differences, only that the *just noticeable difference* between intervals must be larger, and the range is bound within 20 Hz–1,000 Hz (Chafe, 1993). However, this has not stopped researchers from exploring the potential of vibrotactile pitch with respect to music, either as in isolation (Morley and Rowe, 1990), or more commonly in context through *Pitch Ranking* or *Melodic Contour Identification* tasks

1 www.scopus.com

10.3389/fcomp.2023.1085539

TABLE 1 Description of the tactile displays analyzed in this study.

Device name	Purpose	Listening situation	Nr. Act.	Actuators type	Signals used	DSP	Mapping scheme	Body area actuated	Wear-able	Evaluation measure	Evaluation population	Nr. Part.
Pump-and-Vibe (Haynes et al., 2021)	Music Enhancement	Lab Study	8	Mixed	Sawtooth wave	F0 extraction, Beat extraction	Complex Mapping, Pitch to Position	Arm	Yes	Experience	Normal Hearing	20
Tactile Phonemic Sleeve (TAPS) (Reed et al., 2021)	Speech Enhancement	Lab Study	24	Contact speakers	Sine wave	N/A	Complex Mapping	Arm	Yes	Discrimination	Normal Hearing	7
N/A (Pavlidou and Lo, 2021)	HAD, Speech Enhancement	Lab Study	3	ERM	Sine wave	Phoneme extraction	Complex Mapping	Hands	Yes	N/A	N/A	N/A
mosaicOne_C (Fletcher, 2021)	HAD, Music Enhancement	General	4	ERM	Sine wave	N/A	N/A	Wrist	Yes	N/A	N/A	N/A
Syntacts (Pezent et al., 2021)	SW/HW platform	General	N/A	Unspecified	Various Waveforms	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A (DeGuglielmo et al., 2021)	Music Enhancement	Lab Study	4	LRA	Sine wave	Spectrum isolation, Envelope extraction	Pitch to Position	Wrist	Yes	Discrimination	Normal Hearing	3
N/A (Turchet et al., 2020)	Music Enhancement	Artistic	4	ERM	Sine wave	N/A	Complex Mapping	Arm	Yes	Experience	Normal Hearing	20
N/A (Fletcher and Zgheib, 2020)	HAD, Sound Localization	Lab Study	2	EVT	Sine wave	Multiband compression, ILD Enhancement, Envelope extraction	Auditory Frequency to Tactile Frequency	Wrist	No	Discrimination	Normal Hearing	32
mosaicOne_B (Fletcher et al., 2020)	HAD, Pitch Discrimination	Lab Study	12	ERM	Various Waveforms	F0 extraction	Pitch to Position	Arm	Yes	Music Listening Performance	Normal Hearing	12
Tactile Tone (Shin et al., 2020)	HAD, Music Training	Singing along music	9	ERM	Sine wave	F0 extraction	Complex Mapping	Hands	Yes	Music Listening Performance	CI Users	2
N/A (Sharp et al., 2020)	HAD, Music Enhancement	Lab Study	6	Voice coil	Music	N/A	N/A	Hands	Yes	Discrimination	Normal Hearing, HA Users, CI Users, D/HOH	10, 2, 6, 2

(Continued)

TABLE 1 (Continued)

Device name	Purpose	Listening situation	Nr. Act.	Actuators type	Signals used	DSP	Mapping scheme	Body area actuated	Wear-able	Evaluation measure	Evaluation population	Nr. Part.
EarVR (Mirzaei et al., 2020)	HAD, Sound Localization	VR, Lab Study	2	ERM	Square wave	Amplitude Thresholding	Auditory Amplitude to Tactile Amplitude	Ears	Yes	Discrimination	Normal Hearing, D/HOH	20, 20
Tasbi (Pezent et al., 2020)	AR/VR Human-Computer Interaction	N/A	6	LRA	Sine wave	N/A	N/A	Wrist	Yes	N/A	N/A	N/A
N/A (Luciá et al., 2020)	HAD, AV Enhancement	Lab Study, Film	3	ERM	Square wave	Note onset detection	N/A	Hands	Yes	Experience	Normal Hearing, D/HOH, HA Users, CI Users	9, 4, 1, 2
Sound forest (Frid and Lindetorp, 2020)	Digital musical instrument (DMI)	Public Exhibition	1	Voice coil, Bass shaker	Sine wave, Noise	N/A, LPF	N/A	Whole Body, Feet	No	Experience	Normal Hearing	4
N/A (Luo and Hayes, 2019)	HAD, Music Enhancement	Lab Study	1	ERM	Sine wave	F0 extraction	Auditory Frequency to Tactile Frequency	Wrist	Yes	Music Listening Performance	Normal Hearing	8
GLOS (Giulia et al., 2019)	HAD, Speech Enhancement	Lab Study	5	ERM	Sine wave	Speech-to-Text	Complex Mapping	Fingers	Yes	Speech Performance	Normal Hearing	3
N/A (Trivedi et al., 2019)	HAD, Music Enhancement	Lab Study	4	Voice coil	Piano	Frequency Thresholding	Pitch to Position	Leg	Yes	Music Listening Performance	Normal Hearing	5
N/A (Hove et al., 2019)	Music Enhancement	General	1	Subwoofer, Bass shaker	Music	LPF	N/A	Back	No	Experience	Normal Hearing	40
body:suit:score (West et al., 2019)	Music Composition	N/A	60	ERM	Sine wave	Envelope extraction, F0 extraction	Amplitude to Position, Pitch to Position	Whole Body	Yes	N/A	N/A	N/A
N/A (Cieśla et al., 2019)	HAD, Speech Enhancement	Lab Study, Concert	2	Voice coil	Speech	F0 extraction, Envelope extraction	Auditory Frequency to Tactile Frequency	Fingers	No	Discrimination	Normal Hearing	12
N/A (Nakada et al., 2018)	Music Training	Lab Study	1	Voice coil	Sine wave	N/A	Auditory Frequency to Tactile Frequency	Hands	Yes	N/A	N/A	N/A
N/A (Egloff et al., 2018)	Music Enhancement	Lab Study	3	Voice coil	Sine wave	N/A	N/A	Flank, Fingers	No	Discrimination	Normal Hearing	18

(Continued)

10.3389/fcomp.2023.1085539

TABLE 1 (Continued)

Device name	Purpose	Listening situation	Nr. Act.	Actuators type	Signals used	DSP	Mapping scheme	Body area actuated	Wear-able	Evaluation measure	Evaluation population	Nr. Part.
MuSS-Bits++ (Petry et al., 2018)	Music Training	Lab Study, Music teaching for DHOH	1	Voice coil	Music	Frequency extraction, Envelope extraction	Auditory Frequency to Tactile Frequency	Hands	Yes	Music Listening Performance	D/HOH	11
N/A (Huang et al., 2018)	HAD, Music Enhancement	Lab Study	1	Voice coil	Music	LPF	N/A	Fingers	Yes	Discrimination	CI Users	17
LIVEJACKET (Hashizume et al., 2018)	Music Enhancement	Lab Study	22	Mixed	Music	N/A	Instrument to Position	Whole Body	Yes	Experience	Normal Hearing	12
VibGrip++ (Kanebako and Minamizawa, 2018)	Music Enhancement	Lab Study	5	Piezo	Music	N/A	N/A	Hands	Yes	N/A	N/A	N/A
Hedonic Tactile Player (Vallgårda et al., 2017)	Exploratory	General	3	ERM	Square wave, Sine wave, Triangle wave	N/A	N/A	Whole Body	Yes	Experience	Normal Hearing	3
Basslet (Lofelt, 2017)	Music Enhancement, Gaming	General	1	Voice coil	Various Waveforms	LPF	N/A	Wrist	Yes	N/A	N/A	N/A
N/A (Huang et al., 2017)	HAD, Speech Enhancement	Lab Study	1	Voice coil	Speech	F0 extraction, Envelope extraction	N/A	Fingers	Yes	Discrimination	CI Users	10
N/A (Florian et al., 2017)	HAD, Music Enhancement, Dance	Lab Study	5	ERM	Sine wave	Envelope extraction	N/A	Arm	Yes	Experience	D/HOH	45
N/A (Tranchant et al., 2017)	HAD, Music Enhancement, Dance	Lab Study	1	Bass shaker	Music	LPF	N/A	Whole Body, Feet	No	Music Listening Performance	Normal Hearing, D/HOH	14, 7
Smart Finger Braille (Ozioko et al., 2017)	HAD, Speech Enhancement	Lab Study	6	ERM	Sine wave	N/A	N/A	Fingers	Yes	Experience	Normal Hearing	3
Auris System (Araujo et al., 2017)	HAD, Music Enhancement	Lab Study	11	Mixed	Music	F0 extraction, Envelope extraction	Auditory Frequency to Tactile Frequency	Whole Body	No	Experience	Normal Hearing, D/HOH	13

(Continued)

TABLE 1 (Continued)

Device name	Purpose	Listening situation	Nr. Act.	Actuators type	Signals used	DSP	Mapping scheme	Body area actuated	Wear-able	Evaluation measure	Evaluation population	Nr. Part.
MuSS-Bits (Petry et al., 2016)	HAD, Music Training, Music Playing	Exploratory	1	ERM	Sine wave	Envelope extraction	Auditory Amplitude to Tactile Amplitude	Wrist	Yes	Experience	D/HOH	11
Mood Glove (Mazzoni and Bryan-Kinns, 2016)	Film Enhancement	Film	8	ERM	Sine wave	N/A	Complex Mapping	Hands	Yes	Experience	Normal Hearing	10
N/A (Hopkins et al., 2016)	Exploratory	Lab Study	1	EVT	Sine wave	N/A	N/A	Fingers, Feet	No	Discrimination	Normal Hearing, D/HOH	70, 14
N/A (Armitage and Ng, 2016)	Music Enhancement, Music Composition	Artistic	16	ERM	Square wave	Binaural auditory rendering	Auditory Spatialization to Tactile Spatilization	N/A	Yes	N/A	N/A	N/A
Silent Rave Furniture (Jack et al., 2015)	Music Enhancement	Concert	9	Voice coil	Sine wave, Noise	Vocoding	Pitch to Position	Back, Hands, Feet	No	Experience	D/HOH	1
N/A (Hayes, 2015)	Music Enhancement	Artistic	6	Mixed	Square wave, Music	N/A	Pitch to Position	Whole Body	No	N/A	N/A	N/A
CollarBeat (Sakuragi et al., 2015)	Music Enhancement	Lab Study	2	Voice coil	Sine wave, Music	N/A	N/A	Whole Body	No	Experience	Normal Hearing	6
N/A (Young et al., 2015)	Exploratory	Lab Study	6	Voice coil	Various Waveforms	N/A	N/A	Hands	Yes	Discrimination	Normal Hearing	30
N/A (Knutzen et al., 2014)	Music Playing	Lab Study	4	Voice coil	Various Waveforms	F0 extraction	Gesture to Tactile Parameters	Fingers	Yes	Discrimination, Experience	Normal Hearing	5
N/A (Branje et al., 2014)	Film Enhancement	Lab Study	16	Voice coil	Music, Noise	Envelope extraction, Spectrum isolation	Pitch to Position	Back	No	Experience	Normal Hearing	59

(Continued)

TABLE 1 (Continued)

Device name	Purpose	Listening situation	Nr. Act.	Actuators type	Signals used	DSP	Mapping scheme	Body area actuated	Wear-able	Evaluation measure	Evaluation population	Nr. Part.
(Hwang et al., 2013)	Music Enhancement	Lab Study	1	DMA	Sine wave	Spectrum isolation	Auditory Amplitude to Tactile Amplitude	Fingers	No	Experience	Normal Hearing	24
Haptic chair (Nanayakkara et al., 2013)	HAD, Music Enhancement	Lab Study	4, 8	Voice coil	Music	N/A	N/A	Whole Body	No	Experience	D/HOH	43, 12
N/A (Young et al., 2013)	DMI	Lab Study	6	Voice coil	Various Waveforms	N/A	N/A	Fingers	No	Discrimination	Normal Hearing	10
Haptic chair (Nanayakkara et al., 2012)	Speech Enhancement	Lab Study	4	Contact speakers	Speech	N/A	N/A	Whole Body	No	Speech Performance	D/HOH	20
Emoti-chair (Baijal et al., 2012)	Music Enhancement	Lab Study, Workshop	16	Voice coil	Music	Spectrum isolation	Pitch to Position	Back	No	Experience	Normal Hearing, D/HOH	12, 6
N/A (Tessendorf et al., 2011)	Sound Localization	Lab Study	4	ERM	Sine wave	N/A	N/A	Head	No	N/A	N/A	N/A
Emoti-chair (Karam et al., 2010)	Music Enhancement	Lab Study	16	Voice coil	Music	Spectrum isolation, Instrument isolation, Pitch shift	Pitch to Position, Instrument to Position	Back	No	N/A	N/A	N/A
N/A (Branje et al., 2010)	Music Enhancement	Lab Study	1	Voice coil	Square wave	N/A	N/A	Waist	No	Discrimination	Normal Hearing	4
N/A (Bothe et al., 2004)	HAD, Speech Enhancement	Lab Study	5	ERM	Sine wave	Spectrum isolation	Pitch to Position	Fingers	Yes	Speech Performance	Normal Hearing	N/A
N/A (Wada et al., 1999)	HAD, Speech Enhancement	Lab Study	64	Piezo	Sine wave	Spectrum isolation, Envelope extraction, Amplitude modulation	Complex Mapping	Fingers	Yes	Discrimination, Speech Performance	Unspecified	Unspecified

(Continued)

10.3389/fcomp.2023.1085539

TABLE 1 (Continued)

Device name	Purpose	Listening situation	Nr. Act.	Actuators type	Signals used	DSP	Mapping scheme	Body area actuated	Wear-able	Evaluation measure	Evaluation population	Nr. Part.
N/A (Spens et al., 1997)	HAD, Speech Enhancement	Lab Study	2	Unspecified	Unspecified	Envelope extraction	N/A	Hands	Yes	Speech Performance	Normal Hearing	Unspecified
N/A (Milnes et al., 1996)	HAD, Speech Enhancement	Lab Study	1	Piezo	Custom AM, FM coding	Amplitude Thresholding, Spectrum isolation	Complex Mapping	Wrist	Yes	Speech Performance	D/HOH	5
N/A (Miyamoto et al., 1995)	HAD, Speech Enhancement	Lab Study	2	Piezo	Unspecified waveform	Spectrum isolation	Pitch to Position	Sternum	Yes	Speech Performance	CI Users	10
N/A (Osberger et al., 1991)	HAD, Speech Enhancement	Lab Study	7	Piezo	Unspecified waveform	Formant tracking	Complex Mapping	Sternum	Yes	Speech Performance	D/HOH	2
N/A (Weisenberger and Broadstone, 1989)	HAD, Speech Enhancement	Lab Study	1	Unspecified	Unspecified	N/A	N/A	Wrist	Yes	Discrimination, Speech Performance	D/HOH	8
N/A (Weisenberger and Broadstone, 1989)	HAD, Speech Enhancement	Lab Study	16	LRA	Sine wave	Spectrum isolation	Pitch to Position	Arm, Abdomen	Yes	Discrimination	Normal Hearing	3
N/A (Yeung et al., 1988)	HAD, Speech Enhancement	Lab Study	16	Solenoids	Square wave	F0 extraction	Pitch to Position, Pitch to Amplitude	Arm	Yes	N/A	N/A	N/A
N/A (Engebretson and O'Connell, 1986)	HAD, Speech Enhancement	Lab Study	16	Solenoids	Square wave	Spectrum isolation	Pitch to Position	N/A	N/A	N/A	N/A	N/A

A complete explanation of the labels can be found in section 3.1.

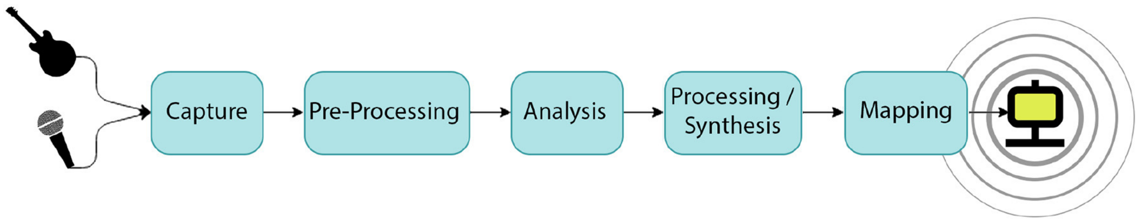

FIGURE 1
Archetype of a vibrotactile augmentation system.

(Hopkins et al., 2021). The last musical dimension that researchers focus on in terms of its vibrotactile properties is the timbre—the tempo-spectral characteristics of the stimulation that translates into the perceived quality of the sound, anecdotally refereed to as the *color of sound or tone color*. As with pitch discrimination, spectral content discrimination is inferior to its auditory counterpart, but there is evidence that humans can identify different spectral characteristics as discrete sensations (Russo, 2019).

Using tactile stimulation to augment auditory signals was first explored as hearing assistive devices focusing on improving speech perception for the D/HOH communities. The first commercial device that promised such results was called *Tickle Talker* and was developed in 1985 by *Cochlear Pty. Ltd* under the supervision of G. M. Clark—inventor of the cochlear implant (CI) (Cowan et al., 1990). The *Tickle Talker* was a *multichannel electrotactile speech processor* that presented speech as a pattern of electrical sensations on 4 fingers. The stimuli presented were processed similarly to the one for early-day CIs. Several other devices emerged in the mid-late 90s that explored the possibilities of using tactile stimulation to enhance speech for the hearing impaired; these will be discussed further in the current article.

What the early devices had in common with modern ones is the fundamental principle they rely on *multisensory integration*, pioneered by B. Stein & A. Meredith (Stein et al., 1993). This mechanism links auditory and tactile sensations and describes how humans form a coherent, valid, and robust perception of reality by processing sensory stimuli from multiple modalities (Stein et al., 1993). The classical rules for multisensory integration demand that enhancement occurs only for stimuli that are temporally coincident and propose that enhancement is strongest for those stimuli that individually are least effective (Stein et al., 1993). This is especially useful for CI users that are shown to be better multisensory integrators (Rouger et al., 2007).

For this integration to occur, the input from various sensors must eventually converge on the same neurons. In the specific case of auditory-somatosensory stimuli, recent studies demonstrate that multisensory integration can, in fact, occur at very early stages of cognition, resulting in supra-additive integration of touch and hearing (Foxe et al., 2002; Kayser et al., 2005). This translates to a lower level of robust synergy between the two sensory apparatuses that can be exploited to synthesize experiences impossible to achieve by unisensory means. Furthermore, research on auditory-tactile interactions has shown that tactile stimulus can influence the auditory stimulus and vice versa (Ro et al., 2009; Okazaki et al., 2012, 2013; Aker et al., 2022). It can therefore be observed that auditory and haptic stimuli are capable of modifying or altering the perception of each other when presented in unison, as described by Young et al. (2016) and Aker et al. (2022), and studied extensively with respect to music experiences (Russo, 2019).

Positive results from the speech experiments, as well as advancement in transducer technology, inspired researchers to explore the benefits of vibrotactile feedback in a musical context. Most of the works fall into two categories: *Musical Haptics*, which focuses mainly on the augmentation of musical instruments, as presented by Papetti and Saitis (2018), and *vibrotactile augmentation* of music listening, generally aimed at D/HOH. The focus of this article will be on the latter. A common system architecture can be seen in Figure 1, with large variability for each step, depending on the goal. For example, in DeGuglielmo et al. (2021) a 4 actuator system is used to enhance music discrimination in a live concert scenario by creating a custom mapping scheme between the incoming signal and the frequency and amplitude of the transducers. A contrasting goal is presented by Reed et al. (2021), where phoneme identification in speech is improved by using a total of 24 actuators. These two examples have different objectives, thus the systems have different requirements, but the overall architecture of both follows the one shown in Figure 1. Throughout the article, each block will be analyzed in detail, and each system description can be seen in Table 1.

Currently, research into tactile displays is expanding from speech to various aspects of music enhancement, although the field of research is still relatively new; for example, over 80% of the articles included in this review are newer than 2009. Nevertheless, the technology is slowly coming out of research laboratories and into consumers' hands, with bands such as *Coldplay* offering *SubPacs*[2] for their D/HOH concert audience.

1.3. Definitions

Before presenting the objectives, it is worth introducing some general terminology and describing the interpretation used throughout this article. The first clarification is with respect to how words such as *tactile, vibrotactile and haptic* are used in this article:

2 https://subpac.com/

- **Haptics** refer to "*the sensory inputs arising from receptors in skin, muscles, tendons, and joints that are used to derive information about the properties of objects as they are manipulated*," explains (Jones, 2009). It is worth highlighting that haptic sensations involve tactile systems and proprioceptive sensory mechanisms.
- **Tactile** refers to the ability of the skin to sense various stimulations, such as physical changes (mechanoreceptors), temperature (thermoreceptors), or pain (nociceptors), according to Marzvanyan and Alhawaj (2019). There are six different types of mechanoreceptors in the skin, each with an individual actuation range and frequency, and together they respond to physical changes, including touch, pressure, vibration, and stretch.
- **Vibrotactile** refers to the stimulation presented on the skin that is produced by oscillating devices. The authors of Marzvanyan and Alhawaj (2019) present evidence that two types of mechanoreceptors respond to vibrotactile stimulation; namely, Pacinian receptors and, to a lesser extent, Meissner corpuscles. These receptors have been frequently analyzed and characterized in terms of their frequency and amplitude characteristics, but it is not excluded that there are other mechanoreceptors responsible for tactile percepts.

These definitions should provide a basic understanding of the taxonomy necessary to interpret this article; nevertheless, an individual study is recommended for a better understanding of the field of physiology and neuroscience. Furthermore, the authors recommend choosing the most descriptive terminology when discussing augmentation in order to avoid potential confusion (e.g., the use of *vibrotactile augmentation* instead of *haptic augmentation* is preferred when describing a system that involves vibrations). Finally, the term **augmentation** will be used as *the process of increasing the cognitive, perceptual or emotional, value, or quality* of the listing experience. Throughout this article, augmentation generally involves the usage of dedicated, specialized hardware (HW) and software (SW) systems.

Tactile augmentation of music is a fairly new multidisciplinary research field; therefore, it is paramount to achieve consensus on terminology and definitions.

2. Methodology

The methodology used to select and analyze the literature will be presented in the following section, starting with the system used to include articles and followed by an explanation of the process used to extract data from the article pool. No a priori review protocol was applied throughout the data collection method. The section ends by presenting known limitations in the methods described. The entire review process followed the *PRISMA-ScR* checklist and structure (Tricco et al., 2018).

2.1. Identifying relevant studies

The *Scopus*® database was queried due to its high Scientific Journal Rankings required for inclusion, as well as significance for the topic. The inclusion selection was a 4 step process:

Step 1 included deciding upon the selection of keywords (below), designed to cover various aspects of tactile music augmentation.

1. Audio-haptic sensory substitution
2. Audio-tactile sensory substitution
3. Cochlear implant music
4. Cochlear implant (vibro)tactile
5. Cochlear implant haptic (display)
6. Electro-haptic stimulation
7. Hearing impaired music augmentation
8. Hearing impaired (vibro)tactile
9. (Vibro)tactile music
10. (Vibro)tactile display
11. (Vibro)tactile augmentation
12. (Vibro)tactile audio feedback.

As of 23.02.2022 a total of 3555 articles were found that contain at least one of the items present in the list above in their title, abstract or keywords section. There was no discrimination between the types of documents that were included, but due to the database's inclusion criteria, PhD thesis and other potentially non-peer-reviewed works are not present.

Step 2 was a selection based on the title alone. Throughout steps 2–4, the eligibility criteria were as follows:

1. Written in English
2. Reporting primary research
3. Must describe devices that are, or could be used for tactile augmentation of music for D/HOH*
4. Must be designed for an audience (as opposite to a performer).

*A device that can be used for tactile augmentation of music for the hearing impaired (as item 3 describes) can be a system that was designed for laboratory studies that focused on the augmentation of speech and not music. Furthermore, a system for musical augmentation of normal hearing people could be used for D/HOH individuals since the tactile receptors do not differ depending on hearing capabilities; more-so there is evidence that the perception of congruent tactile stimulation is elevated in CI users (Rouger et al., 2007). Since the focus is on the technological aspects of these tactile displays and not on their efficiency, it was deemed relevant to include vibrotactile systems that are designed for music augmentation for normal hearing individuals as well.

The selection process was completed with the inclusion of 144 articles for further investigation.

Step 3 represented the reading of abstracts and evaluating their relevance; 102 articles were selected according to the eligibility criteria.

Multiple articles from the same authors describing the same setups were included only once, selecting the most recent publication and excluding the older ones. This would be the case where the authors evaluated multiple hypotheses using the same HW/SW setup.

Step 4 constituted of reading the articles selected in Step 3; the same inclusion criteria were used. This step coincided with data extraction, but articles that had been wrongfully included based on the abstract alone were discarded. The entire inclusion process resulted in 63 articles that were used to construct this review.

2.2. Data items

The relevant articles were studied and a selection of descriptors was noted for each system presented; as mentioned in Section 1.1, the focus is on hardware, software and evaluation practices, making the analysis an agnostic process with respect to their applications, target groups, or success. If the article discussed more than one system, all relevant ones were included and analyzed. The features used to analyze and compare the systems were:

1. **Purpose of the display** represents the end goal of the device, irrespective of the eventual evaluation conducted. All documented purposes for each display were included. A common purpose would be "Speech Enhancement."
2. **Listening situation** refers to the context where the display has been used. A frequent situation is the "Laboratory Study."
3. **Number of actuators** indicates the total number of actuators in the display, regardless of their type.
4. **Actuators type** enumerates the type of actuators used, from a hardware perspective. For example, eccentric rotating mass (ERM) vibration motors are one of the most widely used haptic technologies.
5. **Signals used to feed the actuators** represents the type of audio material used to excite the transducers. In the case of ERMs, most displays will indicate a "sine wave," unless otherwise noted. This is due to the hardware nature of the actuator, capable of producing only sinusoidal oscillations, while the actual signal used is a DC voltage. Some ERM systems can reproduce harmonically complex signals as well, but the cases are few and far apart.
6. **Type of signal processing** (generally called DSP) enumerates the processes applied to an audio input signal to extract relevant information or prepare it for the tactile display. Fundamental frequency (F0) extraction is a common signal processing technique used for tactile displays.
7. **Mapping scheme** describes the features from the auditory input that are mapped to the tactile output.
8. **Area of the body actuated** presents where the actuators are placed on the human body (e.g., hands and chest). Throughout this article *"hand"* is used a combination of 2 or more sub-regions(e.g., palm and fingers)
9. **Whether it is a wearable device or not** (binary Yes/No)
10. **Evaluation measurement** presents the measurement criteria assessed in the evaluation (if applicable). Most items in this descriptor column have been grouped into a meta-category; for example "vocal pitch accuracy" seen in Shin et al. (2020) and "pitch estimation accuracy" from Fletcher (2021) have been grouped into "Music Listening Performance."
11. **Evaluation population** describes the hearing characteristics of the individuals participating in the evaluation (if applicable).
12. **Number of participants** presents the total number of participants in the evaluation described by the previous two items.

These features were chosen in order to create an objective and complete characterization of each system while allowing a high degree of comparability. Furthermore, features regarding evaluation were included to better understand the research field as a whole. A detailed explanation for some relevant categories can be found in Section 3.1. Based on these features, Table 1 was constructed that contains all the systems analyzed.

2.3. Delimitation

To ensure that the inclusion process was feasible, we imposed several constraints on the process, which may have influenced some aspects of the review. First, only one database was used to browse for articles, by only one author. Although inquiring several databases is recommended, the Scopus® database already provided a large sample of publications, and includes many if not all the relevant journal publishers (e.g., IEEE). Furthermore, using the articles indexed in an academic database results in the inevitable exclusion of artistic work that might have limited exposure but with a potentially valuable contribution. Since music is inherently an artistic expression and not an academic work, this limitation could greatly impact the results of the analysis presented in this article. Similarly, articles written in other languages than English were excluded, resulting in many works not having a chance of inclusion.

3. Results

This section provides a thematic analysis of the systems included in the reviewing process. The data is first presented as a table of characteristics that describes each system and its usage, and it will be succeeded by a graphical representation of the most important findings.

3.1. Table explanation

Due to the large variation in the hardware design and evaluation methods used in the literature presented in Table 1, some categories contain meta-descriptors that encapsulate similar features. This section will identify and clarify these situations, as well as provide an additional explanation that would allow readers an easier interpretation of the Table 1.

The first notion worth explaining is the usage of the *"N/A"* acronym. Although most of the time it should be understood as *"Not Available"*—information that is not present in the cited literature, some situations fit the interpretation more to *"not applicable"*. One example can be encountered in the *"Nr. part."* category, where *"N/A"* generally means "not applicable" if the tactile display has not been evaluated at all. A more detailed version of the table, as well as the analysis software, can be found at Tactile Displays Review Repository.[3]

Second, the columns *"Evaluation Population"* and *"Nr. Part."* are linked, and the number of participants separated by commas represents each population described in the previous category. For example, in Sharp et al. (2020), *Evaluation Population* contains: *Normal Hearing, Hearing aid(HA) Users, CI Users, D/HOH* while *"Nr. Part."* contains *10, 2, 6, 2*; this should be red as 10 normal

3 https://github.com/razvysme/TactileDisplaysReview

hearing participants, 2 hearing aid users, 6 cochlear implant users, and 2 deaf or hard-of-hearing individuals.

Third, in the *"Mapping"* column several items called *"Complex Mapping"* that encapsulate all the more elaborate mappings, as well as mappings that have been specially designed for the tactile display in question.

Fourth, the tactile displays analyzed usually fall under more than one category, with respect to those ones mentioned in Section 2.2. For example, Figure 10 shows *purpose of devices*, where several tactile displays fall under HAD as a secondary purpose, on top of their main goal (e.g., Music Enhancement). This means that analyzed items in most of the figures presented in Section 3.2 are not exclusive, resulting in a greater total amount of data points than number of displays included.

Next, a brief description of the actuator types encountered in the review and shown in Figure 4 is presented below. It is worth noting that the list ignores auxiliary systems necessary to operate these actuators (converters, amplifiers, etc.), and describes them based on their practical tactile applications:

- **Voice Coil** actuators get their name from the most common application: moving the paper cone in a speaker and are also known as non-commutated DC linear actuators. They consist of a permanent magnetic element (sometimes replaced by an electromagnet with the same role) and a suspended coil attached to a mobile mass. A variation in this architecture exists, where the permanent magnet is the moving mass, and the coil is static. The current from an amplifier that flows through the coil, creates an electromagnetic field that interacts with the permanent magnet, moving the mass (or the paper cone) accordingly. Voice coil actuators come in various sizes and forces available, have a relatively wide frequency response (in the KHz range), and provide high acceleration.
- **Subwoofers** are a type of voice coil actuators that are optimized to reproduce sound at low frequencies, commonly below 80Hz. Their construction and size vary radically depending on the application, but they do imply the properties described above. One drawback of subwoofers is that they usually require generous amplification to operate. Furthermore, they are optimized for sound reproduction, therefore their tactile characteristics are usually a byproduct of high amplitude playback.
- **Solenoids** are somewhat similar to voice coils, but instead of providing a permanent magnet interacting with an electromagnet, they have a coil creating a magnetic field in order to move a ferrous shaft. Solenoids are generally used to open or close locks, valves, or to apply a constant force on a surface and are not necessarily suitable for oscillating behavior, resulting in a limited frequency response.
- **Piezo-actuators** are mechanisms that vibrate based on the change in the shape of a piezoelectric material and belong to the category of *"resonating actuators"*, which have an efficient operating frequency embedded in their mechanical design. Because of that, piezo-actuators have limited frequency response (generally within 80% of the resonant frequency), but they can be designed to be tiny or in complex shapes, as opposed to the ones presented above. The tactile feedback produced by piezo-actuators is relatively modest, for a given current.
- **Linear Resonant Actuators (LRA)** are mass-spring systems that employ a suspended mass attached to an electromagnetic coil that vibrates in a linear fashion due to the interaction with the permanently magnetized enclosure. Being a resonant system, they need to be driven with signals close to the peak frequency response, similar to piezo-actuators and ERMs described below. Some advanced LRAs have auto-resonance systems that detect the optimal frequency for producing highest amplitude tactile feedback, trading tactile frequency accuracy for perceived intensity.
- **Electro-dynamic shakers (EDS)** is an industrial name given to vibration systems with excellent frequency response characteristics, generally necessary in vibration analysis and acoustics industries. They come in two categories, voice coils and electro-hydraulic shakers, but the systems described in Table 1 only use voice the former type. These can be seen more like bass shakers described below, but with much better frequency response as well as a much higher cost.
- **Eccentric Rotating Mass (ERMs)** are another type of resonating actuators that operate by attaching an unbalanced mass to the shaft of a DC motor. Rotating the mass produces vibrations of different frequencies and amplitudes, typically linked to the amount of current fed to the motor. These types of actuators are very popular due to their low cost, and relatively strong vibration force, but they respond slower than other resonators to a chance in the current (a lag of 40-80ms is commonly expected). Another limitation of ERMs, as well as LRAs and piezo-actuators is that the frequency and amplitude reproduced are correlated due to their resonating design, and thus are generally suggested when limited tactile frequency information is necessary.
- **Dual-Mode Actuators (DMAs)** are relatively new types of actuator that are similar to LRAs, but are designed to operate at two different frequencies simultaneously, usually out of phase with each other. Due to their novelty, the amount of variation and experimentation with them is limited, but Hwang et al. (2013) provides evidence that DMAs outperform LRAs in tactile displays in music as well as HCI applications.
- **Contact speakers** are a sub-category of voice coils that are primarily designed to excite hard surfaces in order to produce sound. They work by moving a suspended coil that has a shaft with a contact surface at the end. The contact surface is usually glued or screw on the desired surface, thus vibrating as the coil oscillates. Contact speakers vary largely in size and power requirements, but generally provide a wide frequency response. In the context of tactile stimulation, they usually have a poor low-frequency representation (below 100 Hz) and are always producing sound (sometimes very loud), due to their focus on the auditory reproduction.
- **Bass shakers** are another sub-category of voice coils that are grouped by marketed applications rather than physical properties. They work by having a large mass attached to the moving coil, usually in a protective enclosure. These devices

are suggested to be used in an audio listening scenario (be it films or games) and should be attached to seating furniture (sofa, chairs, etc). Some vendors provide mounting hardware for drum stools or vibrating platforms designed for stage musicians that complement in-ear headphones to monitor band activity. Bass shakers are generally large, heavy, and require abundant power to operate, but provide a relatively good frequency response up to approx. 350Hz.

3.2. Synthesis of results

Besides Table 1, plots and histograms highlighting the most interesting relationships between characteristics are discussed.

When doing the literature search for scientific publications that included the phrases described in section 2.1, we found that 84% of tactile displays have been introduced after 2010, as shown in Figure 2. This is strong evidence that interest in integrating tactile stimulation into the music-listening activities is blooming.

3.2.1. Distribution of types of actuators used over time

One of the main descriptors of a system is the number of actuators it uses, and it could easily be (wrongfully) assumed that the advancement in transducer technology would encourage researchers to use more actuators in their studies. As shown in Figure 3, the average number of actuators decreases slightly over the years, and the predictor model (blue line) suggests that this number will not increase in the near future. In addition to the occasional outliers with more than 60 actuators, most systems use less than 20 transducers and more than half use less than the average of 8.

Plotting the distribution of actuators over time in Figure 4, shows that voice coils have been used since the 2010s, and are generally preferred for applications that require higher frequency and amplitude accuracy. The drawbacks of this type of actuators are that they are generally larger, more expensive, and would require a more complex Digital-to-Analog (DAC) to operate since they use bipolar signals. Including subwoofers, contact speakers and bass shakers in the "voice coil" category would create a cluster representing 34%, signifying their importance. Another popular choice is eccentric rotating mass (ERM) actuators, which are smaller, cheaper, and simpler to operate than voice coils, but provide limited frequency response, and the amplitude is coupled to the frequency. It is also interesting to observe that older systems used piezo and solenoids–technologies that have a very limited application range with small amplitude, or small frequency response. Lastly, it's important to highlight that one category of electro-dynamic systems is never used for tactile augmentation—electro-hydraulic shakers. These devices can provide large displacement and could be deployed to actuate very large surfaces, but have a limited frequency response, generally below 200 Hz.

3.2.2. Mappings

Figure 5 shows that before 1990 mapping schemes were rather simple (e.g., mapping pitch-to-position or amplitude). Tactile frequency has only recently been brought into discussion probably because of the high computational power required for the analysis stage, but also because of recent advancements in voice coil actuators. This technical progress allowed researchers to explore the tactile frequency as a method of encoding the auditory frequency. The pitch-to-position mapping (the idea of cochlear implants) is the one that is used the longest, probably because the creator of the CI and *Tickle Talker* used the same mapping for both. This could have been an inspiration for further research to produce incremental improvements in these systems, rather than revolutionary approaches, as can be seen in the work of Karam and Fels (2008) and Nanayakkara et al. (2012). Around 2015, new mappings start to be explored, and the popularity of "Pitch-to-position" starts decreasing.

Looking at the relationship between the mapping schemes and the number of actuators in Figure 6 it is interesting to highlight the fact that only one mapping scheme utilizes tactile frequency, in combination with voice coil actuators. Furthermore, *something*-to-position mapping is popular, taking advantage of the larger surface area the body can afford. This hypothesis is reinforced by Figure 12 showing that areas of high sensitivity are excited with a small number of actuators, probably constrained by the actuator size.

3.2.3. Evaluation practices

Observing Figure 7 it can be seen that speech research was the main focus before 2013, while music has been the topic of more investigation since. This can be explained by the advancement in hearing aids and CI technology that solved the speech intelligibility problem to a satisfactory degree. Nevertheless, the need to use tactile augmentation remains present when music is played for D/HOH or CI users due to the problems shown by all hearing assistive devices have with multi-stream, complex signals, as well as with timbre and melody recognition and sound localization in the case of CIs. Simultaneously, subjective experiences of users have become a topic of interest in the second decade of the century, as seen in the work of Mirzaei et al. (2020) or DeGuglielmo et al. (2021), even though tactile augmentation research has been exploring that topic for more than 27 years. A further look at the plot shows a large amount of research on users' experiences between 2000 and 2010. At the same time, there is still interest in finding the limits of the physiological and cognitive systems involved in tactile perception and integration, as evidenced by the large number of studies involving discrimination tasks conducted in the past decade. An extended version of Figure 7, reinforces the fact that speech enhancement has been a focus since the beginning of tactile augmentation, as shown in Figure 8. Although interest in it has decreased in recent years, there are still researchers working on it. The second wave of tactile augmentation research started around 2010 and it focused mostly on music, but new technologies and media consumption methods show interest into tactile stimulation as well (e.g., gaming and enhancement of film and AV).

The overall number of participants is very small, with a mean of 11, resulting in generally low-reliability studies, as displayed in Figure 9. Furthermore, we can see that the hearing aid and CI users studies are mostly below this average. This is a strong indicator that most of the research is preliminary and underlines the necessity for studies to adhere to clinical guidelines (longitudinal, more participants, better control), as well as the need for replication of

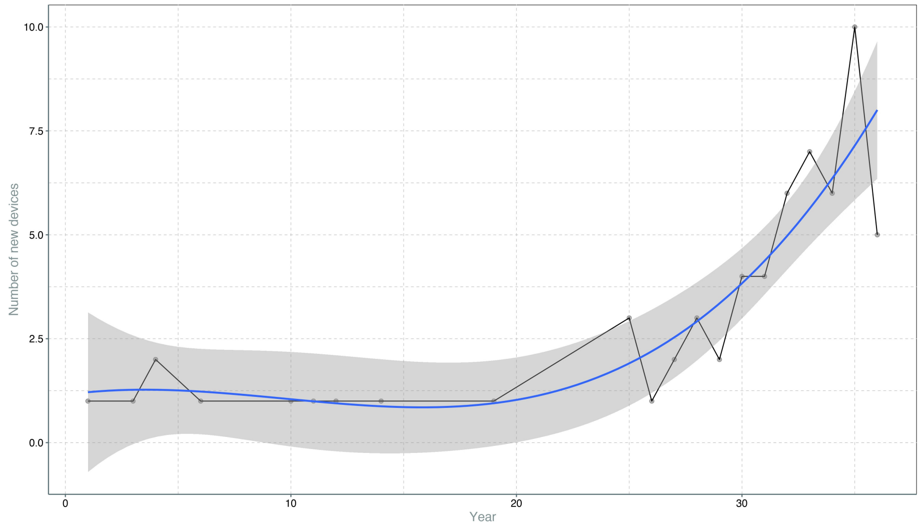

FIGURE 2
Distribution of the number of new systems described in publications every year; black line is the empirical value, the blue line is the predicted value and the grey area represents the 95% confidence interval for the estimated number of studies. The prediction model uses a cubic regression that was chosen due to the lowest AIC and BIC scores (as described by Burnham and Anderson, 2004) when compared to other models of orders 1–5.

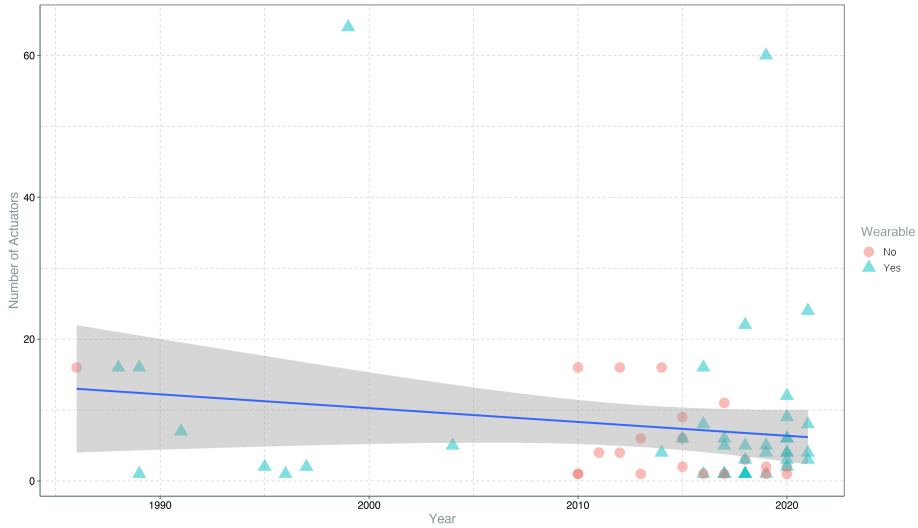

FIGURE 3
Distribution of the number of actuators in a device over years; dashed orange line is the mean over years, the blue line is predicted value, and the grey area represents the 95% confidence interval for the estimated number of actuators; a blue triangle indicates a wearable device, and a red circle indicates a fixed device. The prediction model uses a linear regression that was chosen due to the lowest AIC and BIC scores (as described by Burnham and Anderson, 2004) when compared to other models of orders 1–5.

prior studies. Nevertheless, CI users are mostly evaluating systems focusing on music and speech, while HA users only focus on music as seen in Figure 10. This could be explained by the fact that the speech needs for HA users are fulfilled by the current state-of-the-art in HA tech. However, it could be argued that tactile augmentation could also be interesting for normal hearing and hearing aid users as well; this can be seen in the push for multimodal cinema experiences. Nevertheless, it cannot be overlooked that devices aimed at groups with a particular set of requirements (CI and HA users) are evaluated using a normal hearing population, sometimes exclusively. This practice, while common, might result in studies with lower validity, because the requirements for the target population are not met.

In Figure 11 the relationship between listening situations and evaluation measures is presented. On one hand, the large number of discrimination studies further highlight the incipient state of the research field, which looks to outline the "playing field" by testing the threshold and *just-noticeable differences*. On the other hand, studies focusing on music are increasingly more frequent, indicating that researchers bring forward new systems and ideas

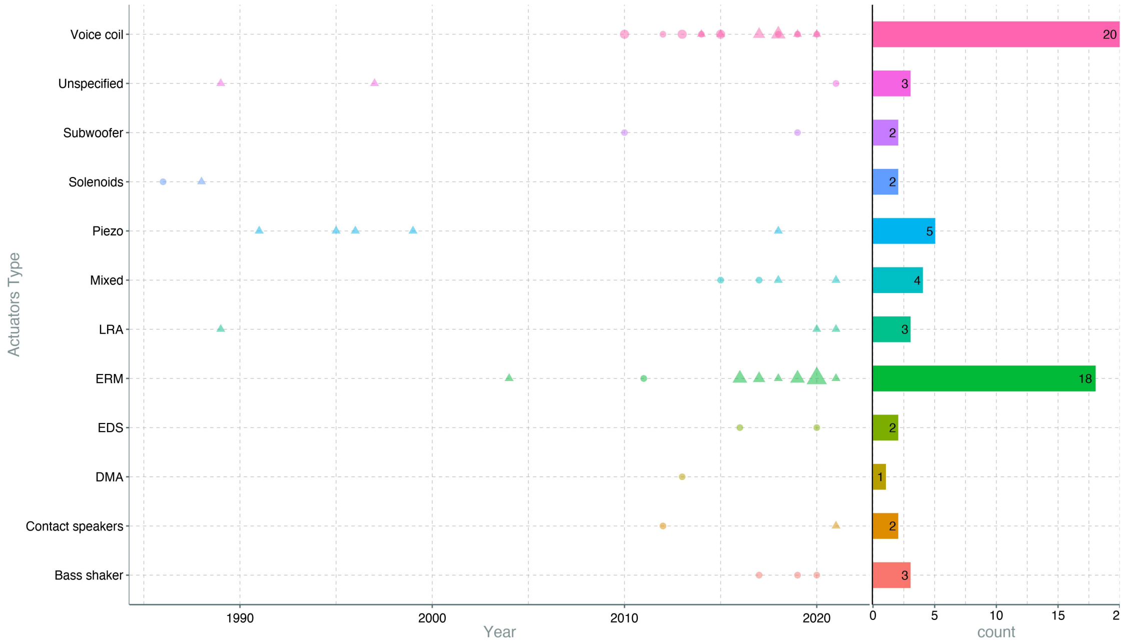

FIGURE 4
Distribution of types of actuators used over time; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices.

that have a more applied research angle to them. This aspect is supported as well by the large emphasis on the users' experience, which is evaluated in various listening situations, not only in research laboratories.

3.2.4. Body regions used for stimulation

Figure 12 shows that systems using fingers use a few actuators, usually one actuator per finger. Similarly, for the wrists, it seems that a small number of actuators is preferred, probably because of physical limitations. On the other hand, when the whole body is used, the number of actuators increases; the same is true for arms. This clearly indicates that the size of the actuator, as well as the spatial resolution of the skin, is a constraining factor in designing complex systems, and advancements in actuator technologies will allow designers to insert more actuators aiming at the high sensitivity areas (hands, wrist, etc). It is known that there is a positive correlation between 2-point discrimination and the contact size of the actuator, so small actuators would require a smaller space between them.

The wrist and fingers are used with mappings that require greater accuracy in terms of frequency, while mappings that rely on simpler encoding (such as amplitude and custom encoding) are generally used with a higher number of actuators, as shown in Figure 12; a potential explanation for this is that the lower physiological accuracy in certain areas is compensated for by a higher number of actuators. Going back to the number of actuators in areas that are using *"auditory frequency to tactile frequency"* mapping, it seems that one or two actuators are sufficient for this type of mapping.

4. Discussion

4.1. Summary of evidence

In this scoping review, we identified 63 primary articles that describe unique vibrotactile displays used for audio augmentation, published from 1986 to 2021. Within this specific research pool, our findings highlight that most of the work in the field of vibrotactile augmentation of sound can be categorized as preliminary, missing the large-scale studies usually associated with clinical research. This conclusion is supported by the low reliability of evaluations presented derived from a low number of participants, as well as the occasional low validity of the said evaluations. The latter is evidenced by experiments conducted with poorly sampled individuals; for example, tactile displays designed for D/HOH are evaluated on normal hearing users. Finally, it should be underlined that much of the available literature covers research conducted under laboratory conditions and not in ecologically valid environments. As such, this contradicts the idea that most long-term benefits are obtained when participants use hearing assistive devices in daily life scenarios. For these reasons we can see an gap in the evaluation and experimental protocols conducted in most studies included, and we suggest that researchers start focusing on larger scale, longitudinal studies that are more akin to clinical ones, when evaluating their audio augmenting tactile displays.

Looking at the hardware aspect, the majority of included studies present tactile displays designed for some regions of the hand. This is expected as hands provide the highest fine motor skills as well as very good spatio-temporal resolution, but it might not be the most practical for devices that are

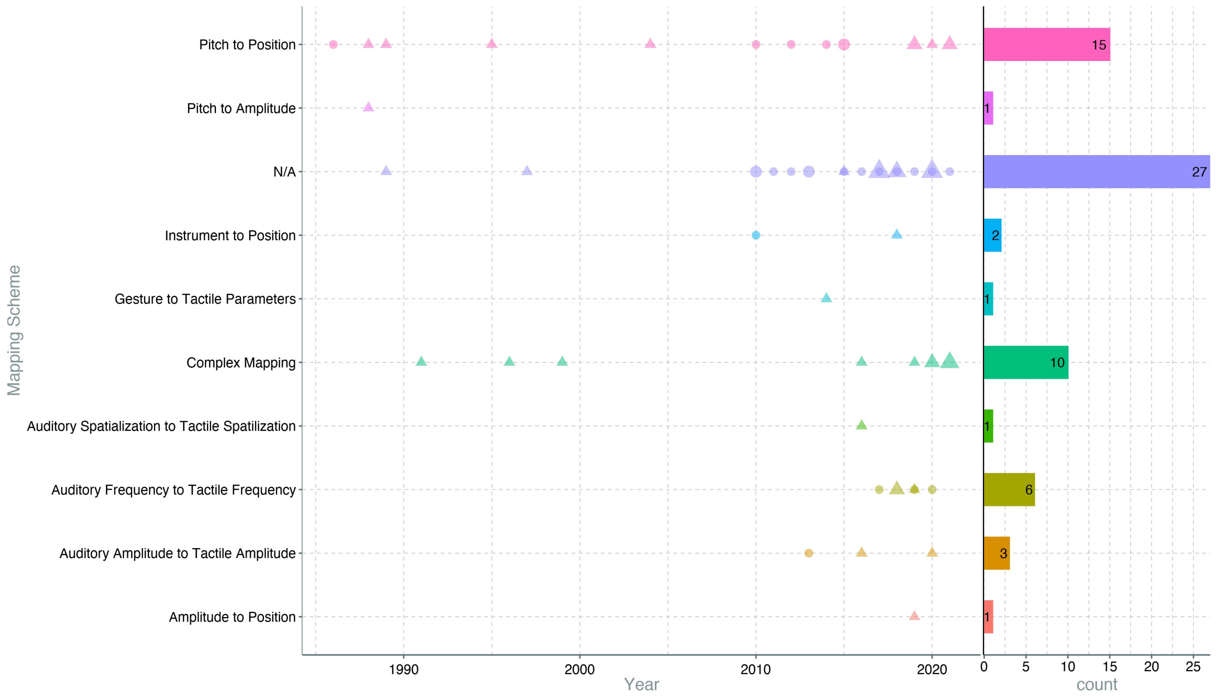

FIGURE 5
Distribution of mapping schemes used over time; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices.

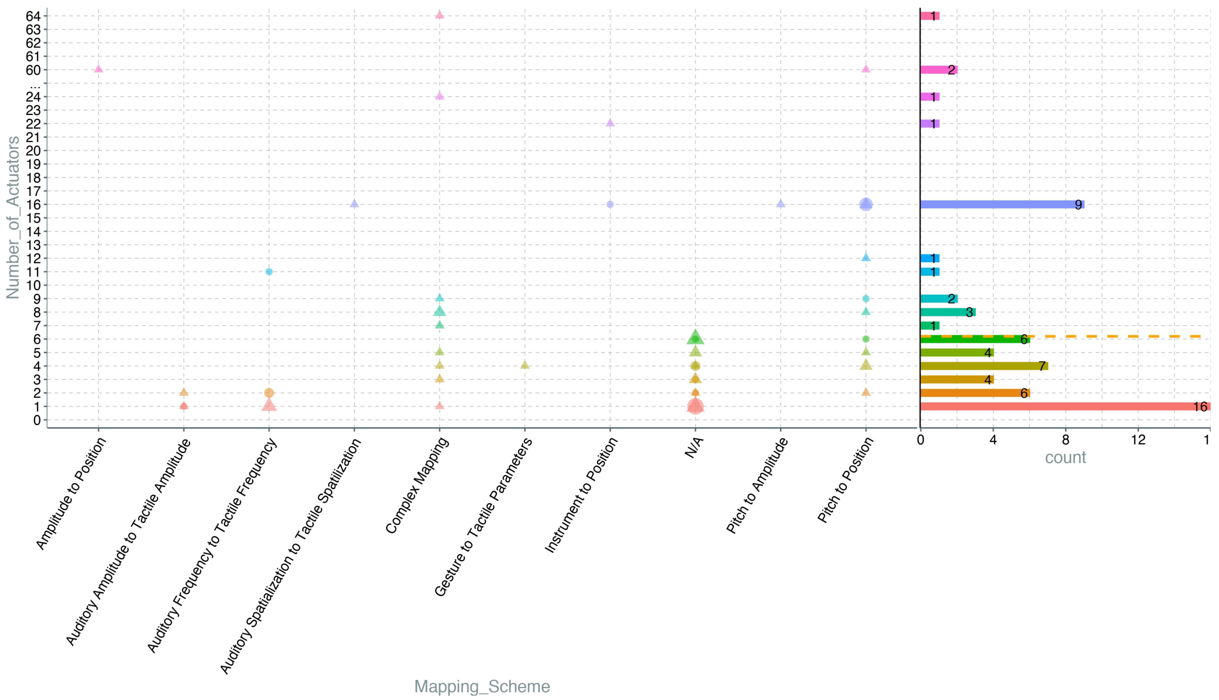

FIGURE 6
Distribution of the number of actuators for each mapping scheme; a triangle indicates a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices. The dashed orange line represents the mean.

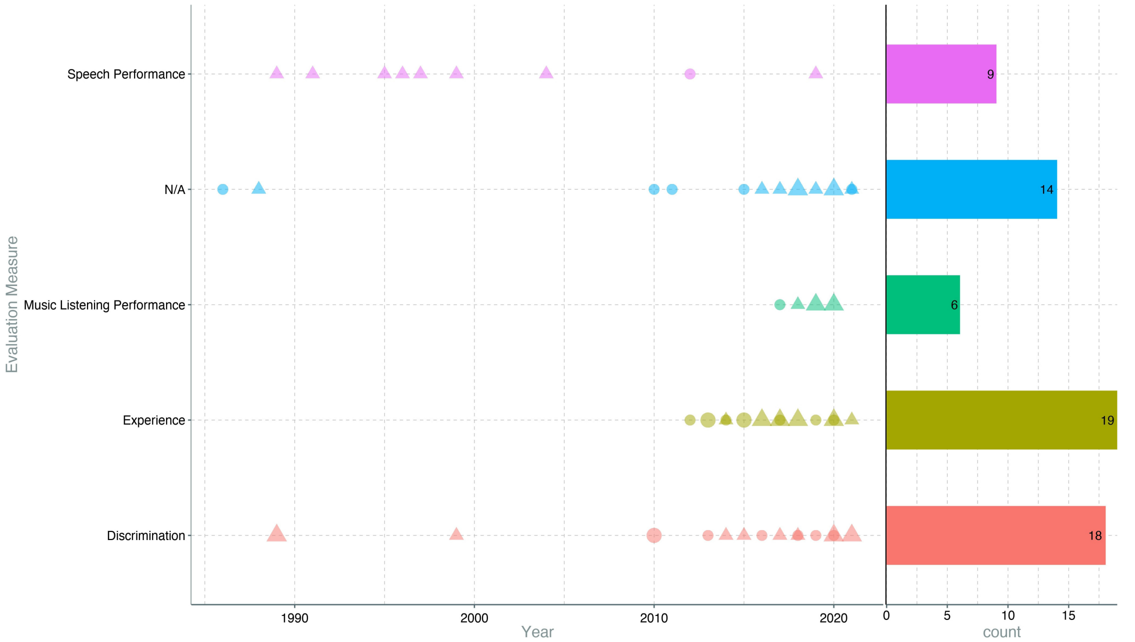

FIGURE 7
Distribution of evaluation measurements over years, a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices.

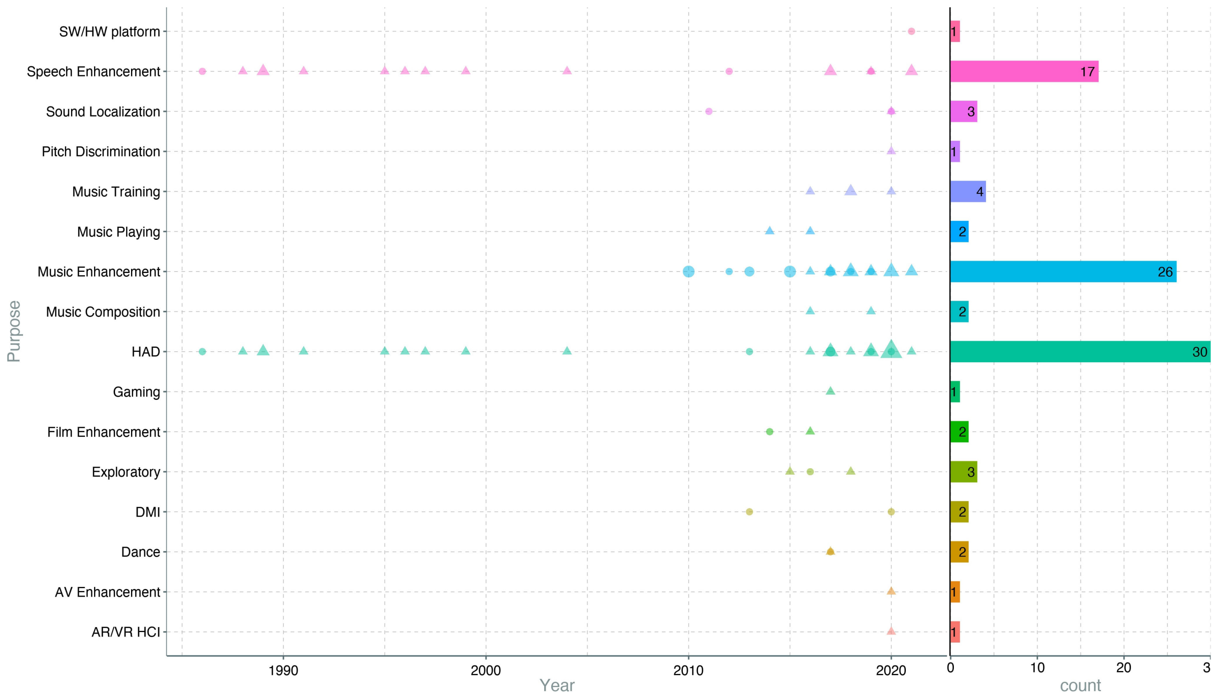

FIGURE 8
Distribution of devices' purpose over year, a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices; items on Y axis are not exclusive.

designed to be used extensively, especially when daily activities are to be executed while using the tactile displays. Furthermore, other areas could present different advantages in terms of duration of stimulation or intensity tolerance for use cases where the finest discrimination properties are not vital. Therefore, we see a great opportunity to branch out and encourage designers and researchers to create displays that afford similar perceptual characteristics and are to be sensed by different body

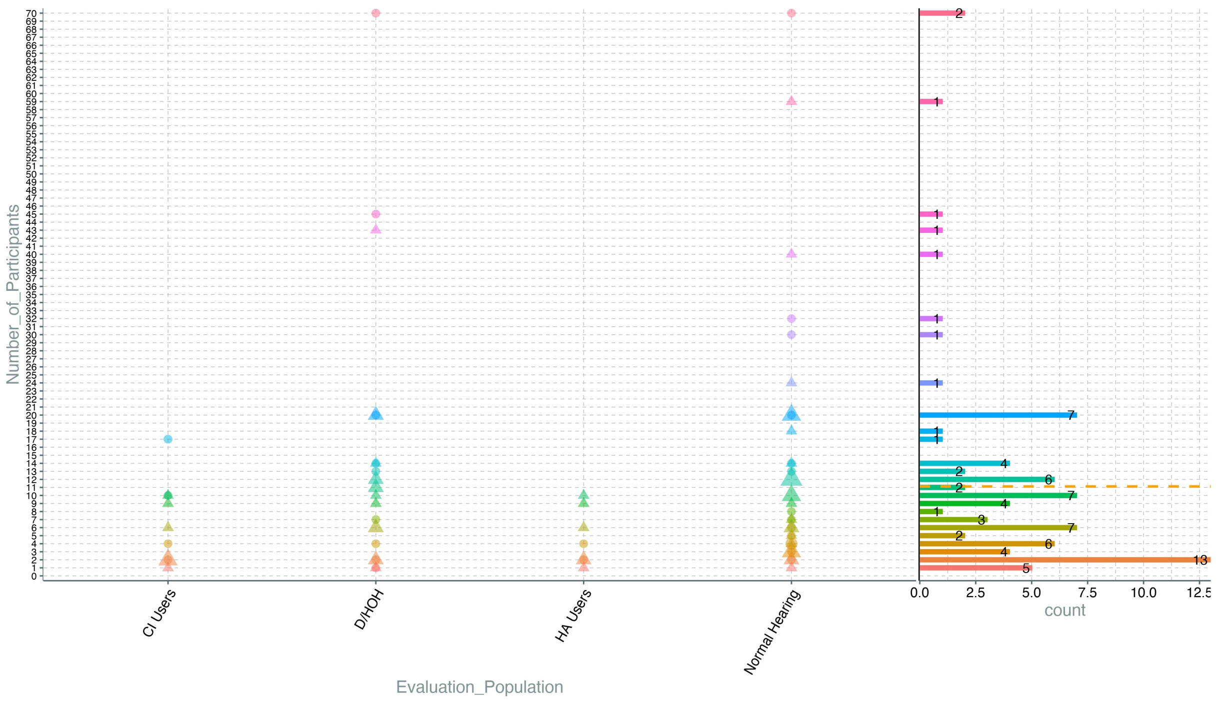

FIGURE 9
Distribution of the number of participants for each of the populations evaluated; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices. *D/HOH* represents persons suffering from a hearing disability, but without any hearing assistive device. The dashed orange line represents the mean.

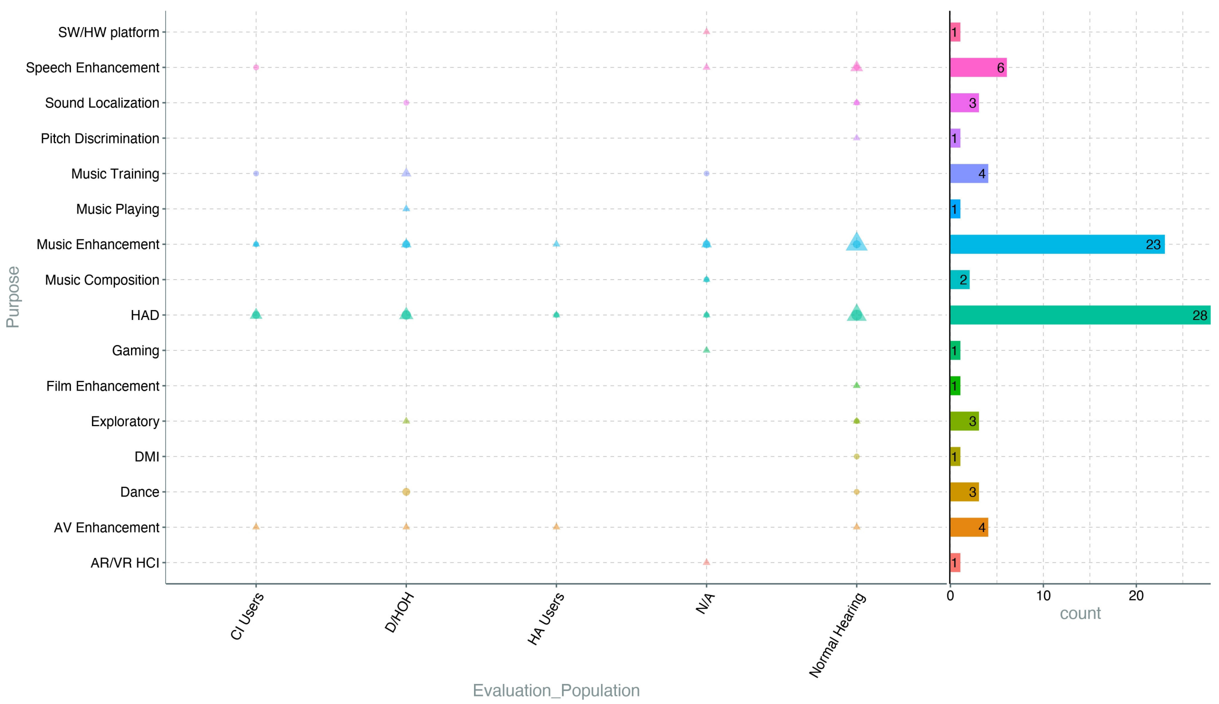

FIGURE 10
Distribution of the purpose for each evaluation population; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices. *D/HOH* represents persons suffering from a hearing disability, but without any hearing assistive device; items on Y axis are not exclusive.

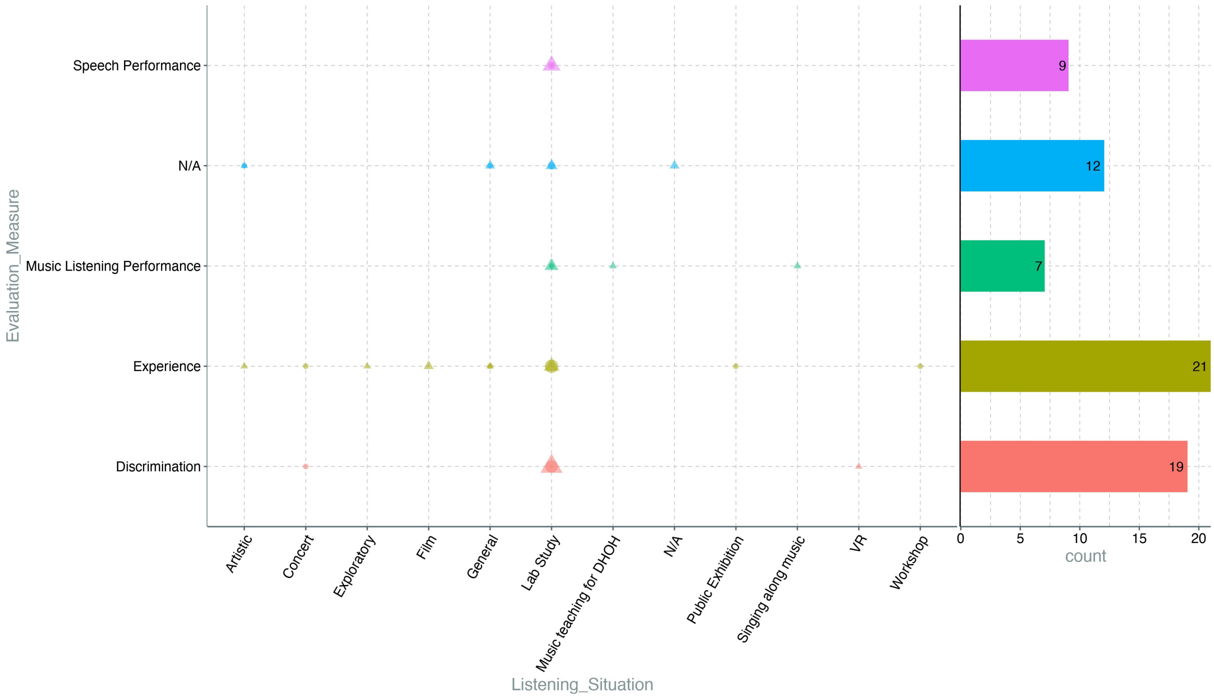

FIGURE 11
Distribution of the evaluation measure goals for each listening situation; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices; items on X axis are note exclusive.

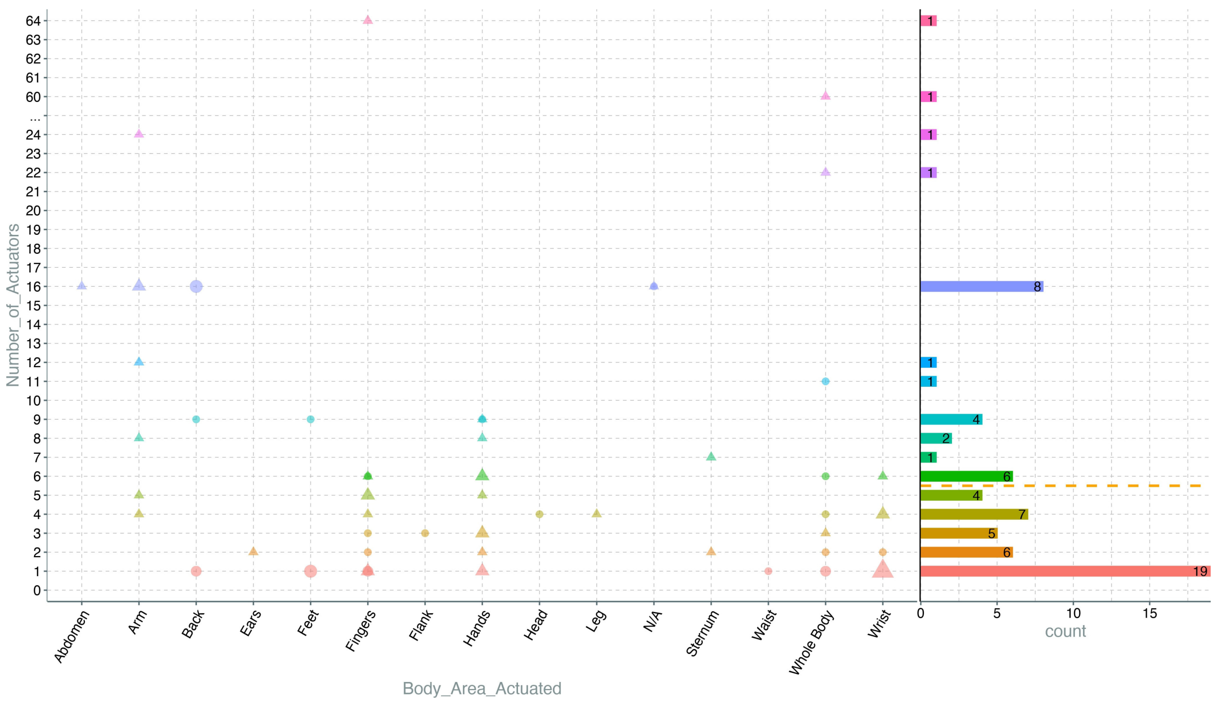

FIGURE 12
Distribution of number of actuators over the body area they actuate; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices. The dashed orange line represents the mean.

regions. Furthermore, versatile designs are strongly encouraged (in terms of mode of interaction, HW/SW and mapping), in order to be adaptable to the inter-user needs; this is especially important for CI users, where the variation in hearing abilities is largest.

On a similar note, our finding indicates that researchers present a large variety of designs in terms of type of actuators, mapping, signal processing, etc. further solidify the exploratory phase of the entire research field, characteristic of an early development stage. There seems to be a consensus on the upper limit of the number of actuators necessary for vibrotactile augmentation, although more than half of the displays identified use less than the average number of transducers, with the mode being 1 actuator. This could be attributed to cost reduction strategies, but since most of the devices are researched almost exclusively in laboratories and are generally far from commercialization, we are confident to suggest that a high number of actuators might not provide substantial benefits to tactile augmentation of audio. With this in mind, we must emphasize the importance of mapping strategies used, both in the time/frequency domain and in the psychoperceptual space, in order to design the best tactile displays for vibrotactile augmentation. Our research pool shows that almost half of the devices studied do not imply any form of mapping between the auditory signal and tactile stimulation, while the mode is *pitch-to-position*—a mapping scheme introduced with the very first audio-tactile augmentation device. We see a great potential for exploration into creative mapping schemes that could have roots in the most commonly encountered ones, as well as radical new ideas that could be generic or case-specific (e.g., bespoke for concert or film scenarios). These new mappings should be carefully designed and evaluated primarily with respect to the target group's hearing profiles (CI, HA, etc.) as well as signal processing used and the eventual acoustic stimulation; if possible, all these aspects should be co-created involving the end users, in order to produce a coherent multisensory experience.

4.2. Limitations

This is the first scoping review focusing on the technological aspect behind the vibrotactile augmentation of music. Although mainly concerned with the hardware and software characteristics of the tactile displays described in Table 1, this article has addressed some elements of the dissemination and evaluation of the devices described. Nevertheless, the scoping nature of this review rules out a detailed description of implementation for each study or evaluates the quality and effectiveness of the included tactile display. Therefore, it is impossible to recommend specific techniques or strategies that would predict better music perception, training, or adjacent metrics for D/HOH and CI implanted people.

While a comprehensive search has been conducted on one of the most relevant databases, this process was carried out by a single reviewer and there was no forward-citation search on the included studies. Furthermore, there was no review of the reference list of included articles or a manual search protocol to scan relevant journals, as it was concluded that most of the articles are indexed by Scopus®. This resulted in the exclusion of any gray literature, as the process of searching for relevant unpublished material was of considerable difficulty.

4.3. Conclusions

The purpose of this scoping review was to investigate and report the current technological state in the field of vibrotactile augmentation, viewed from the perspective of music enhancement for hearing impaired users. A total of 3555 articles were considered for eligibility from the Scopus® database, resulting in the inclusion of 63 studies. The vibrotactile devices in each article was analyzed according to a pre-defined set of characteristics, focusing on hardware and software elements, as well as the evaluation and experiment design, regardless of the hearing profile of their users. The evidence gathered indicates that this research field is in an early phase, characterized by an exploratory approach and preliminary results. A secondary objective of this article was to identify the gaps and trends in the literature that can guide researchers and designers in their practice, and a list of suggestions and recommendations has been presented, based on graphical representations of statistics analysis. The data and the system used to synthesize the review are publicly accessible, and we recommend that readers explore them and generate their own graphs and interpretations.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://github.com/razvysme/TactileDisplaysReview.

Author contributions

RP, NN, and SS contributed to the conception and design of the study, and including the selection of databases to be queried. RP browsed the database and organized the data collected, performed the statistical analysis, and wrote the first draft of the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

Funding

This research was funded by the Nordfosk Nordic University Hub through the Nordic Sound and Music project (number 86892), and as well as the Nordic Sound and Music Computing Network and Aalborg University.

Conflict of interest

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

Publisher's note

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

References

Aker, S. C., Innes-Brown, H., Faulkner, K. F., Vatti, M., and Marozeau, J. (2022). Effect of audio-tactile congruence on vibrotactile music enhancement. *J. Acoust. Soc. Am.* 152, 3396–3409. doi: 10.1121/10.0016444

Araujo, F. A., Brasil, F. L., Santos, A. C. L., Junior, L. D. S. B., Dutra, S. P. F., and Batista, C. E. C. F. (2017). Auris system: providing vibrotactile feedback for hearing impaired population. *Biomed. Res. Int.* 2017, 2181380. doi: 10.1155/2017/2181380

Armitage, J., and Ng, K. (2016). "Feeling sound: exploring a haptic-audio relationship," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Plymouth), *Vol.* 9617, 146–152.

Baijal, A., Kim, J., Branje, C., Russo, F. A., and Fels, D. (2012). "Composing vibrotactile music: a multi-sensory experience with the emoti-chair," in *Haptics Symposium 2012, HAPTICS 2012-Proceedings* (Vancouver, BC), 509–515.

Bothe, H.-H., Mohamad, A., and Jensen, M. (2004). "Flipps: fingertip stimulation for improved speechreading," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),Vol. 3118* (Berlin), 1152–1159.

Branje, C., Maksimowski, M., Karam, M., Fels, D., and Russo, F. A. (2010). "Vibrotactile display of music on the human back," in *3rd International Conference on Advances in Computer-Human Interactions, ACHI 2010* (St. Maarten), 154–159.

Branje, C., Nespoil, G., Russo, F. A., and Fels, D. (2014). The effect of vibrotactile stimulation on the emotional response to horror films. *Comput. Entertainment* 11, 703. doi: 10.1145/2543698.2543703

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding aic and bic in model selection. *Sociol. Methods Res.* 33, 644. doi: 10.1177/0049124104268644

Chafe, C. (1993). "Tactile audio feedback," in *Proceedings of the 19th International Computer Music Conference (ICMC 1993)* Tokyo.

Cieśla, K., Wolak, T., Lorens, A., Heimler, B., Skaryński, H., and Amedi, A. (2019). Immediate improvement of speech-in-noise perception through multisensory stimulation via an auditory to tactile sensory substitution. *Restor. Neurol Neurosci.* 37, 155–166. doi: 10.3233/RNN-190898

Cowan, R. S. C., Sarant, J. Z., Galvin, K. L., Alcantara, J. I., Blamey, P. J., and Clark, G. M. (1990). The tickle talker: a speech perception aid for profoundly hearing impaired children. *Aust. J. Audiol.* 9.

DeGuglielmo, N., Lobo, C., Moriarty, E. J., Ma, G., and Dow, D. E. (2021). "Haptic vibrations for hearing impaired to experience aspects of live music," in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST, 403 LNICST*; Sugimoto.

Egloff, D., Wanderley, M., and Frissen, I. (2018). Haptic display of melodic intervals for musical applications. *IEEE Haptics Sympos.* 2018, 284–289. doi: 10.1109/HAPTICS.2018.8357189

Engebretson, A., and O'Connell, M. (1986). Implementation of a microprocessor-based tactile hearing prosthesis. *IEEE Trans. Biomed. Eng.* 33, 712–716. doi: 10.1109/TBME.1986.325762

Fletcher, M. (2021). Using haptic stimulation to enhance auditory perception in hearing-impaired listeners. *Expert Rev. Med. Devices* 18, 63–74. doi: 10.1080/17434440.2021.1863782

Fletcher, M., Thini, N., and Perry, S. (2020). Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic. *Sci. Rep.* 10, 10354. doi: 10.1038/s41598-020-67140-0

Fletcher, M., and Zgheib, J. (2020). Haptic sound-localisation for use in cochlear implant and hearing-aid users. *Sci. Rep.* 10, 14171. doi: 10.1038/s41598-020-70379-2

Florian, H., Mocanu, A., Vlasin, C., Machado, J., Carvalho, V., Soares, F., et al. (2017). Deaf people feeling music rhythm by using a sensing and actuating device. *Sens. Actuators A Phys.* 267, 431–442. doi: 10.1016/j.sna.2017.10.034

Foxe, J., Wylie, G., Martinez, A., Schroeder, C., Javitt, D., Guilfoyle, D., et al. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* 88, 540–543. doi: 10.1152/jn.2002.88.1.540

Frid, E., and Lindetorp, H. (2020). "Haptic music-exploring whole-body vibrations and tactile sound for a multisensory music installation," in *Proceedings of the Sound and Music Computing Conferences 2020-June* (Turin), 68–75.

Gilmore, S. A., and Russo, F. A. (2020). Neural and behavioral evidence for vibrotactile beat perception and bimodal enhancement. *J. Cogn. Neurosci.* 33,1673. doi: 10.1162/jocn_a_01673

Giulia, C., Chiara, D., and Esmailbeigi, H. (2019). "Glos: Glove for speech recognition," in *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2019* (Berlin: IEEE), 3319–3322.

Hashizume, S., Sakamoto, S., Suzuki, K., and Ochiai, Y. (2018). "Livejacket: wearable music experience device with multiple speakers," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 10921 LNCS* (Frankfurt), 359–371.

Hayes, L. (2015). "Skin music, An audio-haptic composition for ears and body," in *C and C 2015-Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition* (Glasgow), 359–360.

Haynes, A., Lawry, J., Kent, C., and Rossiter, J. (2021). Feelmusic: enriching our emotive experience of music through audio-tactile mappings. *Multimodal Technol. Interact.* 5, 29. doi: 10.3390/mti5060029

Hopkins, C., Maté-Cid, S., Fulford, R., Seiffert, G., and Ginsborg, J. (2016). Vibrotactile presentation of musical notes to the glabrous skin for adults with normal hearing or a hearing impairment: thresholds, dynamic range and high-frequency perception. *PLoS ONE* 11, e0155807. doi: 10.1371/journal.pone.0155807

Hopkins, C., Maté-Cid, S., Fulford, R., Seiffert, G., and Ginsborg, J. (2021). Perception and learning of relative pitch by musicians using the vibrotactile mode. *Musicae Scientiae* 27, 3–26. doi: 10.1177/10298649211015278

Hove, M., Martinez, S., and Stupacher, J. (2019). Feel the bass: music presented to tactile and auditory modalities increases aesthetic appreciation and body movement. *J. Exp. Psychol. Gen.* 149, 1137–1147. doi: 10.1037/xge0000708

Huang, J., Lu, T., Sheffield, B., and Zeng, F.-G. (2018). Electro-tactile stimulation enhances cochlear-implant melody recognition: effects of rhythm and musical training. *Ear. Hear.* 41, 106–113. doi: 10.1097/AUD.0000000000000749

Huang, J., Sheffield, B., Lin, P., and Zeng, F.-G. (2017). Electro-tactile stimulation enhances cochlear implant speech recognition in noise. *Sci. Rep.* 7, 2196. doi: 10.1038/s41598-017-02429-1

Hwang, I., Lee, H., and Choi, S. (2013). Real-time dual-band haptic music player for mobile devices. *IEEE Trans. Haptics* 6, 340–351. doi: 10.1109/TOH.2013.7

Jack, R., Mcpherson, A., and Stockman, T. (2015). "Designing tactile musical devices with and for deaf users: a case study," in *Proceedings of the International Conference on the Multimedia Experience of Music*; Treviso.

Jiam, N. T., and Limb, C. J. (2019). Rhythm processing in cochlear implant–mediated music perception. *Ann. N. Y. Acad. Sci.* 1453, 22–28. doi: 10.1111/nyas.14130

Jones, L. A. (2009). *Haptic*,. Berlin; Heidelberg: Springer Berlin Heidelberg.

Kanebako, J., and Minamizawa, K. (2018). Vibgrip++: haptic device allows feeling the music for hearing impaired people. *Lecture Notes Electr. Eng.* 432, 449–452. doi: 10.1007/978-981-10-4157-0_75

Karam, M., Branje, C., Nespoli, G., Thompson, N., Russo, F., and Fels, D. (2010). "The emoti-chair: an interactive tactile music exhibit," in *Conference on Human Factors in Computing Systems-Proceedings* (Atlanta, GA), 3069–3074.

Karam, M., and Fels, D. (2008). "Designing a model human cochlea: issues and challenges in crossmodal audio-haptic displays," in *Proceedings of the 2008 Ambi-Sys Workshop on Haptic User Interfaces in Ambient Media Systems, HAS 2008*; Brussels.

Kayser, C., Petkov, C., Augath, M., and Logothetis, N. (2005). Integration of touch and sound in auditory cortex. *Neuron* 48, 373–384. doi: 10.1016/j.neuron.2005.09.018

Knutzen, H., Kvifte, T., and Wanderley, M. M. (2014). "Vibrotactile feedback for an open air music controller," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Marseille), *Vol. 8905*.

Lauzon, A. P., Russo, F. A., and Harris, L. R. (2020). The influence of rhythm on detection of auditory and vibrotactile asynchrony. *Exp. Brain Res.* 238, 825–832. doi: 10.1007/s00221-019-05720-x

Lofelt (2017). *Basslet*. Available online at: https://eu.lofelt.com/ (accessed on June 30, 2018).

Luci,á, M., Revuelta, P., Garci, á,., Ruiz, B., Vergaz, R., Cerdán, V., et al. (2020). Vibrotactile captioning of musical effects in audio-visual media as an alternative for deaf and hard of hearing people: an EEG study. *IEEE Access* 8, 190873–190881. doi: 10.1109/ACCESS.2020.3032229

Luo, X., and Hayes, L. (2019). Vibrotactile stimulation based on the fundamental frequency can improve melodic contour identification of normal-hearing listeners with a 4-channel cochlear implant simulation. *Front. Neurosci.* 13, 1145. doi: 10.3389/fnins.2019.01145

Marzvanyan, A., and Alhawaj, A. F. (2019). *Physiology, Sensory Receptors*. Tampa, FL: StatPearls Publishing LLC.

Mazzoni, A., and Bryan-Kinns, N. (2016). Mood glove: a haptic wearable prototype system to enhance mood music in film. *Entertain Comput.* 17, 9–17. doi: 10.1016/j.entcom.2016.06.002

Milnes, P., Stevens, J., Brown, B., Summers, I., and Cooper, P. (1996). "Use of micro-controller in a tactile aid for the hearing impaired," in *Annual International Conference of the IEEE Engineering in Medicine and Biology-Proceedings* (Amsterdam: IEEE), 413–414.

Mirzaei, M., Kán, P., and Kaufmann, H. (2020). Earvr: using ear haptics in virtual reality for deaf and hard-of-hearing people. *IEEE Trans. Vis. Comput. Graph.* 26, 2084–2093. doi: 10.1109/TVCG.2020.2973441

Miyamoto, R., Robbins, A., Osberger, M., Todds, S., Riley, A., and Kirk, K. (1995). Comparison of multichannel tactile aids and multichannel cochlear implants in children with profound hearing impairments. *Am. J. Otol.* 16, 8–13.

Morley, J. W., and Rowe, M. J. (1990). Perceived pitch of vibrotactile stimuli: effects of vibration amplitude, and implications for vibration frequency coding. *J. Physiol.* 431, 8336. doi: 10.1113/jphysiol.1990.sp018336

Nakada, K., Onishi, J., and Sakajiri, M. (2018). "An interactive musical scale presentation system via tactile sense through haptic actuator," in *HAVE 2018—IEEE International Symposium on Haptic, Audio-Visual Environments and Games, Proceedings* (Dalian: IEEE), 13–16.

Nanayakkara, S., Wyse, L., Ong, S., and Taylor, E. (2013). Enhancing musical experience for the hearing-impaired using visual and haptic displays. *Human Comput. Interact.* 28, 115–160. doi: 10.1080/07370024.2012.697006

Nanayakkara, S., Wyse, L., and Taylor, E. (2012). "The haptic chair as a speech training aid for the deaf," in *Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012* (Melbourne, VIC), 405–410.

Okazaki, R., Hachisu, T., Sato, M., Fukushima, S., Hayward, V., and Kajimoto, H. (2013). Judged consonance of tactile and auditory frequencies. *World Haptics Conf.* 2013, 663–666. doi: 10.1109/WHC.2013.6548487

Okazaki, R., Kajimoto, H., and Hayward, V. (2012). "Vibrotactile stimulation can affect auditory loudness: a pilot study," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications* (Kyoto), 103–108.

Osberger, M., Robbins, A., Todd, S., and Brown, C. (1991). Initial findings with a wearable multichannel vibrotactile aid. *Am. J. Otol.* 12, 179–182.

Ozioko, O., Taube, W., Hersh, M., and Dahiya, R. (2017). Smartfingerbraille: a tactile sensing and actuation based communication glove for deafblind people. *IEEE International Symposium on Industrial Electronics* (Edinburgh: IEEE), 2014–2018.

Pacchierotti, C., Sinclair, S., Solazzi, M., Frisoli, A., Hayward, V., and Prattichizzo, D. (2017). Wearable haptic systems for the fingertip and the hand: taxonomy, review, and perspectives. *IEEE Trans. Haptics* 10, 9006. doi: 10.1109/TOH.2017.2689006

Papetti, S., and Saitis, C. (2018). *Musical Haptics: Introduction*. Cham: Springer.

Pavlidou, A., and Lo, B. (2021). "Artificial ear-a wearable device for the hearing impaired," in *2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2021, 2021-January* (Athens: IEEE).

Petry, B., Illandara, T., Elvitigala, D., and Nanayakkara, S. (2018). "Supporting rhythm activities of deaf children using music-sensory-substitution systems," in *Conference on Human Factors in Computing Systems-Proceedings, 2018-April*; Montreal, QC.

Petry, B., Illandara, T., and Nanayakkara, S. (2016). "Muss-bits: sensor-display blocks for deaf people to explore musical sounds," in *Proceedings of the 28th Australian Computer-Human Interaction Conference, OzCHI 2016* (Launceston, TAS), 72–80.

Pezent, E., Cambio, B., and Ormalley, M. (2021). Syntacts: open-source software and hardware for audio-controlled haptics. *IEEE Trans. Haptics.* 14, 225–233. doi: 10.1109/TOH.2020.3002696

Pezent, E., O'Malley, M., Israr, A., Samad, M., Robinson, S., Agarwal, P., et al. (2020). "Explorations of wrist haptic feedback for ar/vr interactions with tasbi," in *Conference on Human Factors in Computing Systems-Proceedings*; Hawaii.

Reed, C. M., Tan, H. Z., Jiao, Y., Perez, Z. D., and Wilson, E. C. (2021). Identification of words and phrases through a phonemic-based haptic display. *ACM Trans. Appl Percept.* 18, 8725. doi: 10.1145/3458725

Remache-Vinueza, B., Trujillo-León, A., Zapata, M., Sarmiento-Ortiz, F., and Vidal-Verdú, F. (2021). Audio-tactile rendering: a review on technology and methods to convey musical information through the sense of touch. *Sensors* 21, 6575. doi: 10.3390/s21196575

Ro, T., Hsu, J., Yasar, N., Elmore, C., and Beauchamp, M. (2009). Sound enhances touch perception. *Exp. Brain Res.* 195, 135–143. doi: 10.1007/s00221-009-1759-8

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., and Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7295–7300. doi: 10.1073/pnas.0609419104

Russo, F. A. (2019). "Multisensory processing in music," in *The Oxford Handbook of Music and the Brain*; Oxford.

Rydel, A., and Seiffer, W. (1903). Untersuchungen über das vibrationsgefühl oder die sog. knochensensibilität (pallästhesie). *Archiv für Psychiatrie und Nervenkrankheiten* 37, 8367. doi: 10.1007/BF02228367

Sakuragi, R., Ikeno, S., Okazaki, R., and Kajimoto, H. (2015). "Collarbeat: whole body vibrotactile presentation via the collarbone to enrich music listening experience," in *International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, ICAT-EGVE 2015* (Kyoto), 141–146.

Setzepfand, W. (1935). Zur frequenzabhängigkeit der vibrationsempfindung des menschen. *Z. Biol.* 96, 236–240.

Sharp, A., Bacon, B., and Champoux, F. (2020). Enhanced tactile identification of musical emotion in the deaf. *Exp. Brain Res.* 238, 1229–1236. doi: 10.1007/s00221-020-05789-9

Shin, S., Oh, C., and Shin, H. (2020). "Tactile tone system: a wearable device to assist accuracy of vocal pitch in cochlear implant users," in *ASSETS 2020-22nd International ACM SIGACCESS Conference on Computers and Accessibility*; New York, NY.

Shull, P. B., and Damian, D. D. (2015). Haptic wearables as sensory replacement, sensory augmentation and trainer-a review. *J. Neuroeng. Rehabil.* 12, 59. doi: 10.1186/s12984-015-0055-z

Sorgini, F., Cali,ò, R., Carrozza, M. C., and Oddo, C. M. (2018). Haptic-assistive technologies for audition and vision sensory disabilities. *Disabil. Rehabil. Assist. Technol.* 13, 394–421. doi: 10.1080/17483107.2017.1385100

Spens, K.-E., Huss, C., Dahlqvist, M., and Agelfors, E. (1997). A hand held two-channel vibro-tactile speech communication aid for the deaf: characteristics and results. *Scand. Audiol. Suppl.* 26, 7–13.

Stein, B., Stein, P., and Meredith, M. (1993). *The Merging of the Senses. A Bradford Book*. Cambridge, MA: MIT Press.

Swerdfeger, B. A., Fernquist, J., Hazelton, T. W., and MacLean, K. E. (2009). "Exploring melodic variance in rhythmic haptic stimulus design," in *Proceedings-Graphics Interface*; Kelowna, BC.

Tessendorf, B., Roggen, D., Spuhler, M., Stiefmeier, T., Tröster, G., Grämer, T., et al. (2011). "Design of a bilateral vibrotactile feedback system for lateralization," in *ASSETS'11: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee), 233–234.

Tranchant, P., Shiell, M., Giordano, M., Nadeau, A., Peretz, I., and Zatorre, R. (2017). Feeling the beat: Bouncing synchronization to vibrotactile music in hearing and early deaf people. *Front. Neurosci.* 11, 507. doi: 10.3389/fnins.2017.00507

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). Prisma extension for scoping reviews (prisma-scr): checklist and explanation. *Ann. Internal Med.* 169, 467–473. doi: 10.7326/M18-0850

Trivedi, U., Alqasemi, R., and Dubey, R. (2019). "Wearable musical haptic sleeves for people with hearing impairment," in *ACM International Conference Proceeding Series* (New York, NY), 146–151.

Turchet, L., West, T., and Wanderley, M. M. (2020). Touching the audience: musical haptic wearables for augmented and participatory live music performances. *Pers. Ubiquitous Comput.* 25, 749–769. doi: 10.1007/s00779-020-01395-2

Vallgårda, A., Boer, L., and Cahill, B. (2017). The hedonic haptic player. *Int. J. Design* 11, 17–33.

Wada, C., Shoji, H., and Ifukube, T. (1999). Development and evaluation of a tactile display for a tactile vocoder. *Technol. Disabil.* 11, 151–159. doi: 10.3233/TAD-1999-11305

Weisenberger, J., and Broadstone, S. (1989). Evaluation of two multichannel tactile aids for the hearing impaired. *J. Acoust. Soc. Am.* 86, 1764–1775. doi: 10.1121/1.398608

West, T., Bachmayer, A., Bhagwati, S., Berzowska, J., and Wanderley, M. (2019). "The design of the body:suit:score, a full-body vibrotactile musical score," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Orlando, FL), *Vol. 11570 LNCS*, 70–89.

Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. *Acta Physiol. Scand.* 31, 284–289. doi: 10.1111/j.1748-1716.1954.tb01139.x

Yeung, E., Boothroyd, A., and Redmond, C. (1988). A wearable multichannel tactile display of voice fundamental frequency. *Ear Hear.* 9, 342–350. doi: 10.1097/00003446-198812000-00011

Young, G., Murphy, D., and Weeter, J. (2013). "Audio-tactile glove," in *DAFx 2013-16th International Conference on Digital Audio Effects*; Maynooth.

Young, G., Murphy, D., and Weeter, J. (2015). "Vibrotactile discrimination of pure and complex waveforms," in *Proceedings of the 12th International Conference in Sound and Music Computing, SMC 2015* (Maynooth), 359–362.

Young, G., Murphy, D., and Weeter, J. (2016). Haptics in music: the effects of vibrotactile stimulus in low frequency auditory difference detection tasks. *IEEE Trans. Haptics* 10, 135–139. doi: 10.1109/TOH.2016.2646370

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 11 May 2023
DOI 10.3389/fcomp.2023.1153232

Check for updates

OPEN ACCESS

EDITED BY
Kjetil Falkenberg,
Royal Institute of Technology, Sweden

REVIEWED BY
Alex Michael Lucas,
Queen's University Belfast, United Kingdom
Stefania Serafin,
Aalborg University Copenhagen, Denmark

*CORRESPONDENCE
Fabio Morreale
✉ f.morreale@auckland.ac.nz

SPECIALTY SECTION
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Computer Science

RECEIVED 29 January 2023
ACCEPTED 18 April 2023
PUBLISHED 11 May 2023

CITATION
McMillan A and Morreale F (2023) Designing accessible musical instruments by addressing musician-instrument relationships. *Front. Comput. Sci.* 5:1153232. doi: 10.3389/fcomp.2023.1153232

Designing accessible musical instruments by addressing musician-instrument relationships

Andrew McMillan and Fabio Morreale*

Te Whare o ngā Pūkōrero Pūoro (School of Music), Waipapa Taumata Rau (The University of Auckland), Tāmaki Makaurau (Auckland), New Zealand

This article explores the concept of intimacy in the relationship between a musician and their instrument, specifically in the context of designing digital and accessible musical interfaces (DMI/AMI) with disabled musicians. We argue that current DMI/AMI design frameworks are lacking in their consideration of this relationship and that this deficiency can prevent designers from understanding the specific needs and desires of disabled musicians. The paper presents an autoethnographic study of the lived experience of the first author, a disabled musician, to provide insight into the evolution of his musician-instrument relationships and his definition of "success" in this context. The authors propose that incorporating these types of lived experiences into the AMI design process, and considering cultural probes or provocations related to phenomenological experiences and characteristics that contribute to a successful musician-instrument relationship, could lead to more effective and tailored DMI/AMI designs with disabled musicians.

KEYWORDS

accessible musical instruments, musician-instrument relationship, digital musical instruments, design framework, conceptual mapping

1. Introduction

Designers of Digital Musical Interfaces (DMI) and Accessible Musical Instruments (AMI)[1] have provided several frameworks to assist with the design of these categories of instruments (Jordà, 2004; Johnston et al., 2008; Overholt, 2009; Morreale et al., 2014). These frameworks have proved helpful in defining the musician's goals, physical abilities, and environments and matching them to existing technology. However, these frameworks lack propositions on how to inquire into and design for the intimacy of the musician-instrument relationship. As we argue in this paper, we hold that this enquiry is essential in designing bespoke DMIs, especially with disabled musicians.

Research on musician-instrument relationships has been studied in relation to non-disabled musicians connecting with conventional instruments (Nijs et al., 2009; Simoens and Tervaniemi, 2013; Waters, 2021), but specific investigations with disabled musicians and DMIs/AMIs are currently missing. Also, current DMI/AMI frameworks have

1 Whether *I* in both acronyms refers to *Interfaces* or *Instruments* is still the subject of a scholarly debate (see, for instance, Jensenius and Lyons, 2017) that we don't intend to contribute to. Some AMIs are not digital; thus AMI are not fully included within the DMI set.

been primarily constructed by non-disabled designers. Although they often forge constructive relationships with disabled musicians, they lack the lived experience of disability. We propose that this deficiency not only prevents designers from having an intimate understanding of a disabled musician's needs and desires for a successful musician-instrument relationship but also—and probably *mostly*—what "successful" actually means.

In this paper, through the lived experiences of the first author, Andrew McMillan, a musician with acquired disability, we propose that questions need to be asked about the relationship with an instrument when defining guidelines for AMIs. Following an autoethnographic methodology, we provide insights into the evolution of Andrew's musician-instrument relationships. Specifically, we account for the evolution of his practice and his relationship; first, as a non-disabled musician performing traditional instruments, then as a disabled musician performing DMIs. We also discuss "success" in this context by outlining Andrew's objectives in pursuing an equivalent disabled musician-instrument relationship. As a disabled musician and DMI designer and researcher, Andrew has the unique vantage point of being able to intimately investigate the gap between the critical aspects that drive the successful musician-instrument relationship for a person using AMI and place the information gathered from his insights into design discourse.

The goal of this article is twofold. First, it identifies the necessity to account for the relationship between a disabled musician and their instrument in the design process of AMIs. Second, it offers methodological suggestions on how to elicit discussions around a disabled musician's phenomenological experiences and characteristics that contribute toward a successful musician-instrument relationship. We identify that cultural probes and conceptual metaphors are excellent tools for this purpose.

The rest of this article is structured as follows. In the next section, we cover the related work in this article's various areas of interest. The successive section is the core part of this article, in which we present an autoethnographic account of the first author's experience and how his relationship with musical instruments evolved after a life-changing accident. We then discuss the implications of this work and offer suggestions on how to elicit meaningful musician-instrument relationships at the design stage.

2. Background

This section reports the state of the art in the areas relevant to this paper: DMIs and AMI design and the relationship between musicians and their instruments. We also describe cultural probes and conceptual metaphors as methodologies for research investigations.

2.1. DMI and AMI design frameworks

The HCI (Human Computer-Interaction) legacy of the idea of musicians being "users" of DMIs has been recently challenged by Rodger et al. (2020), who point out that musicians should not be considered users, but rather *agents* in musical ecologies. This concept resonates with Brown (2016), who proposed that creative practices can be framed as an agency network, which includes human and non-human agents. When dealing with AMIs, the human agents participating in the musical ecology have further complexities that need to be addressed in the design. An example of addressing these complexities in Special Education Needs settings is discussed by Blatherwick et al. (2017), who suggested a number of aspects to be considered when designing for and with mixed-ability students. Different musicians have different needs of their instruments; thus, it is inherently complex to know what is demanded or expected from the instrument. Waters (2021) points out that an HCI expert and a musician have different views on the demands of the design process or framework. He states that designers approach a problem by breaking it down into component functions that they can solve whereas musicians are focused on deciding or making choices in the moment of doing (music). The way designers create and intend for instruments to be played is often altered or extended by musicians, a phenomenon that has been referred by HCI and NIME scholars as *appropriation* (Dix, 2007; Masu et al., 2016; Zappi and McPherson, 2018). Therefore, assessing the success of an instrument on the basis of its functions or behaviors as the designer knows them to attain predetermined goals is restrictive (Rodger et al., 2020). The variable nature of a musician's activities challenges the idea of a *prototypical user* that is often the implicit subject of design frameworks. Rodger et al. (2020) proposed to use a specification-type framework that tailors various methods to meet the needs of individual agents. The authors suggest that musical instruments do not have a prototypical user; the musician develops effective skills within their environment through a multiplicity of processes afforded by the design and specificities.[2]

This article is focused explicitly on a sub-category of DMIs—Accessible Musical Instruments (AMI). AMIs have evolved over the years from adapting conventional instruments to bespoke instruments and interfaces. Two recent studies, a survey of inclusive instruments (Frid and Ilsar, 2021) and a review of inclusive musical interfaces (Frid, 2019), added important knowledge to this research area. Categories of technologies that apply to AMIs are listed by Larsen et al. (2016) as they discuss the prospect of musical instruments for people with disabilities.[3] Instruments mentioned by the authors include Soundbeam[4], EyeMusic (Hornof and Sato, 2004), EyeGuitar (Vickers et al., 2010), Skoog[5], and TouchTone (Bhat, 2010). In their practice, some AMIs designers work closely with disabled musicians on bespoke instruments. This is the case with Drake Music.[6] Some examples of

2 Rodger et al. (2020) use the term *specificities* to refer to configurations of things or objects created for particular needs of users, agents, and ecologies. In music, these configurations can be involve instruments, musicians-as-agents, and ecological contexts.

3 There is an ongoing debate around the expressions *people with disabilities* or *disabled people* across various public and social spaces. Andrew is confortable in using the expression *people with disabilities* as it well represents his feeling around his disability not being an external factor to him.

4 www.soundbeam.co.uk

5 www.skoogmusic.com

6 www.drakemusic.org

their creations include The Kellycaster[7], a bespoke type of guitar co-designed by John Kelly, a disabled musician who has a self-described *"punk at heart"* approach to music (Harrison et al., 2019), and Charles Matthews. Another example of an interface that has been used for AMIs is that of the MI.Mu Gloves[8], an interface designed to play a sophisticated version of an air guitar, air synth and air drums. There are also current developments that lie somewhere in between bespoke and generic instruments, such as the Touch Chord[9] and Jamboxx.[10] Although these instruments can meet the needs of a large range of users with disabilities, they do require some physical abilities to use each instrument. For instance, the Touch Chord requires the ability to activate sounds by touching a sensor placed on a board and the Jamboxx requires the ability for the disabled musician to be able to produce enough breath to activate it.

With the many varieties of components and interfaces available and being utilized in DMI and AMI design, challenges exist in evaluating bespoke design in instruments. Instruments specifically designed to be inclusive have been highlighted by Lucas et al. (2019). Their methodological approach consisted in using quantitative data and qualitative observations whilst also taking in observations and viewpoints from the participant. This holistic approach covers the many variables around designing with individuals with complex needs around accessibility and function.

2.2. Conceptual metaphors and cultural probes

Conceptual metaphors help explain or describe concepts that contribute toward assessing what elements and functionalities of an interface a user might intuitively relate to. This understanding can therefore assist a designer in forming design concepts. Metaphors have been used to understand embodied interactions, such as how abstract sound concepts like pitch, volume, and tempo might be associated with body movements through a subject's actions (Antle et al., 2008; Bakker et al., 2012). Kim and Maher (2020) discuss how conceptual metaphors inform design decisions when mapping interface elements to function for consumer devices. When considering conceptual metaphors as an approach to participatory design, Wilkie et al. (2013) introduce the connections between an image schema and embodied cognition theories with interaction. In the above examples, participants considered their responses from sets of concepts and items (e.g., words, images) developed by the researchers. Waters (2021) adopted conceptual metaphors in his investigations by referring to the connection between the musician understanding the analog world whilst imagining or conceptualizing that into the digital realm. Waters describes a continuum of inseparable Deleuze-Guattarian's *assemblages* (Deleuze and Guattari, 1988) as a player-instrument-social expectation.

7 www.drakemusic.org/technology/instruments-projects/the-kellycaster

8 www.mimugloves.com

9 www.humaninstruments.co.uk/instruments

10 www.jamboxx.com

Another methodological tool discussed in our work is the cultural probes (Gaver et al., 1999). Cultural probes are design artifacts designed to stimulate reflections on the relationship between humans and technology and, in the DMI space, between musicians and their instruments. Tahıroğlu et al. (2020) used this tool to investigate how technology changed the *mode-of-being* of musical instruments. Instrumentalists familiarize themselves with sensing the unique characteristics of an instrument (De Souza, 2017), and we can consider this relationship as a probe into musical possibilities (Tahıroğlu et al., 2020). The idea that traditional instruments can serve as probes also applies to DMIs. Cultural probes can be employed to deepen the understanding of musicians' experiences to eventually support the design of the DMI to better assist musicians in achieving their goals. As a result, DMIs themselves become probes that we can use to understand our and other musicians' experiences relating to the human condition, the instrument, and the music we wish to make (Tahıroğlu et al., 2020). Interfaces and technical components have also been used as probes to explore specific research questions (Jack et al., 2020; Guidi and McPherson, 2022). Waters (2021) provides a good example of what a probe could help discover from the account of a drummer describing his experience of playing their instrument:

> *"Bodily exploration of movement on top of, across, and within, the interchangeable pathways of the drum kit. Physical restrictions considered, I tend to think of this style of playing as waves of circular phrasing moving above, around, and passing through the kit, the presence of which is felt both in the feet and the arms/hands, as well as the knees, chest, and stomach."*

2.3. Musician-instrument relationship

A musician-instrument relationship is developed via the intimacy that exist between a musician and their instrument as well as the perceptual transparency between the two entities (Nijs et al., 2009). This perceptual transparency, which is acquired with practice, refers to the musician's perception of their instrument as an extension of their body (Rabardel, 1995; Leman, 2007; Morreale et al., 2018). These elements contribute to a sense of flow in their performances as they help provide a sense of control, motivation, and wellbeing contributing to longevity in a musician's application to the instrument (Csikszentmihalyi and Csikszentmihalyi, 1992; Simoens and Tervaniemi, 2013). Simoens and Tervaniemi (2013) quote the violinist Stephen Bryant in an interview *"I don't like the thought of anyone else playing it [his violin]. It's such a close relationship"*. This comment exemplifies the intimate relationship a musician can have with their instrument or their tools. The musician-instrument relationship can thus be similar to that one might have with another person they care deeply for. When playing or performing, this close relationship creates a unity between the instrument and the musician contributing to creative expressions in the interplay between the two (Nijs et al., 2009; Waters, 2021). This unity stems from the studies in embodied music cognition, which provide a framework for studies of the musician-instrument relationship. According to Leman (2007), the human body is considered the natural mediator between the musician's mind and the physical environment that contains musical energy. Considering how a musician might communicate ideas into this physical environment creating and containing musical energy, the musician must first establish a musician-instrument relationship as an entity that players can establish dynamic relations with Jordà (2004).

A musician's subjective experience when performing an instrument thus establishes and strengthens their relationship with it. The experience can be considered from a perspective *"in which the interaction between musician and musical environment, the nature of human activities, and the quality of subjective experience are addressed"* (Nijs et al., 2009). A successful interaction in which the subjective experience is addressed contributes to the feeling that the musical instrument has become part of the body. In other words, a musician experiences an embodied relationship with an instrument. The embodied experience creates entanglements between humans, machines, objects, and environments, including social structures (Mice and McPherson, 2022). The resulting entanglements are part of embodiment theory in how we interact with objects and environments and are intrinsically a part of them as they are a part of us. Entanglement theories specific to HCI are discussed by Frauenberger (2019). The authors suggest that humans are inseparable from the technologies we engage with; a relation that is described by a philosophical concept knows as *relational ontology* (Rosenberger and Verbeek, 2015). A specific area of embodiment theory related to sensor-motor skill is discussed by Guidi and McPherson (2022) in their investigation of skilled musicians using an unfamiliar interface (an augmented guitar pick). The musician's motor skills can be impacted by changes to the interface's physical or sonic characteristics (Morreale et al., 2019) affecting interactions and the embodied experience.

The musician regulates the goal-directed activity structure of music performance, which is inspired and influenced through the musician's subjective experience of the musical ecology during the performance.

Forms of entanglements also apply to interactions between the musician and the musical environment created throughout the performance (Nijs et al., 2009). Partially conflicting with the view of Rodger et al. (2020) on goal-directed activities, Nijs et al. (2009) suggest that the musician regulates the goal-directed activity structure of music performance, which is inspired and influenced by the musician's subjective experience of the musical ecology during the performance. Both elements (goal direction and subjective experience) influence each other and are optimized through an iterative process. These viewpoints are elaborated within the framework of embodied music-cognition (Nijs et al., 2009). This framework is based on Ecological Philosophy, Activity Theory, and Flow and Presence Research. Ecological philosophy describes a musician-instrument relationship in which the musician and their instrument merge and boundaries are no longer experienced. This functional transparency occurs when a musician directly responds to the musical environment and is lost when the instrument is altered (Morreale et al., 2018). Using the activity theory framework, a musician acts as the mediator that establishes an intimate relationship with their instrument (Nijs et al., 2009). When the subject, object, and environment combine, a state of flow is produced, which results in enjoyment, engagement, and increased motivation (Csikszentmihalyi and Csikszentmihalyi, 1992). Achieving this flow state enables a musician to establish a long-term constructive relationship with their instrument. An embodied experience of the instrument becoming an extension of the musician contributes to withdrawing of the instrument from consciousness. The state of flow is achieved through this immersion of the subject and object through any musical activities (Nijs et al., 2009). Each successful experience of flow and transparency contributes to strengthening long-term relationship bonds between a musician and their instrument (Csikszentmihalyi and Csikszentmihalyi, 1992). A positive, satisfying, and constructive relationship between a musician and the instrument relies on not being hindered or obstructed by technical difficulties, performance anxiety, or accessibility in playing an instrument. These obstacles will ultimately affect not only the positive experience of creative practice but interrupt the flow state, which is the ultimate loss of awareness of oneself when performing a task (Nakamura et al., 2002). The hindrance of any obstacle must be reduced to establish and maintain engagement and ultimate embodiment with an instrument. Waters (2021) and Simoens and Tervaniemi (2013) provide examples of how interfaces might interrupt or obstruct a musician or performer from engaging in the process of making music.

Through the processes explained above that lead to embodiment, an experience is established where the musician and the instrument are operating in what Nijs et al. (2009) consider an *instrumental genesis*. The instrumental genesis is a two-fold movement: the instrument constraints the creative possibilities of a musician, who then generates automatic responses within the musical environment. A *relationship of reciprocal affordances* is then established, which integrates the instrument with a musician through the activities, contributing to the feeling of the instrument being a natural extension of the musician within a musical environment.

Music, musical instruments, technology, and social and cultural structures do not develop in isolation. Each plays a role in contributing to the thinking and designing of the day. For example, the valve that was adopted into brass instruments relates to metal tubes and plumbing used at the time (Waters, 2021). Also, the transgression or migration (of musical instruments) from region to region or even rural to urban/city environments affects change in construction and even the playing style of some instruments (Bates, 2012). Bates (2012) suggests that humans are connected with instruments through the multiple functions of instruments (e.g., dance, celebration, experimentation, rebellion) and our relationship with them, from instrument making to performance. These connections impact how our expressions, playing techniques, and instruments develop. As Bates (2012) points out, our embodiment of musical instruments and their sound materialize in how parts of instruments are named and can refer to the human anatomy, such as neck, body, and head. With respect to the technology surrounding a musician's connection and physical relationship with their instrument, there are significant differences between conventional acoustic or electric instruments and DMIs. With acoustic or electric instruments, it can be argued that there is a more transparent connection between the musician and the instrument. The musician faces directly transmitting energy through their body to excite the resonance of an instrument and produce the sound. Jordà (2004) describes this connection as fuzzy and unclear. DMIs, on the other hand, *listen* for gestures performed on an interface and, via mapping, changes sound parameters. Tahıroğlu et al. (2020) suggests that our relationships with DMIs can be *"decoupled from the established relationships we have with more traditional musical instruments."* This suggestion resonates with the notion that coupling is considered important for establishing a successful musician-instrument relationship (Nijs et al., 2009). Our relationships with DMIs and musical norms are shaped and transformed through opportunities (Tahıroğlu et al., 2020). This, in turn, impacts compositions, performances, and musical experiences, and our intentions forge relationships with machine/interface technologies. It is through the creative intent and gestures that a musician uses to connect with the technologies available that establish and determine the strength of the musician's relationship with the instrument. This has been described as a *"dynamic re-formation of gestural and expressive intent"* (Van Nort, 2011). How physical parameters respond to our intentions plays a part in establishing expressive, strong, and meaningful relationships that help create shared authorship, agency, and intentionality between a musician and an instrument. Experiencing or feeling the effects of input devices is proposed as important by Hunt et al. (2000) because interactions with interfaces can determine a musician's experience.

3. Autoethnographic account

In this section, we offer Andrew's autoethnographic account of the evolution of his relationship with his musical instruments. This account provides insights into the evolution of his musician-instrument relationships before and following an injury resulting in a life-changing disability. Autoethnography combines autobiography and ethnography to systematically describe and analyse personal experiences and observations in retrospect (Ellis et al., 2011). Autoethnography offered Andrew the methodological tools to thoroughly observe, challenge, and interrogate his thoughts, beliefs, and assumptions on the evolving relationships with his musical instruments. These experiences, reflections, and analyses produce insights for further discussions and development as they are cross-referenced with existing research and collaborators. Following typical autoethnographic procedures, Andrew will use the first-person singular in the rest of this section.

3.1. Relations with the instrument before the accident

Before the accident, I was able to pursue my creative practice on various instruments. I was primarily a saxophone player but also played flute, clarinet, and piano. I was also an electronic musician, composer, and dedicated improviser. All of these creative activities or practices were undertaken with relative physical ease. I was performing many gigs on saxophone and sometimes piano and keyboards, composing and performing for theater and dance, and running recording sessions in my studio. I ran workshops in free-improvised performances with other musicians and artists from disciplines such as dance, spoken word, moving image, and graphic art. I played a key part in organizing events and took up many opportunities to attend and perform at various festivals around Aotearoa (New Zealand). Alongside these creative activities across different communities, I was completing an Honors degree in Composition. Unfortunately, the accident occurred before I completed my degree. When considering my relationship with my instruments before the accident, what mostly emerges is the deep connection with each instrument I played. This connection is described in the following journal entry, in which I recalled the ritualistic aspect that involved setting up my saxophone:

> "When considering the relationship between an artist and their instrument, I would like to start by considering my relationship with my saxophone before my accident. I fondly remember, when preparing to play either for practice, rehearsal, or performance, the ritual of taking the saxophone from the case and assembling it. This formal process allowed me to connect with the instrument in a way that would not only prepare the instrument to be played but prepare me to play the instrument. As each part of the saxophone was placed together, I would feel as if I was connecting with the instrument. This would be especially apparent when testing the reed on the mouthpiece and strongly felt as I placed my fingers on the keys. As I felt the pressure of the springs of the keys against the pads of my fingers, this sensation of touching the instrument before producing the first sound gave me an extremely strong connection. This is where my connection or response begins in how the relationship between the artist and their instrument can be considered. Similarly, when sitting to play the piano, a shorter ritual would take place. I would feel that the act of approaching and sitting down at the piano, opening the lid if necessary, and then pausing, considering the musical possibilities before placing my hands on the keys was important. Once placing my hands on the keys, there would be a further pause to engage with the sensation of the keys beneath his fingers. Through making this connection with these instruments and forging a strong relationship, it now felt as if we were not separate entities, I was not a user, and the instrument was not merely a tool. Still, we were bound together, ready to produce a creative output. That creative output was music."

The ritual described in the note above shows how the relationship is not best described in terms of musician-tools-music but rather between the (musician-instrument)-music, where musician-instrument constitutes a unique element. The tools needed to create music are intimately bound to the musician. I do not perceive them as separate entities but rather as one entity. Although these rituals may sound romantic, they are extremely fond memories for me and offer insight into what I consider necessary conditions to engage positively with a new DMI. In outlining how a bespoke DMI would be, I need to account for the sensations that take place even before playing it, as well as how those sensations contribute to creating an intimate relationship with the instrument and performing with satisfying action and response.

3.2. Experiments with instruments as a disabled musician

An accident left me paralyzed with tetraplegia from the chest down, dramatically affecting my creative practice. This accident put on hold the completion of my degree. Whilst in my first year of rehabilitation, I had the opportunity to create music for theater and dance. I created this music using my laptop, which had accessible aids[11] I continued composing and producing music for theater and dance, along with getting back to organizing performances and events in Tamaki Makaurau (Auckland). After a few years, I started looking into ways to perform, compose, and produce music. These early engagements with live creative practice mainly involved using my voice, a small percussion instrument that could fit on my lap, and a slide whistle. Although rewarding in some ways, there was always an unpleasant feeling of novelty to performing in this manner as most of these creative activities were undertaken in freely improvised collaborative settings. I eventually completed my Honors degree and joined a Master's degree in Composition and interactive technologies. During my Masters degree, I discovered how to use the laptop as an instrument. These early experiments involved using webcam tracking, voice processing, and mapping of the joystick, mouse, and keyboard. I also experimented with some sensor technology, but due to the complexities around setup, calibration, and reliability, it eventually failed to become a permanent part of my performance setup. In 2019, I started playing the guitar, placing it on a custom-made case that converts to a stand supporting the guitar horizontally across my lap. I processed the guitar signal through a distortion pedal and a volume pedal. With this setup, I often performed in an ensemble with a bass player and a drummer, playing rock-influenced free improvisations. To play the guitar, I use two splints attached to my hands, one for picking and strumming, the other for replacing a slide on the strings. The range of sounds I am able to produce in this way includes slide guitar tones, strum "bar" chords across an open tuning, sustained feedback, and short and sharp percussive tones.

3.2.1. Issues that hinder musician-instrument relationship

Instruments from my early experimentation felt "novelty-like." Although they provided me with the opportunity to creatively reconnect with fellow artists, they failed to create a feeling that was similar to performing on instruments before my accident. Using the laptop as an instrument opened up opportunities that moved away from this feeling of being included as a novelty for participation. However, the lack of haptic/tactile feedback results in feeling disconnected from the creative process and disrupts the intimacy of my relationship with the instrument. To date, the guitar has been the most rewarding instrument in terms of feeling connected with an instrument in a physical or haptic/tactile way. However, the size and the cumbersome setup of the instrument have significant limitations when it comes to musical control and expressivity. Playing single notes is possible, but not with accuracy and speed. As a result of what I can play, collaboration with other artists is possible only in experimental music settings. For instance, I can play chords only as bar chords across the strings and single notes cannot be timed accurately. In brief, I am prevented from performing in more conventional music settings. The lack of connection with the laptop or other computer interfaces tested so far results in a limited relationship with the instruments. By contrast, having easy access to the guitar, and being able to play it with some ease (i.e., a low entry level and then the opportunity to progress with a high ceiling), has alluded to the possibility of feeling the same sense of purpose, accomplishment, and satisfaction I had before my accident. As much as the desire of discovering new ways of playing and considering music is understandable, I still desire to be able to construct conventional musical ideas and expressions. This is part of the relationship that I wish to build with an instrument and will help establish relationships with musicians performing on conventional instruments.

I extensively used designed guidelines from previous design frameworks in identifying my specific aims with the design of my DMI and in creating prototypes (Overholt, 2009; Morreale et al., 2014). What I have not found available are suggestions centered on the elicitation of specific musician-instrument relationships. In other words, these frameworks offered rich insight into the specificities of the instruments but failed to offer insights into how to design for specific musician-instrument relationships. This is the gap I evidence in current frameworks, and I urge more research in DMI and AMI design: the musician-interface relationship should have priorities over specific functionalities.

3.2.2. Consideration on AMI design

When creating bespoke instruments, my experiences before and after the accident kept informing successive design choices. Each of the approaches I have worked with has indeed offered useful takeaways. The accurate and deterministic response of a computer keyboard, the interpretation of gestures through reliable, consistent and intuitive sensors, the ease of control of the mouse joystick, and the physicality of the guitar are all important features I came to appreciate and that I will consider for my future AMIs.

In the list below, I indicate a series of instrument characteristics that I found crucial in establishing a relationship with an instrument. The intention is not to offer a comprehensive list of design guidelines. This quest would be pointless as these features worked for the specificities of my condition. However, I hope these sorts of suggestions could pave the way for future directions in the design of AMIs.

- Enable a physical connection with the instrument. Interfaces can be analog (i.e., keys, strings) or digital (i.e., breath controllers, pressure sensors).
- Create a linear or predictable connection between energy, the energy injected in the instrument, and the energy generated by the instrument.

11 Accessible aids include a special joystick mouse, a small USB keyboard so that both the joystick and keyboard could fit on a tray on my lap. These worked in conjunction with bespoke designed typing splints. I also had "Dragon Dictate" dictation software and a headphone microphone to communicate speech to text for emails, word/text documents, or navigating the internet.

- Ensure an intuitive, natural, and predictable response from the instrument that can generate a visceral connection and symbiotic relationship.
- Consider ease of setup, possibly in a way that any person can assist in setting up the instrument.
- Ensure the accessibility to the instrument: Positioning oneself and the instrument to play or move away or out from the instrument when finished playing should require little to no assistance once the instrument is set up.
- Design for a low entry level so that a musician can start playing the instrument with relative ease while offering the opportunity to progress and develop those skills to a high ceiling for more complex performances.
- Enable accuracy and the repetition of musical ideas so that more intricate musical lines can be created beyond simply improvising sonic textures or effects.

Several of these points have been illustrated in other DMI frameworks. However, addressing them is more complex when designing AMIs. The added complexities come from physical (e.g., limited mobility and reduced tactility) and cognitive impairments (e.g., ability to concentrate, understand tasks, issues with speech and language, behavioral difficulties; Blatherwick et al., 2017).

4. Discussions

Existing DMI and AMI design frameworks primarily focus on technical and goal-based solutions around the environments they investigate. However, the complexities surrounding the subjective aspects of the musician-instrument relationship have been mostly overlooked. Understanding these complexities involves exploring musicians' emotional and physical connections with their instruments. These aspects are undoubtedly more challenging to measure and not readily available compared to technical specifications commonly addressed in design frameworks. This observation is aligned with what Born (2020) called *analytical ontology*. With this term, Born referred to the assumptions often implicit in Music Information Retrieval research and practice about *what music is*. Many music subtleties (mostly of non-Western music), which are difficult to extract and analyse, are simply ignored. We see a similar issue in the design and design frameworks of DMI and AMI.

Scholarly investigations on connectivity and embodiment of musician-instrument relationships, how they contribute to creative practice, and the phenomenology of the experience that goes with them have primarily focused on conventional instruments and non-disabled musicians. In successful musician-instrument relationships where functional transparency takes place (Nijs et al., 2009), key concepts such as flow (Csikszentmihalyi and Csikszentmihalyi, 1992) contribute to the embodiment and entanglements of instruments with the musician. Musician-instrument relationships of disabled musicians seem underrepresented in current research. An opportunity exists for designers and musicians to explore ways to investigate the complexities surrounding the feelings and values within a musician's relationship with an instrument, how they are accessed, and lastly, the technological solutions to be considered.

Andrew's comment on the relation $(musician - instrument) - music$ resonates with the embodied relationship described by Ihde (1990). The American philosopher talks about embodiment needing to be *constituted*, or learnt, for a technology to become transparent. To do so, Ihde explains, the technology *"must be technically capable of being seen through."* We identify a designer's paramount role in the musician-instrument relationship to be constituted in the case of disabled musicians. Thus, the relationship might be best represented as $[(musician - designer) - instrument] - music$. For this relationship to exist, the designer must develop an understanding of the physical limitation and possible technological options and an intimate appreciation of the musician's desires, goals, and expected output. It is not only the physical limitations that significantly differ among disabled musicians but also their intended relationship with the instrument. Creating instruments with generic inputs and outputs that can easily adapt to a wide range of users is problematic for AMI designers due to individual nature of disabilities and impairments.

The question remains, how can these technical solutions best satisfy a musician-instrument relationship? How do designers find the best way to match these solutions to the specific demands of the needs of an individual's musician-instrument relationship? A combination of processes and ecologies, termed *specificities* by Rodger et al. (2020), offers insights into constructing a relationship between the musician and the instrument beyond goals, environments, and solution approaches. The work of Rodger et al. (2020) is focused on an already developed system or prototype to be evaluated rather than investigating a musician-instrument relationship before the design. We argue that considerations about a successful musician-instrument relationship should be given at the conceptual design stage.

One possibility is to integrate conceptual metaphors with cultural probes, technology probes, or research products (Jack et al., 2020). Probes can play a part in discovering more details around the complexities of emotions that establish the musician-instrument relationship. Understanding these complexities might help describe what feelings produce engagement and embodiment of the instrument to fulfill the desire for expressions through a successful design. One example is the conceptual metaphor of asking what acoustic instrument they would wish to play if they could (Waters, 2021). And if so, why that instrument? What are the elements around that instrument that excites them?

A designer could create a probe that cultivates an answer from the musician they are designing with, who could describe how they would like to engage with an instrument. An example could be for a designer to ask the musician to describe, draw, or demonstrate movements that communicate musical intentionalities or sensations and how to physically embody them in the design of the instrument and the gesture-sound mapping. This probe would help inform the musician-instrument relationship and could assist in how the designer approaches further developments. The probe could consider pathways, physical restrictions, phrasing or movements, connections to and with an instrument that would likely be felt or resonate, and what instrument/sonic responses would feel the most natural or interconnected. Conversely, asking a musician to freely associate and describe the complexities around the subjective feelings they are looking for when playing a musical instrument could create a set of probes for any designer to work with. However, do the answers for these types of probes establish an understanding for a designer of the intimate relationship a musician may wish to experience? This is what we are seeking to understand through future research.

One option would be to start with questions that explore the phenomenology surrounding the musicians' desires when wishing to engage with an instrument and experience its response. We propose the following questions to inform the design of the probes:

1. How do they wish for the instrument to "feel," and what psychological or subjective connections and responses would they consider rewarding?
2. How do they intuitively and instinctively understand or control the instruments' responses? What are the physical limitations and opportunities?
3. Which movements (and how do they) feel the most natural, intuitively connected to music/sound production parameters and ultimately rewarding
4. How does creating combinations of these movements intuitively engage with the imagination around making music/sound?
5. What feedback responses are the most effective or appropriate to feel connected to the instruments' responses?
6. How can a link between the psychological and the physical connection(s) and responses lead to "functional transparency"?

The musician can respond to these questions in any way they wish to communicate. Through words, pictures, demonstrations of movement, or communication in any form of media that best suits their interpretation of a set of questions. This could be prescribed from past experiences, imagination, and ideals. For instance, the following quotes describe Andrew's answers to the first question points:

- *"The ability to be intuitively accurate through my responses"*
- *"To feel the music around me, or be moved within a musical environment and be able to place sound from my mind into that environment with as little conscious cognitive thought and physical impedance as possible"*
- *"To create the idea of a sound or note and play accurately without needing to think about the physical tasks required to produce it"*
- *"To have responses from the instrument that feel innately connected to the energy put into it; this means having the opportunity to put energy into the instrument in a haptic/tactile way, and have that represented as a sonic realization or representation directly connected to that input of energy"*

From here, the design can refer to the principles of activity theory[12] as framing to analyse the responses. Starting from these responses, designers would create further probes for investigation and development and move toward constructing a prototype.

12 Activity Theory principles are outlined as: (1) the unity of consciousness and activity, (2) object-orientedness, (3) hierarchical structure of an activity, (4) iternalization and externalization, (5) mediation, and (6) continuous development.

5. Conclusion

Existing design frameworks for DMI largely focus on technological solutions and goal-based approaches while overlooking the subjective aspects of the musician-instrument relationship. Additionally, research on the embodiment and connectivity of musician-instrument relationships in creative practice has primarily focused on conventional instruments and non-disabled musicians, leaving the experiences of disabled musicians underrepresented. In this article, we proposed that DMIs and AMIs design strategies should extend their scope of investigation to also account for the connections that musicians might have with their instruments, which are difficult to measure and not typically addressed in design frameworks. Notably, the intention of this article was not to provide a comprehensive framework. This article is intended to surface fundamental points that are needed to be tackled when designing for and with disabled musicians. These points can be integrated as complementary features in existing DMI and AMI design frameworks. Specifically, we proposed that musician-instrument relationships should be considered in the design of AMIs at an early stage before considering goals and technological solutions. We also indicated methods to surface the intended musician-instrument relationship. We discussed the possibility of borrowing conceptual metaphors and cultural probes as tools to be used within a DMI and AMI framework to account for the subjective experience and the ecology and specificity that determine the disabled musician's desires for a musician-instrument relationship.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

AM ideated the project and presented the autoethnographic account. FM offered theoretical support on DMI/AMI and guidance throughout the project. Both authors wrote the article. Both authors contributed to the article and approved the submitted version.

Funding

This work was funded by University of Auckland CAI RDF grant (9483-3728468) and University of Auckland PREeSS grant (312820-9441-AMCM028).

Conflict of interest

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

Publisher's note

References

Antle, A. N., Droumeva, M., and Corness, G. (2008). "Playing with the sound maker: do embodied metaphors help children learn?," in *Proceedings of the 7th International Conference on Interaction Design and Children* (Chigago, IL), 178–185. doi: 10.1145/1463689.1463754

Bakker, S., Antle, A. N., and Van Den Hoven, E. (2012). Embodied metaphors in tangible interaction design. *Pers. Ubiquit. Comput.* 16, 433–449. doi: 10.1007/s00779-011-0410-4

Bates, E. (2012). The social life of musical instruments. *Ethnomusicology* 56, 363–395. doi: 10.5406/ethnomusicology.56.3.0363

Bhat, S. (2010). "Touchtone: an electronic musical instrument for children with hemiplegic cerebral palsy," in *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction* (Cambridge, MA), 305–306. doi: 10.1145/1709886.1709955

Blatherwick, A., Woodbury, L., and Davis, T. (2017). "Design considerations for instruments for users with complex needs in SEN settings," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Copenhagen: Aalborg University Copenhagen), 216–221.

Born, G. (2020). Diversifying MIR: knowledge and real-world challenges, and new interdisciplinary futures. *Trans. Int. Soc. Mus. Inform. Retriev.* 3, 193–204. doi: 10.5334/tismir.58

Brown, A. R. (2016). "Understanding musical practices as agency networks," in *Proceedings of the Seventh International Conference on Computational Creativity*.

Csikszentmihalyi, M., and Csikszentmihalyi, I. S. (1992). *Optimal Experience: Psychological Studies of Flow in Consciousness*. Cambridge: Cambridge University Press.

De Souza, J. (2017). *Music At Hand: Instruments, Bodies, and Cognition*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780190271114.001.0001

Deleuze, G., and Guattari, F. (1988). *A Thousand Plateaus: Capitalism and Schizophrenia*. London: Bloomsbury Publishing.

Dix, A. (2007). "Designing for appropriation," in *Proceedings of HCI 2007 The 21st British HCI Group Annual Conference University of Lancaster* (Lancaster), 1–4. doi: 10.14236/ewic/HCI2007.53

Ellis, C., Adams, T. E., and Bochner, A. P. (2011). Autoethnography: an overview. *Histor. Soc. Res.* 273–290. doi: 10.17169/fqs-12.1.1589

Frauenberger, C. (2019). Entanglement hci the next wave? *ACM Trans. Comput. Hum. Interact.* 27, 1–27. doi: 10.1145/3364998

Frid, E. (2019). Accessible digital musical instruments—a review of musical interfaces in inclusive music practice. *Multimodal Technol. Interact.* 3:57. doi: 10.3390/mti3030057

Frid, E., and Ilsar, A. (2021). "Reimagining (accessible) digital musical instruments: a survey on electronic music-making tools," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Shanghai). doi: 10.21428/92fbeb44.c37a2370

Gaver, B., Dunne, T., and Pacenti, E. (1999). Design: cultural probes. *Interactions* 6, 21–29. doi: 10.1145/291224.291235

Guidi, A., and McPherson, A. (2022). "Quantitative evaluation of aspects of embodiment in new digital musical instruments," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Auckland: The University of Auckland). doi: 10.21428/92fbeb44.79d0b38f

Harrison, J., Chamberlain, A., and McPherson, A. P. (2019). "Accessible instruments in the wild: engaging with a community of learning-disabled musicians," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow), 1–6. doi: 10.1145/3290607.3313037

Hornof, A., and Sato, L. (2004). "Eyemusic: making music with the eyes," in *Proceedings of the 2004 Conference on New Interfaces for Musical Expression* (Hamamatsu), 185–188.

Hunt, A., Wanderley, M. M., and Kirk, R. (2000). "Towards a model for instrumental mapping in expert musical interaction," in *ICMC* (Berlin).

Ihde, D. (1990). *Technology and the Lifeworld: From Garden to Earth*. Bloomington, IN: Indiana University Press.

Jack, R., Harrison, J., and McPherson, A. (2020). "Digital musical instruments as research products," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, eds R. Michon and F. Schroeder (Birmingham: Birmingham City University), 446–451.

Jensenius, A. R., and Lyons, M. J. (eds.) (2017). *A NIME Reader: Fifteen Years of New Interfaces*. Berlin: Springer. doi: 10.1007/978-3-319-47214-0

Johnston, A., Candy, L., and Edmonds, E. (2008). Designing and evaluating virtual musical instruments: facilitating conversational user interaction. *Design Stud.* 29, 556–571. doi: 10.1016/j.destud.2008.07.005

Jordà, S. (2004). Instruments and players: Some thoughts on digital lutherie. *J. New Mus. Res.* 33, 321–341. doi: 10.1080/0929821042000317886

Kim, J., and Maher, M. L. (2020). Conceptual metaphors for designing smart environments: device, robot, and friend. *Front. Psychol.* 11:198. doi: 10.3389/fpsyg.2020.00198

Larsen, J. V., Overholt, D., and Moeslund, T. B. (2016). "The prospects of musical instruments for people with physical disabilities," in *NIME, Vol. 16* (Brisbane, QLD), 327–331.

Leman, M. (2007). *Embodied Music Cognition and Mediation Technology*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/7476.001.0001

Lucas, A. M., Ortiz, M., and Schroeder, D. F. (2019). "Bespoke design for inclusive music: the challenges of evaluation," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, eds M. Queiroz and A. X. Sedó (Porto Alegre: UFRGS), 105–109.

Masu, R., Conci, A., Menestrina, Z., Morreale, F., and De Angeli, A. (2016). "Beatfield: an open-meaning audiovisual exploration," in *COOP 2016: Proceedings of the 12th International Conference on the Design of Cooperative Systems* (Trento: Springer), 309–314. doi: 10.1007/978-3-319-33464-6_20

Mice, L., and McPherson, A. P. (2022). "Super size me: Interface size, identity and embodiment in digital musical instrument design," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA), 1–15. doi: 10.1145/3491102.3517626

Morreale, F., Armitage, J., and McPherson, A. (2018). Effect of instrument structure alterations on violin performance. *Front. Psychol.* 9:2436. doi: 10.3389/fpsyg.2018.02436

Morreale, F., De Angeli, A., and O'Modhrain, S. (2014). "Musical interface design: an experience-oriented framework," in *NIME* (London), 467–472.

Morreale, F., Guidi, A., and McPherson, A. P. (2019). "Magpick: an augmented guitar pick for nuanced control," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, eds M. Queiroz and A. X. Sedó (Porto Alegre: UFRGS), 65–70.

Nakamura, J., Csikszentmihalyi, M., et al. (2002). The concept of flow. *Handb. Posit. Psychol.* 89:105.

Nijs, L., Lesaffre, M., and Leman, M. (2009). "The musical instrument as a natural extension of the musician," in *The 5th Conference of Interdisciplinary Musicology* (Paris: LAM-Institut jean Le Rond d'Alembert), 132–133.

Overholt, D. (2009). The musical interface technology design space. *Organised Sound* 14, 217–226. doi: 10.1017/S1355771809000326

Rabardel, P. (1995). *Les Hommes et les Technologies; Approche Cognitive des Instruments Contemporains*. Paris: Armand colin.

Rodger, M., Stapleton, P., Van Walstijn, M., Ortiz, M., and Pardue, L. (2020). "What makes a good musical instrument? A matter of processes, ecologies and specificities," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Birmingham: Birmingham City University), 405–410.

Rosenberger, R., and Verbeek, P.-P. (2015). "A field guide to postphenomenology," in *Postphenomenological Investigations: Essays on Human-Technology Relations* (Lanham, MD), 9–41.

Simoens, V. L., and Tervaniemi, M. (2013). Musician-instrument relationship as a candidate index for professional well-being in musicians. *Psychol. Aesthet. Creativ. Arts* 7:171. doi: 10.1037/a0030164

Tahıroğlu, K., Magnusson, T., Parkinson, A., Garrelfs, I., and Tanaka, A. (2020). Digital musical instruments as probes: how computation changes the mode-of-being of musical instruments. *Organised Sound* 25, 64–74. doi: 10.1017/S1355771819000475

Van Nort, D. (2011). Human: machine: human: gesture, sound and embodiment. *Kybernetes* 40, 1179–1188. doi: 10.1108/03684921111160403

Vickers, S., Istance, H., and Smalley, M. (2010). "Eyeguitar: making rhythm based music video games accessible using only eye movements," in *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology* (Taipei), 36–39. doi: 10.1145/1971630.1971641

Waters, S. (2021). The entanglements which make instruments musical: rediscovering sociality. *J. New Music Res.* 50, 133–146. doi: 10.1080/09298215.2021.1899247

Wilkie, K., Holland, S., and Mulholland, P. (2013). Towards a participatory approach for interaction design based on conceptual metaphor theory: a case study from music interaction. *Music Hum. Comput. Interact*. 259–270. doi: 10.1007/978-1-4471-2990-5_15

Zappi, V., and McPherson, A. (2018). Hackable instruments: supporting appropriation and modification in digital musical interaction. *Front. ICT* 5:26. doi: 10.3389/fict.2018.00026

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 June 2023
DOI 10.3389/fpsyg.2023.1158172

Check for updates

OPEN ACCESS

EDITED BY
Lei Wang,
Tongji University, China

REVIEWED BY
Bruno Gingras,
University of Vienna, Austria
Chen Ji Shang,
Xinjiang University of Finance and Economics, China

*CORRESPONDENCE
Kat R. Agres
✉ katagres@nus.edu.sg

[†]These authors have contributed equally to this work and share first authorship

RECEIVED 03 February 2023
ACCEPTED 05 May 2023
PUBLISHED 06 June 2023

CITATION
Agres KR, Dash A and Chua P (2023) AffectMachine-Classical: a novel system for generating affective classical music. *Front. Psychol.* 14:1158172.
doi: 10.3389/fpsyg.2023.1158172

AffectMachine-Classical: a novel system for generating affective classical music

Kat R. Agres[1,2*†], Adyasha Dash[1†] and Phoebe Chua[3]

[1]Yong Siew Toh Conservatory of Music, National University of Singapore, Singapore, Singapore, [2]Centre for Music and Health, National University of Singapore, Singapore, Singapore, [3]Augmented Human Lab, DISA, National University of Singapore, Singapore, Singapore

This work introduces a new music generation system, called AffectMachine-Classical, that is capable of generating affective Classic music in real-time. AffectMachine was designed to be incorporated into biofeedback systems (such as brain-computer-interfaces) to help users become aware of, and ultimately mediate, their own dynamic affective states. That is, this system was developed for music-based MedTech to support real-time emotion self-regulation in users. We provide an overview of the rule-based, probabilistic system architecture, describing the main aspects of the system and how they are novel. We then present the results of a listener study that was conducted to validate the ability of the system to reliably convey target emotions to listeners. The findings indicate that AffectMachine-Classical is very effective in communicating various levels of Arousal ($R^2 = 0.96$) to listeners, and is also quite convincing in terms of Valence ($R^2 = 0.90$). Future work will embed AffectMachine-Classical into biofeedback systems, to leverage the efficacy of the affective music for emotional wellbeing in listeners.

KEYWORDS

automatic music generation system, algorithmic composition, music MedTech, emotion regulation, listener validation study, affective computing

1. Introduction

There is now overwhelming evidence that music supports health and well-being in various ways, from motivating physical activity, to promoting mental health and fostering social connection (MacDonald et al., 2013; Fancourt and Finn, 2019). Music is particularly effective for supporting and mediating emotion states. Indeed, one of the primary reasons people report listening to music is to change or enhance their emotions (Thayer et al., 1994; Lonsdale and North, 2011; Saarikallio, 2012). Given the affordances of music to support health and wellbeing, as well as advances in machine learning and computational techniques, there has recently been a call to action to compose music with the use of computational technologies for healthcare applications (Agres et al., 2021). Compared to the use of human-composed, pre-recorded music, which spans many genres and emotions but is fixed and difficult to adjust in real-time, generative music composition systems are able to support real-time interactivity—they are able to flexibly manipulate musical features almost instantaneously according to the listener's current neural or physiological state, or given their real-time input and preferences. These systems therefore show promise in delivering personalized, cost-effective (and free of copyright), non-invasive, non-pharmaceutical methods for helping individuals improve their emotion states. Given the global mental health crisis (e.g, nearly 20% of adults in the USA live with a mental illness, or 52.9 million

Americans in 2020[1]), music medtech systems are projected to be extremely valuable tools for supporting emotional wellness, and mental health more broadly.

More generally, in an age where computational systems are now being used extensively to generate impressive natural language and visual art, such as the technologies available through OpenAI,[2] it is no surprise that there has been a recent surge of interest in the development of automatic music generation systems (AMGSs; also known as algorithmic composition systems). Like human music composition and improvisation, AMGSs generally aim to create harmonic, timbral, and rhythmic sequences in an organized, musically-coherent fashion. This area, which sits at the intersection of computing, music theory/composition, and computational creativity, is relatively nascent, however, compared to the computational creation of visual art. This work aims to not only chip away at this gap, but offer a new automatic music generation system—AffectMachine-Classical—that is capable of producing controllable *affective* music. AffectMachine-Classical offers an effective and flexible means of conveying emotions in real-time, and the system has been developed to be embedded into biofeedback systems such as brain-computer-interfaces (see for example Ehrlich et al., 2019), making it a potentially powerful tool for therapeutic applications such as emotion and mood regulation in listeners, augmentation of physical activity during rehabilitation, as well as commercial use cases such as soundtrack design and providing silent videos with novel music free of copyright.

1.1. Related work

A review of automatic music generation is out of the scope of this article (for a review and summary of the state-of-the-art, see Herremans et al., 2017; Carnovalini and Rodà, 2020; Dash and Agres, 2023), however we will briefly summarize the main approaches to automatic music generation. Previous approaches to developing music generation systems largely fall into two categories: learning-based methods, and rule-based methods. While there has been a recent trend toward learning-based approaches, they present several challenges for affective music generation. First, ecological (or realistic) music pieces typically exhibit hierarchical, long-term structure, as well as polyphony. For example, a melodic phrase may extend over multiple measures of music, and involve several different instruments or voices. Further, music typically has an overall form that allows for musical and stylistic coherence throughout the piece. As such, a generative model must be able to capture harmonic, rhythmic and temporal structure, as well as the interdependency between voices (Dong et al., 2018). Second, learning-based approaches require large music datasets with emotion labels for training, a resource that is still scarce in the community, although we note that acoustic models that are able to link musical excerpts directly to natural language descriptions are beginning to emerge (Huang et al., 2022), and may be a promising direction for future work. Style transfer models have had success as an alternative to models capable of generating novel affective music from scratch—for example, Ramirez et al. (2015) used machine learning models to apply appropriate expressive transformations on the timing and loudness of pre-composed input musical pieces based on desired arousal and valence. In addition, Williams et al. (2017) used affective feature mappings to transform seed material generated by a neural network trained on short musical excerpts, and Hu et al. (2020) used convolutional neural networks to extract stylistic features from therapeutic music pieces and incorporate them into user-selected songs. Despite these promising applications, style transfer models and similar approaches are subject to several important limitations. For example, although leveraging pre-composed music greatly simplifies the challenge of producing affective music, the pre-existing music chosen is subject to copyright. Recent progress in conditional music generation from text has resulted in models that are able to generate high-fidelity music based on natural language descriptions (Agostinelli et al., 2023), which may potentially sidestep copyright issues. However, style transfer models and generative models are not yet able to support flexible and continuous generation for real-time interactivity, which is essential in biofeedback systems or any other systems meant to compose music in real-time to mediate the user's affective states.

In comparison to learning-based approaches, rule-based approaches rely on hand-designed functions to map affective signals to musical parameters. As such, they are able to sidestep the challenges associated with learning-based approaches by building in knowledge of how affective states map to musical parameters, as well as typical expectations regarding harmonic, rhythmic, and temporal structure. Additionally, the design of rule-based affective music generation systems benefits from an extensive body of theoretical and empirical work going back almost a century that investigates how different aspects of musical structure contribute to emotional expression (Gabrielsson and Lindström, 2010). For example, the system described in Wallis et al. (2011) was primarily informed by Gabrielsson and Lindström (2001), and generates novel music algorithmically by mapping seven musical parameters (e.g., note density, musical mode) to either valence or arousal in the most continuous possible way. Even though several salient musical parameters such as tempo, voice leading, and voice spacing, were not mapped for simplicity, the system was sufficient for participants to hear corresponding changes in the emotion of the music when changes were applied to the valence and arousal parameter settings. Similarly, the adaptive music engine described in Gungormusler et al. (2015) manipulates musical parameters including tempo, articulation, and timbre based on empirical validation of music-emotion structural rules carried out by Livingstone and Brown (2005). Most recently, Ehrlich et al. (2019) developed a system that loops over a I-IV-V-I harmonic progression, and modifies the musical mode, tempo, rhythmic roughness (a measure of the amount of variation in note lengths within a measure), overall pitch, and relative loudness of subsequent notes based on the desired level of valence and arousal. Their listening study confirmed a high correspondence between the system's arousal and valence settings and the emotions listeners perceived (Ehrlich et al., 2019).

1 Statistics from the National Institute of Mental Health: https://www.nimh.nih.gov/health/statistics/mental-illness.

2 https://openai.com/

Compared to existing rule-based systems, AffectMachine is more sophisticated, by taking into account traditional features, such as tempo, rhythmic roughness/note density, mode, etc., as well as additional features such as voice leading and a fine-grained mapping between valence/arousal and musical features such as the chord progression. In addition, because our system is capable of producing music in real-time based on given arousal and valence, it has a flexibility not exhibited by most other music generation systems.

1.2. Emotion perception in music

Music listening is often a rich emotional and cognitive experience (Altenmüller and Schlaug, 2012), and numerous studies have explored the relationship between music and emotional expression. For example empirical studies have been carried out to better understand both the emotions that can be expressed through music (e.g., Gabrielsson and Juslin, 2003), as well as the musical factors that contribute to perceived emotional expression (e.g., Gabrielsson and Lindström, 2010). Research has shown that various musical cues, such as tempo, mode, dynamics, pitch range, rhythm, and articulation, can influence the perceived emotion in music (Gabrielsson and Juslin, 2003; Schubert, 2004; Juslin and Västfjäll, 2008; Juslin and Sloboda, 2013). For example, studies have found that fast tempos are associated with positive emotions such as joy and excitement, while slow tempos are associated with negative emotions such as sadness and melancholy (Juslin and Laukka, 2003). Similarly, major modes are generally associated with positive emotions, while minor modes are associated with negative emotions (Juslin and Laukka, 2003), although this can depend on musical enculturation (Swaminathan and Schellenberg, 2015). Other musical cues, such as dynamics/loudness and pitch range, can also influence the perceived emotion in music. For instance, loudness has been found to correlate strongly with perceived and induced arousal, while high pitch ranges are associated with excitement and low pitch ranges with sadness (Balkwill and Thompson, 1999; Swaminathan and Schellenberg, 2015). Overall, these findings suggest that musical features play a crucial role in influencing perceived emotion in music. The connection between musical features and emotion has also led to a surge of research in Music Information retrieval (MIR) which aims to identify the high-level emotions of music from its low-level features (see, for example, Yang et al., 2018), an area often referred to as music emotion recognition.

Studies have found that listeners tend to exhibit agreement in their judgment of the general emotions expressed by a piece of music, and that these judgments are only marginally affected by demographic factors such as musical training, age, and gender (Juslin and Laukka, 2004), although differences in emotion perception have been emerging in recent work examining the impact of factors such as age and musical training (Cohrdes et al., 2020; Koh et al., 2023). In addition, music is often unable to reliably communicate finely differentiated emotions (Juslin, 1997). Sloboda (2004) offers an explanation for this phenomenon, suggesting that music is to a large extent abstract and ambiguous, and while it may be able to suggest varying levels of energy or resemble certain gestures and actions, these emotional contours are often fleshed out in a subjective manner.

Other recent studies have explored the neural mechanisms underlying emotional responses to music, with a particular focus on the role of the brain's reward system. For example, Salimpoor et al. (2015) describes how listening to music activates the brain's reward system, leading to the release of dopamine and other neurotransmitters associated with pleasure and reward. This suggests that our emotional responses to music are not simply a matter of subjective experience, but are also rooted in the underlying biology of the brain, e.g., dopamine is released in concert with prediction mechanisms in the brain during music listening (Huron, 2008; Salimpoor et al., 2015; Ferreri et al., 2019). Overall, these and other recent studies continue to deepen our understanding of the complex relationship between music and emotion, and suggest that systems able to flexibly manipulate musical features have great potential for emotion-focused well-being applications such as affective music generation systems.

Taken together, the literature suggests that (i) to a large extent, music can be systematically modified to express desired emotions, and that (ii) the effectiveness of affective music generation systems should be fairly robust across listeners.

1.3. AffectMachine-Classical

The current music generation system, AffectMachine-Classical, uses a probabilistic, rule-based approach to generate affective classical music in real time. The system was developed with the help of a classically-trained composer finishing his studies at a major Conservatory of Music, and the system's generated music generally aims to follow the stylistic conventions of Western tonal classical music[3].

Various approaches have been used to measure and describe the affective qualities of musical stimuli, ranging from widely used measures such as Russell (1980)'s circumplex model and the Geneva Emotional Music Scale (GEMS; Zentner et al., 2008) to bespoke methods developed for specific studies (e.g., Costa et al., 2000; Lindström, 2006). Following much of the existing work on affective music systems (e.g., Wallis et al., 2011; Ehrlich et al., 2019), we opted to represent emotion in AffectMachine using the circumplex model, in which emotions can be understood as points within a two-dimensional space. The first dimension is arousal, which captures the intensity, energy, or "activation" of the emotion, while the second is valence, which captures the degree of pleasantness. For example, excitement is associated with high arousal and high valence, while contentment would be associated with low arousal and high valence. The circumplex model has several advantages over alternative measures of emotion. Firstly, to provide accurate and fine-grained feedback to a user about his or her emotional state, music generated by AffectMachine should ideally vary smoothly over the entire space of emotions, making continuous models of emotion such as the circumplex model a natural choice over

3 A separate version of AffectMachine is being developed and tested in a popular-music genre, to afford listeners some variety and choice based on their musical preferences.

categorical models of emotion such as GEMS. Secondly, allowing musical features to change gradually over time could help lend the music a more natural sound. Finally, the generalizability of the circumplex model also enables us to make use of previous research which may have used less common measures of emotion, by interpreting their results in terms of arousal and valence.

AffectMachine provides a model that is able to fluidly generate affective music in real time, either based on manually-input or predetermined arousal and valence values (e.g., as a sort of affective playlist for emotion mediation, or trajectory through emotion space), or based on the real-time feedback or physiological state of the user (e.g., EEG activity captured from the user and mapped to arousal and valence). In this way, AffectMachine offers a flexible yet powerful way to sonify (real-time) emotion states, and to influence the emotion states of the listener. The system may be used for health and wellness applications, such as generating affective playlists for emotion mediation. Further, AffectMachine may also be integrated into Brain-Computer Interface (BCIs) devices, or other systems capable of providing biofeedback, to assist the user in achieving a desired emotion state through neuro/biofeedback and affective music listening.

The main contributions of this work are: (1) the design of a novel rule-based affective music generation system to compose non-monotonic classical music, and (2) validation of the proposed system for expressing different emotions through a listener study. In the next section of this paper, we describe the features of AffectMachine-Classical (Section 2). We then describe the listener study and discuss the findings and implications of our results (Section 3), before providing our general conclusions and suggested future directions (Section 4).

2. AffectMachine-Classical system description

In this section, we describe the parameters and design of our novel affective music generation system, AffectMachine-Classical, which produces affective music in a classical style. AffectMachine was developed to be embedded in a BCI or neurofeedback system, to both generate emotion-inducing music in real-time, and to allow for neural or physiological signals (such as EEG) to *drive* the music generation system. That is, the system was developed to both induce emotion in listeners, and provide users with real-time feedback on their current emotional state, in which the generated music is a reflection (or sonification) of the listeners' emotion state (when AffectMachine is embedded in a BCI or neurofeedback system). In the present paper, we remove AffectMachine from any embedded, interactive contexts (e.g., BCI), and examine the standalone AffectMachine, focusing on the efficacy of AffectMachine for generating music that conveys the intended emotion.

The automatic music generation system was developed in Python, and takes a sequence of arousal and valence states as input and encodes a corresponding sequence of harmonic, rhythmic, and timbral parameters in the form of a MIDI event stream as output. The MIDI event stream is then sent to a digital audio workstation (DAW) over virtual MIDI buses to be translated into sound. For the present version of AffectMachine, we use the Ableton DAW for its wide selection of instruments and its ability to support live multi-track recording. Arousal and valence are continuous values within the range [0, 1] that can either be sampled from sensors (such as EEG) or manually provided. All musical parameters are updated each bar in accordance with the current arousal and valence values.

Developing a rule-based affective music generation system requires first identifying a set of musical parameters and affective states, then designing functions that map parameter values to target states. For this reason, the harmonic, rhythmic, and timbral parameters were selected based on previous work establishing their influence on musical expression of emotions, and developed in collaboration with conservatory students formally trained in music composition.

In the subsections below, we present the details of the AffectMachine-Classical system.

2.1. Harmonic parameters

Previous rule-based music generation systems have controlled the mode parameter by choosing a fixed harmonic progression (e.g., I-IV-V-I) and in a few cases, by varying the musical mode from which the chords are drawn (e.g., each musical mode was mapped to a certain level of valence), with Lydian typically identified as the mode that expresses the highest valence, and Locrian or Phrygian as the mode that expresses the lowest valence, as per Schmuckler (1989). A simpler, and much more common, version of this logic is to switch between the major and minor modes.

In the AMG system, we introduce a completely novel way of controlling mode by using a bespoke probabilistic chord progression matrix inspired by the theme and variation form found in (human-composed) classical music. The music loops through an 8-bar theme with fixed chord functions for each bar, but the specific chords used, as well as their probabilities, are determined by the target level of valence desired. To our knowledge, this approach has never before been implemented in a computational music generation system. The chord set available for each level of valence was based on previous empirical work, as well as the musical insights from conservatory students formally trained in music composition. Previous empirical work has established that valence tends to be positively related to the major mode, and negatively related to the presence of dissonance (e.g., diminished and augmented intervals; Costa et al., 2004; Costa and Nese, 2020). Generally, the chords progressions in our system exhibit greater dissonance with higher probability as valence decreases. Arousal had no influence on the chord progression selected.

Unlike previous systems which are constrained to a specific harmonic progression, the AMG system is extremely flexible—the only constraint being that the music has to progress through the 8-bar theme. (Note that the majority of human-composed music also adheres to a repeating X-bar structure). This novel approach is therefore beneficial by allowing a greater range of musical possibilities (and "interestingness" of the composition). At the same time, the music is able to achieve greater coherence of musical structure than what is commonly found in machine learning-based approaches by using chord substitutions in an 8-bar theme to express the desired level of valence, and by ending each iteration of the theme with a cadence.

To craft the 8-bar theme, the valence range was divided into 10 regions, with one probabilistic chord progression composed for each region to match the intended level of valence. For example, at higher levels of valence, the chord progressions are composed in the major mode as it is typically associated with expressions of positive valence. As valence decreases, the likelihood of chords with greater tension or dissonance (such as those with diminished or minor intervals) increases. For a given bar (e.g., 1–8) and level of valence (e.g., 0–1), there are a set of possible chords, each with a particular probability of occurrence from 0.1 to 0.8. At any given bar and valence level, there are typically multiple chords (between one and five) to choose from.

2.2. Pitch characteristics of voices

2.2.1. Voice leading

Voice leading refers to the art of creating perceptually independent musical lines (e.g., tenor line, soprano line, etc.) that combine to form a coherent piece (Huron, 2001), and is a steadfast component of the majority of human-composed polyphonic music. Despite the importance of voice-leading, automatic generation of polyphonic music with multiple voices or tracks is a challenge that research is only just beginning to address, primarily with learning-based generative methods (e.g., Dong et al., 2018), and many of these systems either fail to address voice leading altogether or use highly simplified versions of voice leading.

In the AffectMachine system, we implement a novel rule-based music generation system that draws on both traditional rules of voice leading as well as heuristics used by human musicians, to create pieces that exhibit perceptually independent musical lines with nontrivial complexity and variability. (Note that in our system, we utilize and refer to voices, not in the strict sense of counterpoint, but similar to the use of voices in a string quartet, where one voice or instrument is capable of playing a chord.) By mapping these rules to differing levels of arousal and valence, we also provide more cues for listeners to identify the emotion being conveyed by the music, and enable finer-grained control over the mapping between affective states and musical parameters. This is an extremely important aspect and benefit of our approach.

AffectMachine-Classical was developed to generate music with four parts or voices. The four voices/algorithms we employ were selected to fill out the acoustic space from low bass frequencies to the higher soprano range. While the instruments do not map strictly to the counterpoint definition of voices (e.g., with independent bass, tenor, alto, and soprano lines), they do span the frequency spectrum from low to high, and work together to convey a cohesive melody and harmony. The bass voice is carried by the string section, and always plays the root note of the current chord. The principal melody is played by the soprano voice, which is carried by the clarinet and doubled at higher valence settings by the marimba. Both inner voices are carried by the piano, with the tenor voice playing a full chord voicing in the middle register, and the alto voice providing harmonic accompaniment by means of single notes adhering to voice leading principles (the details are described below). Instrumentation is explained in more detail in the section on timbral parameters.

While there are numerous principles that govern voice leading, or the creation of perceptually independent parts (Huron, 2001), we select several straightforward rules that provide sufficient melodic diversity while minimizing unpleasant or artificial-sounding melodic lines. The three parts that are determined through voice leading logic are the tenor, alto, and soprano voices. For the principal melody, our primary goal was to avoid unexpected dissonance. Hence, the note sequence is a randomly selected sequence of chord tones. For the tenor voice, which plays the full chord voicing of the chord progression, we follow the heuristic outlined in Wallis et al. (2011)—that pianists tend to voice new chords in a manner that is as similar as possible to the previous chord, in terms of interval and placement on the keyboard. We calculate dissimilarity between two notesets (N, N') as per Equation (1) and select the least dissimilar chord voicing to be played the first inner voice.

$$dissimiliarity = \sum_i \sum_j \mid N_i - N'_j \mid \forall i \in N, \forall j \in N' \qquad (1)$$

The alto voice is monophonic, playing one note at a time according to a step motion rule, where the initial note is a randomly selected chord tone. This rule states that if the next note in the *melody* is of a different pitch, the pitch motion of the alto voice should be by diatonic step (e.g., move up or down the diatonic scale). These rules are encoded in the form of transition matrices. There are four possible states: −1, indicating a diatonic step down the scale; 1, indicating a diatonic step up the scale; 0, indicating no pitch motion; and CT, indicating a jump to a randomly selected chord tone (CT). The arousal range was divided into two equal regions, with one matrix composed for each region to generate appropriate melodies for each level of arousal. The transition matrices were developed such that at higher levels of arousal, melodies are more likely to consist of scale patterns, mitigating the risk of the music being too dissonant or unpleasant due to the increased tempo and note density. Note that our system does not directly avoid parallel fifths/octaves (due to the complexity of the system and the presence of many features), but because this is a probabilistic system, movements of fifths in multiple voices at the same time are relatively rare.

2.2.2. Pitch register

Research in the psychology of music has associated pitch height and pitch register with emotional expression for almost a century (Hevner, 1937); yet pitch height is often not explicitly incorporated into automatic music generation systems. Higher pitches generally tend to be associated with positively-valenced emotions such as excitement and serenity (Collier and Hubbard, 1998), while lower pitches tend to be associated with negatively-valenced emotions such as sadness.

In AffectMachine-Classical, the pitch register of the lowest voice is consistent (at C3). For the remaining voices, the pitch register can vary within a permissible range determined by the current valence level.

To implement changes in pitch register, we again divided the valence range into ten equally spaced regions and tuned the lower and upper bounds of allowable pitches by ear. Both the lower and upper bounds of the range of permissible pitches increase gradually as valence increases. The range of permissible pitches starts at [C1, C5] in the lowest valence region, and gradually moves to [G3, C6] in the highest valence region.

2.3. Time and rhythm parameters

2.3.1. Rhythm

In most automatic generation approaches, the rhythmic content of the music is either fixed (e.g., a repeating pattern or a pre-composed rhythm template is used), or the temporal duration of notes (the rhythmic content) is based on a machine-learning generative process that affords little musical cohesion. This tends to either make the music sound extremely repetitive, or rather incoherent and unpleasant for most listeners.

To surmount this issue, the different voices/parts/tracks in AffectMachine-Classical use different rhythmic logic, e.g., one voice uses probabilistic rhythms while another uses composed rhythms. In this way, our new approach finds a nice and aesthetically-pleasant balance between composed and probabilistic elements.

As mentioned above, AffectMachine-Classical was developed to generate music using four parts or voices. The bass voice (string section) and first tenor voice (piano) employ a fixed rhythmic pattern—they are both played on the first beat of each bar. For the soprano voice (clarinet and marimba), we divided the arousal range into three regions: low ($Arousal < 0.4$), moderate ($0.4 \geq Arousal < 0.75$) and high ($Arousal > 0.75$). Much like the implementation of mode, for a given bar (e.g., 1–8) and arousal region, there is a set of two possible rhythmic patterns or "licks" with equal probability of occurrence. The rhythmic pattern is represented in code as a list of binary values indicating whether each beat (subdivision) is associated with a note activation.

Finally, for the alto voice (piano), we incorporate rhythmic roughness, which is a measure of how irregular the rhythm of a piece of music is. Music with smooth, regular rhythms are typically perceived as higher in valence. In AffectMachine-Classical, we use note density as a proxy for rhythmic roughness (Wallis et al., 2011). As arousal increases, roughness decreases, and note density increases. When roughness is 0, each bar is populated with eight notes of equal length. However, this often results in overly dense-sounding output, because tempo is also high at higher levels of arousal. Hence, we limit the lowest roughness to 0.3.

2.3.2. Tempo

Tempo, or beats per minute, determines how quickly the notes of each bar are played. Alternatively, tempo can be thought of as a measure of note duration—the faster the tempo, the shorter the note duration. In AffectMachine-Classical, tempo is determined by a simple linear relationship with arousal, and ranges from 60 bpm at $Arousal = 0$–200 bpm at $Arousal = 1$.

2.4. Timbral and loudness parameters

Two parameters contributed to variations in timbre: (i) the instrumentation of AffectMachine-Classical, and (ii) the velocity of notes, which refers to the force with which a note is played.

2.4.1. Velocity range

Similar to the algorithmic composition system developed by Williams et al. (2017), we mapped coordinates with higher arousal to brighter and harder timbres that were created by increasing MIDI key velocity. In MIDI, velocity is measured on a scale from 0 to 127. In our system, the range of permissible MIDI key velocities is [40, 70] at Arousal = 0, and the lower and upper bounds of the range increase linearly with arousal to [85, 115] at Arousal = 1. A uniform distribution over the range is used to determine the velocity for each bar.

$$Velocity = unif\, 40 + aro * 45, 70 + aro * 45 \qquad (2)$$

2.4.2. Velocity variation

Patterns of velocity variation have affective consequences. For example, research has found that large changes in velocity (loudness) suggest fear, while small variations convey happiness and pleasantness (Scherer and Oshinsky, 1977; Krumhansl, 1997; Juslin and Laukka, 2001; Gabrielsson and Lindström, 2010). Further, rapid changes in velocity may be associated with playfulness or fear (Krumhansl, 1997).

In our experimentation with the system, we found that frequent changes in velocity tend to result in unpleasantly disjointed, artificial-sounding music, and we therefore attempt to limit large, rapid (e.g., unexpected-sounding) variations in velocity. Furthermore, changes in velocity become more frequent as tempo (which is linearly related to arousal, as per Section 2.3.2) increases. Therefore, to strike a balance between enabling sufficient variation in velocity, and incorporating those variations in as natural a way as possible, we limited the maximum change in velocity allowable within each bar. The variation in velocity is dependent on the arousal level and bar of the progression. Specifically, we set an overall minimum and maximum loudness level, and the allowed deviation becomes smaller as arousal decreases. The magnitude of variation in velocity is random, within the allowable range (which is set for each bar), and there are no changes in velocity within each bar.

2.4.3. Instrumentation

Four virtual instruments were employed in the system (piano, a string section, clarinet, and marimba), and used to convey a classical musical style. As mentioned previously, the lowest voice is conveyed by the string section, while both inner voices are carried by the piano. The principal melody is placed in the uppermost voice, which is played by the clarinet. The marimba is used to double over the clarinet at high levels of valence ($Valence \geq 0.8$) due to its cheerful-sounding timbre (and because, during experimentation with the system, marimba was found to nicely complement the timbre of the clarinet, which could sound slightly shrill at higher pitch heights). After all other harmonic, rhythmic, and timbral parameters have been determined, instrument samples in the DAW (Ableton) are used to generate the final output audio.

3. AffectMachine-Classical listener study

3.1. Method

A listening study was conducted in order to validate the efficacy of AffectMachine-Classical for generating affective music. We first used our system to generate brief musical examples from different points around the arousal-valence space of the circumplex model (Russell, 1980). Listeners then provided arousal and valence ratings for each of these excerpts to examine whether the target emotion (in terms of arousal and valence) was indeed perceived as intended by listeners.

3.1.1. Participants

The listening study was conducted with 26 healthy participants (average age = 22 yrs, SD = 4 yrs) including 11 male and 15 female participants. Twelve of the 26 participants reported having prior musical training. All the participants were given verbal and written instructions about the listening study prior to providing their written consent. The study was approved by the Institutional Review Board (IRB) of the National University of Singapore (NUS).

3.1.2. Stimuli

AffectMachine-Classical was designed to compose affective music that can span the entire valence-arousal plane. For the validation study, musical stimuli were generated from 13 different points around the valence and arousal plane. These were meant to represent different emotional states around the space, and covered the corners, middle of each quadrant, and the neutral middle point of the space. The points are: {valence, arousal} = [{0,0}; {0,0.5}; {0,1}; {0.25;0.25}; {0.25,0.75}; {0.5,0}; {0.5,0.5}; {0.5,1}; {0.75,0.25}; {0.75,0.75}; {1,0}; {1,0.5}; {1,1}]. There is a precedent in the literature for selecting these points in the arousal-valence plane for the validation of a music generation system (Ehrlich et al., 2019).

To account for the probabilistic nature of the system, three different musical stimuli were generated from each of the thirteen points, resulting in a total of 39 musical excerpts. This mitigates the risk that artifacts in any particular stimulus might bias listener ratings, for more robust results. The average duration of the music stimuli is 23.6 s. The stimuli were composed based on either an 8- or 16-bar progression to allow the music to reach a cadence. Note that because AffectMachine was designed to generate music continuously and flexibly based on the listener's physiological state or real-time arousal and valence values, the music does not always reach a full cadence at the end of an 8-bar sequence (e.g., sometimes the tonic/cadence is only reached at the beginning of the subsequent 8-bar sequence). In the present case, we are not testing the ability of the music to have well-formed cadences *per se*, but to convey a target emotion. That is, the examples do not necessarily end with a musical cadence; rather, they are excerpts from what could be an infinitely-long musical creation. Therefore, while generating stimuli with a fixed duration is possible, this often results in stimuli that end abruptly, which might influence a listener's emotional response to the stimuli. Sixteen bars were used for stimuli with a fast tempo (e.g., high arousal excerpts), as 8 bars produced too brief a time duration for these excerpts. All musical stimuli were presented to each participant in randomized order to avoid order effects across participants. The music stimuli used in this validation study are available online at: https://katagres.com/AffectMachineClassical_stimuli.

3.1.3. Experimental protocol

The experiment was conducted one participant at a time in a quiet room with minimal auditory and visual distractions. The experimenter first provided verbal and written instructions about the experiment, and then the participant provided written, informed consent to participate in the study. During the listening study, the participant sat in front of a computer and listened to the music stimuli over headphones, with the sound level adjusted to a comfortable listening volume.

Before the listening task, the participant was asked to complete a demographic questionnaire which included questions about his/her age, prior musical training, ethnicity, etc. Subsequently, the participant rated his/her current emotional state.

The music listening study began with two practice trials, followed by the 39 experimental trials in randomized order. After listening to each stimulus, the participant was asked to indicate the *perceived* emotion of the stimulus (that is, the emotions they felt that the music conveyed) on a visual 9-point scale known as the Self-Assessment Manikin (SAM; Bradley and Lang, 1994). These ratings were collected for both arousal and valence. Briefly, valence refers to the degree of the pleasantness of the emotion, while arousal refers to the activation or energy level of the emotion. The SAM scale ranged from "very unpleasant" (1) to "extremely pleasant" (9) for valence, and from "calm" (1) to "excited" (9) for arousal. Participants were allowed to take as long as they required to make these ratings, but were only permitted to listen to each musical stimulus once. The total duration of the experiment was ∼40 min, and participants were compensated with $6 SGD (equivalent to $4.50 USD) for their time.

3.2. Results and discussion

In order to evaluate the efficacy of the music generation system, we analyzed the user ratings collected during the music listening study. We aimed to investigate (1) whether the music generated by the system is able to express the desired level of valence and arousal to the listeners, and (2) whether perceived valence and arousal are dependent on the listeners' prior musical training/knowledge. In this regard, we present our results in two subsections: (1) arousal and valence ratings, and (2) the impact of prior musical training on emotion ratings. We do not consider demographic factors such as age and ethnicity for further analysis due to the limited sample size.

As is commonly found in listener studies of emotion in music, we observed that the average valence and arousal ratings varied across listeners. This variance is often attributed to individual differences in musical preferences and training, and the listeners' demographic and cultural profile (Koh et al., 2023). In order to mitigate the differences across listeners, we normalized the perceptual ratings from each user (see Equations 3 and 4 below). Here, $Max_{Valence}$ refers to the maximum possible valence rating (i.e., 9), and $Min_{valence}$ refers to the minimum possible valence rating (i.e., 1). The same Max and Min values apply to Arousal. The normalized valence and normalized arousal ratings, ranging between 0 and 1, are used for further analysis. In the remainder of the article, the normalized valence and normalized arousal ratings will be referred to as valence and arousal ratings, respectively.

$$Normalized_{Valence} = \frac{Rated_{Valence}}{(Max_{Valence} - Min_{Valence})} \tag{3}$$

$$Normalized_{Arousal} = \frac{Rated_{Arousal}}{(Max_{Arousal} - Min_{Arousal})} \tag{4}$$

3.2.1. Arousal and valence ratings

To investigate whether AffectMachine is able to accurately express the intended emotion through music, we compared participants' averaged (normalized) emotion ratings for the musical stimuli with the valence or arousal parameter settings used during the music generation process. For example, the averaged valence ratings for all stimuli generated with the parameter settings {valence, arousal} = [{0,0}; {0,0.5}; {0,1}] were used to evaluate the system's performance when valence is set to zero. The bar graphs depicting the averaged ratings (along with standard errors) are presented in Figure 1. As expected, a strong increasing trend is seen for both the average valence and arousal ratings with respect to their corresponding parameter settings. With regard to the valence ratings, we observe the majority of ratings to fall between the < 0.25 and > 0.75 parameter settings. It is common to see a higher density of responses in the middle of psychometric rating scales (e.g., with both ends of the scale receiving proportionally fewer responses; Leung, 2011). This could also indicate that the extremes of the valence parameter values are less distinguishable by listeners. On the other hand, a better correspondence is observed between average arousal ratings and the respective parameter values at all levels of arousal.

To test the relationship between average valence and arousal user ratings and parameter settings, we performed linear regression analyses (illustrated in Figure 1). The coefficient of determination is $R^2 = 0.90$ ($F = 27, p < 0.05$) for valence, and $R^2 = 0.96$ ($F = 74, p < 0.01$) for arousal, which confirms that both parameters are very effective in conveying their intended dimension of emotion. The results also show a stronger linear relationship for arousal (between average arousal ratings and parameter settings) in comparison to valence. This finding, in which arousal is more reliably expressed via music than valence, has previously been found in the literature (Wallis et al., 2011; Ehrlich et al., 2019). These results show that the music generated by AffectMachine-Classical generally conveys the intended levels of valence and arousal to listeners.

Next, we investigate whether the perception of valence is influenced by changes in the arousal parameter setting, and conversely whether the perception of arousal is influenced by changes in the valence parameter setting. To do so, we analyse the dependence of average emotion ratings on both the valence and arousal parameter settings together. Figure 2 visualizes this dependence by presenting the interpolated average valence (left) and arousal ratings (right) as a function of the emotion parameter settings. The stars in the figure represent the 13 points around the valence and arousal plane used to generate musical stimuli. As can be seen in the figure on the left, the perceived valence is lower than the actual valence parameter setting (for $V > 0.7$) for excerpts expressing arousal values < 0.4. That is, excerpts generated to express high valence convey only moderate valence when the arousal setting is low. This may be due in part to the effect of a slower tempo. Ratings at low valence settings are, however, in accordance with their respective parameter values. In contrast, we observe uniform correspondence between the arousal parameter values and arousal ratings regardless of the valence parameter setting. Our study replicates a phenomenon that has been previously described in Wallis et al. (2011)—the authors found asymmetrical "crossover" effects between arousal and valence such that while perceived valence correlates with intended arousal, perceived arousal does not correlate significantly with intended valence.

To investigate these linear dependencies, we performed multiple linear regression between the valence and arousal parameter settings (independent variables) and average valence/arousal rating (dependent variable). The results indicate that perceived valence ratings are significantly influenced by both the valence ($F = 63, p < 0.001$) and arousal ($F = 11, p < 0.01$) parameter settings, which is in line with what we observed in Figure 2. Perceived arousal ratings, however, only show a significant dependence on the arousal settings ($F = 153, p < 0.001$). This observation is in line with findings from the literature which show that modeling the arousal component of emotion is more straightforward than the valence component (Yang et al., 2008; Wallis et al., 2011). Nevertheless, the obtained R^2 values are high $R^2 > 0.85$ for both average valence and arousal ratings. This confirms that irrespective of the emotion component, the majority of variability in average ratings during multiple regression analysis is explained by the valence and arousal settings values.

In summary, the listener study validates the ability of AffectMachine-Classical to generate music that expresses desired levels of emotion, measured in terms of arousal and valence. This confirms that the system has the potential to be deployed in applications that benefit from affective music—for example, the AffectMachine-Classical could be integrated with biofeedback systems wherein the music driven by the users' neural (or other physiological) signals can be used to reflect their emotional state. This direction is promising for developing more sophisticated emotion mediation systems with applications in healthcare (Agres et al., 2021). In the next section, we analyse the impact of participants' prior musical training on emotion ratings.

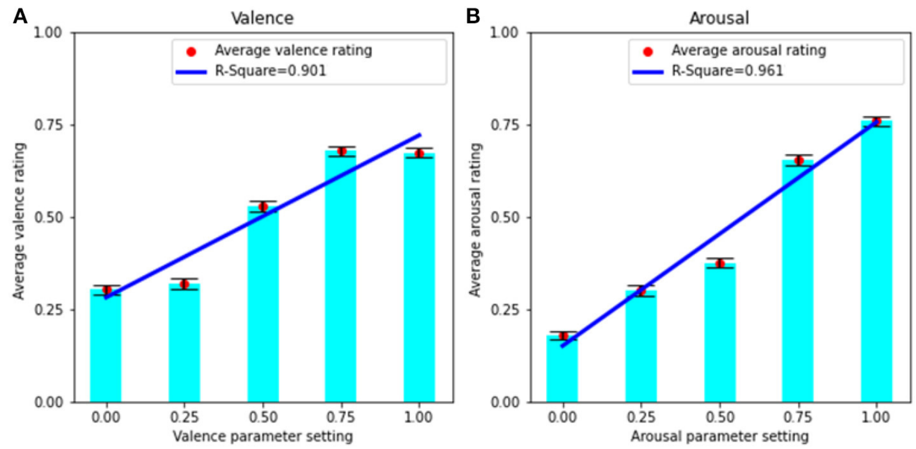

FIGURE 1
(A) Linear regression between parameter settings and average valence ratings. **(B)** Linear regression between parameter settings and average arousal ratings. Error bars display standard error.

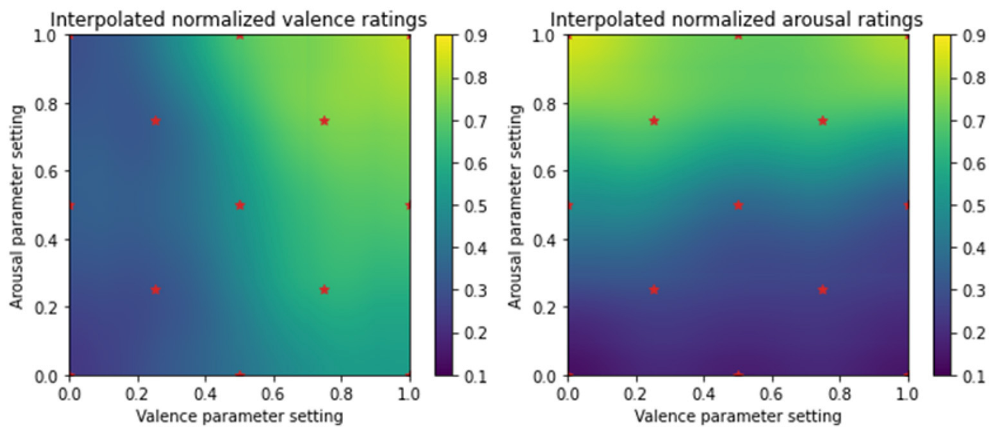

FIGURE 2
Average (interpolated) valence and arousal ratings as a function of the valence and arousal parameters. Vertical color bars represent the colors corresponding to different values of normalized average ratings over the range of 0.1–0.9.

3.2.2. Impact of prior musical training on emotion ratings

In this section, we present a comparison of user ratings provided by participants with and without prior musical training. Participants indicated whether they had prior musical training in the demographic questionnaire they completed. Based on participants' response to the question "Do you currently play an instrument (including voice)?" they were divided into two groups—the musical training (MT) group and no musical training (NMT) group. The MT and NMT groups have 12 and 14 participants, respectively.

Figure 3 presents the average emotion ratings corresponding to different levels of emotion parameter values for both the MT and NMT groups. As illustrated in the graphs, a stronger correspondence between the average emotion ratings and parameter-setting values is observed for arousal in comparison to valence, for both groups. As noted above, the average valence ratings demonstrate a saturation effect for lower (< 0.25) and higher (> 0.75) parameter-setting values for both the MT and NMT groups. Figure 3 also shows the linear regression fit for all the cases. The R^2 values reflecting the relationship between emotion ratings and parameter settings are marginally higher for the MT group (R^2 for valence is 0.91, $F = 33, p = 0.01$; R^2 for arousal is 0.97, $F = 111, p < 0.01$) as compared with the NMT group (R^2 for valence is 0.88, $F = 22, p < 0.05$; R^2 for arousal is 0.94, $F = 51$, p<0.01), for both valence and arousal. To compare whether the differences between these linear regression models were significant, we calculated the Akaike Information Criterion (AIC) for both the MT group ($AIC = -10.65$ for valence and $AIC = -14.87$ for arousal) and NMT group ($AIC = -12.20$ for valence and $AIC = -11.74$ for arousal). The statistics show that there is only a marginal difference (in perceived emotion ratings based on the system's emotion settings) between the MT and NMT groups. Although musical expertise has been found to influence the perception of emotion in affective music in

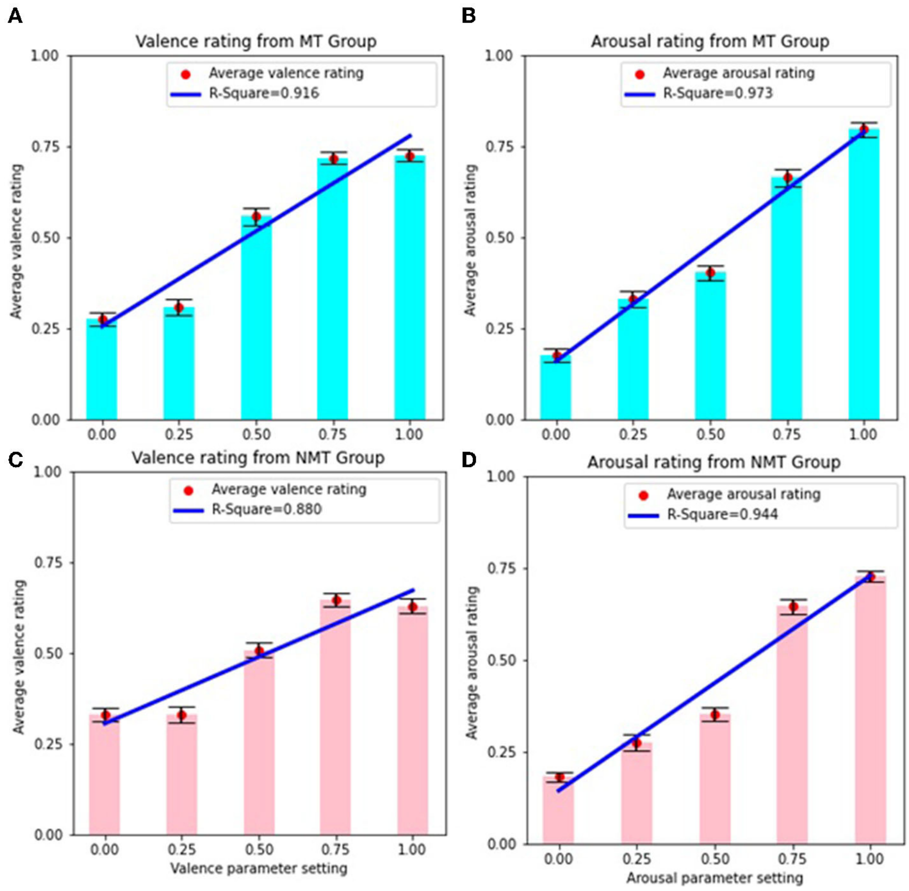

FIGURE 3
(A) Average valence rating and linear regression for MT group. **(B)** Average arousal rating and linear regression for MT group. **(C)** Average valence rating and linear regression for NMT group. **(D)** Average arousal rating and linear regression for NMT group. Error bars depict standard error.

some cases (e.g., see Di Mauro et al., 2018), we find here that both musicians and non-musicians reliably appraise the music created by AffectMachine-Classical as the emotion intended by the system. We note, however, that given the limited sample size in our study, it is difficult to generalize the effects of musical training, and a larger sample size could yield a significant difference between the two listener groups. Nevertheless, we observe that regardless of musical training, all of the participants were able to reliably perceive the emotional expression in the music, which is evident from the high R^2 values observed (> 0.85) for both listener groups.

In addition, we also performed a multiple linear regression to examine the effect of parameter settings in both emotion dimensions on individual perceived emotion ratings. We obtained high R^2 values ($R^2 > 0.8$) for all the scenarios, i.e., for both emotion dimensions for both groups. Furthermore, we observed that perceived valence is significantly influenced by both valence ($F = 89, p < 0.001$ for MT, and $F = 38, p < 0.001$ for NMT) and arousal ($F = 5.8, p < 0.05$ for MT, and $F = 15, p < 0.01$ for NMT) parameter settings for both MT and NMT groups. However, perceived arousal ratings are only influenced by the arousal settings ($F = 216, p < 0.001$ for MT and $F = 110, p < 0.001$ for NMT), and not valence settings, in both groups. These findings are similar to what we observed for all the participants without any grouping.

4. General discussion

In this paper, we present a new computational system for generating affective classical music called AffectMachine-Classical. The system provides a probabilistic, rule-based algorithm for flexibly generating affective music in real-time. AffectMachine's behavior essentially resembles semi-structured musical improvisation, not dissimilar to the approach utilized by Baroque composers/musicians (Moersch, 2009), or how human jazz performers might follow the basic melody outlined by a lead sheet while coming up with reharmonizations, chord voicings, and appropriate accompaniments, on the fly, to help convey the emotions they are aiming to express (Johnson-Laird, 2002; McPherson et al., 2014). To our knowledge, ours is the first affective music generation system to adopt this approach. A key advantage of this method is that the music generated by the system achieves a balance between musical coherence and self-similarity, which may be valuable in research and other contexts that require lengthier pieces of music. Although the issue of artificially generating music that is capable of exhibiting long-term structure has been described as "notoriously difficult" (Carnovalini and Rodà, 2020) and cited as one of the grand challenges for automatic music generation (Herremans et al., 2017; Briot and Pachet, 2020), our system addresses this issue in part by providing a structural frame by means of an 8-bar form in which the music is generated. Melodic coherence is maintained due to the constraints enforced by the algorithms used to generate melodic patterns, and harmonic coherence is achieved through the use of a chord matrix based on the 8-bar form.

AffectMachine was developed to be embedded into real-time biofeedback systems, such as music-based Brain-Computer Interfaces (BCIs), to leverage neurofeedback and adaptive, affective music generation to help the listener achieve a target emotion state. The listener study reported here was conducted to validate the efficacy of the system for generating affective music. Indeed, regardless of musical experience, listeners perceived the target emotion of the musical excerpts (in terms of arousal and valence), as intended by the system.

The results of the listener study indicate a strong relationship between the arousal parameter setting and average arousal ratings ($R^2 = 0.96$), as well as the valence setting and average valence ratings ($R^2 = 0.90$). The correlation between target and perceived emotion was more tempered for valence compared to arousal, as previously found in the literature (e.g., Wallis et al., 2011; Ehrlich et al., 2019). From the results of our listener study, it is evident that AffectMachine is capable of expressing the desired emotional information, and thus holds the potential to be used as an affect guide for mediating/regulating the emotion states of listeners. We would like to emphasize that despite the differences in listeners' prior musical training, individual and cultural preferences, and demographic profile, there was strong evidence suggesting that the system's target emotions were indeed perceived as intended across listeners, which makes Affect Machine-Classical a very promising tool for creating music with reliable emotion perception.

In terms of future directions, as discussed above, AffectMachine will be embedded into biofeedback systems, such as a Brain-Computer-Interface (similar to Ehrlich et al., 2019), to support emotion self-regulation in listeners. Further, our system may be used for wellness applications such as generating affective music "playlists" for emotion mediation. That is, using the flexible music generation system, a user may pre-define an "emotion trajectory" (e.g., a path through emotion space, such as the two-dimensional Valence-Arousal space) to define the emotional qualities of their music over the duration of listening. For example, if a user desires 10 min of music to help him move from a depressed emotion state to a happy emotion state, he may indicate an emotion trajectory from negative arousal/valence to positive arousal/valence over the specified duration, and the system will create bespoke affective music to this specification. Therefore, AffectMachine has the potential to be embedded in various kinds of well-being applications to create highly-personalized, affective music for emotion regulation.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB) of the National University of Singapore (NUS). The patients/participants provided their written informed consent to participate in this study.

Author contributions

KA and AD led the research. KA initiated, supervised the project, and led the manuscript preparation and revision. AD led data collection for the listener study and as well as data analysis and reporting. PC led the system development, under the supervision of KA, and as well as the system description. All authors contributed to writing the paper and approved of the submitted version.

Funding

This work was supported by the RIE2020 Advanced Manufacturing and Engineering (AME) Programmatic Fund (No. A20G8b0102), Singapore.

Acknowledgments

We would like to thank the composer Cliff Tan for his musical ideas, help in refining musical parameters, and assistance with DAW implementation when developing AffectMachine-Classical. We would also like to thank the composer Wen Liang Lim for contributing to our project ideation in the early stages of this work.

Conflict of interest

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

Publisher's note

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

References

Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., et al. (2023). MusicLM: Generating music from text. *arXiv preprint arXiv:2301.11325*. doi: 10.48550/arXiv.2301.11325

Agres, K. R., Schaefer, R. S., Volk, A., van Hooren, S., Holzapfel, A., Dalla Bella, S., et al. (2021). Music, computing, and health: a roadmap for the current and future roles of music technology for health care and well-being. *Music Sci*. 4:2059204321997709. doi: 10.1177/2059204321997709

Altenmüller, E., and Schlaug, G. (2012). "Music, brain, and health: exploring biological foundations of music's health effects," in *Music Health Wellbeing*, eds R. MacDonald, G. Kreutz, and L. Mitchell (New York, NY: Oxford University Press), 12–24. doi: 10.1093/acprof:oso/9780199586974.003.0002

Balkwill, L.-L., and Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Percept*. 17, 43–64. doi: 10.2307/40285811

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Briot, J.-P., and Pachet, F. (2020). Deep learning for music generation: challenges and directions. *Neural Comput. Appl*. 32, 981–993. doi: 10.1007/s00521-018-3813-6

Carnovalini, F., and Rodà, A. (2020). Computational creativity and music generation systems: an introduction to the state of the art. *Front. Artif. Intell*. 3:14. doi: 10.3389/frai.2020.00014

Cohrdes, C., Wrzus, C., Wald-Fuhrmann, M., and Riediger, M. (2020). "The sound of affect": age differences in perceiving valence and arousal in music and their relation to music characteristics and momentary mood. *Music Sci*. 24, 21–43. doi: 10.1177/1029864918765613

Collier, W. G. and Hubbard, T. L. (1998). Judgments of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicology* 17:36. doi: 10.1037/h0094060

Costa, M., Fine, P. A., and Bitti, P. E. R. (2004). Interval distributions, mode, and tonal strength of melodies as predictors of perceived emotion. *Music Percept*. 22, 1–14. doi: 10.1525/mp.2004.22.1.1

Costa, M., and Nese, M. (2020). Perceived tension, movement, and pleasantness in harmonic musical intervals and noises. *Music Percept*. 37, 298–322. doi: 10.1525/mp.2020.37.4.298

Costa, M., Ricci Bitti, P. E., and Bonfiglioli, L. (2000). Psychological connotations of harmonic musical intervals. *Psychol. Music* 28, 4–22. doi: 10.1177/0305735600281002

Dash, A., and Agres, K. R. (2023). Ai-based affective music generation systems: a review of methods, and challenges. *arXiv preprint arXiv:2301.06890*. doi: 10.48550/arXiv.2301.06890

Di Mauro, M., Toffalini, E., Grassi, M., and Petrini, K. (2018). Effect of long-term music training on emotion perception from drumming improvisation. *Front. Psychol*. 9:2168. doi: 10.3389/fpsyg.2018.02168

Dong, H.-W., Hsiao, W.-Y., Yang, L.-C., and Yang, Y.-H. (2018). "Musegan: multi-track sequential generative adversarial networks for symbolic music generation and accompaniment," in *Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 32*, (New Orleans, LA), doi: 10.1609/aaai.v32i1.11312

Ehrlich, S. K., Agres, K. R., Guan, C., and Cheng, G. (2019). A closed-loop, music-based brain-computer interface for emotion mediation. *PLoS ONE* 14:e0213516. doi: 10.1371/journal.pone.0213516

Fancourt, D., and Finn, S. (2019). *What is the Evidence on the Role of the Arts in Improving Health and Well-being? A Scoping Review*. World Health Organization Regional Office for Europe.

Ferreri, L., Mas-Herrero, E., Zatorre, R. J., Ripollés, P., Gomez-Andres, A., Alicart, H., et al. (2019). Dopamine modulates the reward experiences elicited by music. *Proc. Natl. Acad. Sci. U.S.A*. 116, 3793–3798. doi: 10.1073/pnas.1811878116

Gabrielsson, A., and Juslin, P. N. (2003). "Emotional expression in music," in *Handbook of Affective Sciences*, eds R. J. Davidson, K. R. Scherer, and H. H. Goldsmith (Oxford University Press), 503–534.

Gabrielsson, A., and Lindström, E. (2001). The influence of musical structure on emotional expression.

Gabrielsson, A., and Lindström, E. (2010). The role of structure in the musical expression of emotions. *Handb. Music Emot*. 367400, 367–344. doi: 10.1093/acprof:oso/9780199230143.003.0014

Gungormusler, A., Paterson-Paulberg, N., and Haaahr, M. (2015). "barelymusician: An adaptive music engine for video games," in *Audio Engineering Society Conference: 56th International Conference: Audio for Games* (Audio Engineering Society).

Herremans, D., Chuan, C.-H., and Chew, E. (2017). A functional taxonomy of music generation systems. *ACM Comput. Surveys* 50, 1–30. doi: 10.1145/3108242

Hevner, K. (1937). The affective value of pitch and tempo in music. *Am. J. Psychol*. 49, 621–630. doi: 10.2307/1416385

Hu, Z., Liu, Y., Chen, G., Zhong, S.-h., and Zhang, A. (2020). "Make your favorite music curative: Music style transfer for anxiety reduction," in *Proceedings of the 28th ACM International Conference on Multimedia* (Seattle, WA: ACM), 1189–1197. doi: 10.1145/3394171.3414070

Huang, Q., Jansen, A., Lee, J., Ganti, R., Li, J. Y., and Ellis, D. P. (2022). Mulan: a joint embedding of music audio and natural language. *arXiv preprint arXiv:2208.12415*. doi: 10.48550/arXiv.2208.12415

Huron, D. (2001). Tone and voice: a derivation of the rules of voice-leading from perceptual principles. *Music Percept*. 19, 1–64. doi: 10.1525/mp.2001.19.1.1

Huron, D. (2008). *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Percept*. 19, 415–442. doi: 10.1525/mp.2002.19.3.415

Juslin, P. N. (1997). Can results from studies of perceived expression in musical performances be generalized across response formats? *Psychomusicology* 16:77. doi: 10.1037/h0094065

Juslin, P. N., and Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion* 4, 381–412. doi: 10.1037/1528-3542.1.4.381

Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychol. Bull*. 129, 770–814. doi: 10.1037/0033-2909.129.5.770

Juslin, P. N., and Laukka, P. (2004). Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *J. New Music Res*. 33, 217–238. doi: 10.1080/0929821042000317813

Juslin, P. N., and Sloboda, J. A. (2013). *Music and Emotion*. Elsevier Academic Press. doi: 10.1016/B978-0-12-381460-9.00015-8

Juslin, P. N., and Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci*. 31, 559–575. doi: 10.1017/S0140525X08005293

Koh, E. Y., Cheuk, K. W., Heung, K. Y., Agres, K. R., and Herremans, D. (2023). MERP: a music dataset with emotion ratings and raters' profile information. *Sensors* 23:382. doi: 10.3390/s23010382

Krumhansl, C. (1997). An exploratory study of musical emotions and psychophysiology. *Can. J. Exp. Psychol*. 4, 336–353. doi: 10.1037/1196-1961.51.4.336

Leung, S.-O. (2011). A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point likert scales. *J. Soc. Service Res*. 37, 412–421. doi: 10.1080/01488376.2011.580697

Lindström, E. (2006). Impact of melodic organization on perceived structure and emotional expression in music. *Music. Sci*. 10, 85–117. doi: 10.1177/102986490601000105

Livingstone, S. R., and Brown, A. R. (2005). "Dynamic response: real-time adaptation for music emotion," in *ACM International Conference Proceeding Series*, Vol. 123 (Association for Computing Machinery), 105–111.

Lonsdale, A. J., and North, A. C. (2011). Why do we listen to music? A uses and gratifications analysis. *Br. J. Psychol*. 102, 108–134. doi: 10.1348/000712610X506831

MacDonald, R., Kreutz, G., and Mitchell, L. (2013). *Music, Health, and Wellbeing*. Oxford University Press. doi: 10.1093/acprof:oso/9780199586974.001.0001

McPherson, M. J., Lopez-Gonzalez, M., Rankin, S. K., and Limb, C. J. (2014). The role of emotion in musical improvisation: an analysis of structural features. *PLoS ONE* 9:e105144. doi: 10.1371/journal.pone.0105144

Moersch, C. M. (2009). "Keyboard improvisation in the baroque period," in *Musical Improvisation: Art, Education, and Society*, eds G. Solis and B. Nettl (University of Illinois Press), 150–170.

Ramirez, R., Palencia-Lefler, M., Giraldo, S., and Vamvakousis, Z. (2015). Musical neurofeedback for treating depression in elderly people. *Front. Neurosci*. 2015:354. doi: 10.3389/fnins.2015.00354

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol*. 39:1161. doi: 10.1037/h0077714

Saarikallio, S. (2012). Development and validation of the brief music in mood regulation scale (b-MMR). *Music Percept*. 30, 97–105. doi: 10.1525/mp.2012.30.1.97

Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., and McIntosh, A. R. (2015). Predictions and the brain: how musical sounds become rewarding. *Trends Cogn. Sci*. 19, 86–91. doi: 10.1016/j.tics.2014.12.001

Scherer, K. R., and Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motiv. Emot*. 1, 331–346. doi: 10.1007/BF00992539

Schmuckler, M. A. (1989). Expectation in music: investigation of melodic and harmonic processes. *Music Percept*. 7, 109–149. doi: 10.2307/40285454

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Percept*. 21, 561–585. doi: 10.1525/mp.2004.21.4.561

Sloboda, J. (2004). "Musical performance and emotion: issues and developments," in *Exploring the Musical Mind: Cognition, Emotion, Ability, Function* (Oxford University Press). doi: 10.1093/acprof:oso/9780198530121.003.0013

Swaminathan, S., and Schellenberg, E. G. (2015). Current emotion research in music psychology. *Emot. Rev*. 7, 189–197. doi: 10.1177/1754073914558282

Thayer, R. E., Newman, J. R., and McClain, T. M. (1994). Self-regulation of mood: strategies for changing a bad mood, raising energy, and reducing tension. *J. Pers. Soc. Psychol*. 67:910. doi: 10.1037/0022-3514.67.5.910

Wallis, I., Ingalls, T., Campana, E., and Goodman, J. (2011). "A rule-based generative music system controlled by desired valence and arousal," in *Proceedings of 8th International Sound and Music Computing Conference (SMC)* (Padova), 156–157.

Williams, D., Kirke, A., Miranda, E., Daly, I., Hwang, F., Weaver, J., et al. (2017). Affective calibration of musical feature sets in an emotionally intelligent music composition system. *ACM Trans. Appl. Percept*. 14, 1–13. doi: 10.1145/3059005

Yang, X., Dong, Y., and Li, J. (2018). Review of data features-based music emotion recognition methods. *Multim. Syst*. 24, 365–389. doi: 10.1007/s00530-017-0559-4

Yang, Y.-H., Lin, Y.-C., Su, Y.-F., and Chen, H. H. (2008). A regression approach to music emotion recognition. *IEEE Trans. Audio Speech Lang. Process*. 16, 448–457. doi: 10.1109/TASL.2007.911513

Zentner, M., Grandjean, D., and Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion* 8:494. doi: 10.1037/1528-3542.8.4.494

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 20 June 2023
DOI 10.3389/fcomp.2023.1158046

OPEN ACCESS

EDITED BY
Kat Agres,
National University of Singapore, Singapore

REVIEWED BY
Usha Shastri,
Kasturba Medical College, Mangalore, India
C. S. Vanaja,
Bharati Vidyapeeth Deemed University, India

*CORRESPONDENCE
Ramya Vaidyanath
✉ ramyavaidyanath@sriramachandra.edu.in

RECEIVED 03 February 2023
ACCEPTED 05 June 2023
PUBLISHED 20 June 2023

CITATION
Ramadas V, Vaidyanath R, Uppunda AK and Viswanathan S (2023) Computerized music-based intervention module for auditory processing and working memory in older adults. *Front. Comput. Sci.* 5:1158046.
doi: 10.3389/fcomp.2023.1158046

Computerized music-based intervention module for auditory processing and working memory in older adults

Vaishnavi Ramadas[1], Ramya Vaidyanath[1]*, Ajith Kumar Uppunda[2] and Sushma Viswanathan[1]

[1]Sri Ramachandra Institute of Higher Education and Research, Chennai, India, [2]All India Institute of Speech and Hearing (AIISH), Mysore, India

Introduction: The contribution of technology to the field of health is vast, both in diagnosis and management. More so, the use of computer-based intervention has become increasingly widespread over the past decade. Human beings experience a decline in auditory processing and cognitive skills as they age, consistent with deterioration of other bodily functions. In addition, speech perception abilities in both quiet and in the presence of noise are impacted by auditory processing abilities and cognitive skills such as working memory. This pilot study explored the use of music as an intervention for improving these skills and employed a computerized delivery of the intervention module.

Method: A battery of tests was carried out to assess the baseline auditory processing and working memory skills in eight older adults between the ages of 56 and 79 years, all of whom had normal hearing. Following the assessment, a short-term computerized music-based intervention was administered. The style of music chosen was Carnatic classical music, a genre widely practiced in Southern India. The intervention module involved note and tempo discrimination and was carried out for a maximum of 10 half-hour sessions. The multi-level intervention module was constructed and administered using Apex software. Following the intervention, the auditory processing and cognitive skills of the participants were reassessed to study any changes in their auditory processing and working memory skills.

Results and discussion: There were positive changes observed in all the auditory processing and some of the working memory abilities. This paper discusses in detail the systematic structuring of the computerized music-based intervention module and its effects on the auditory processing and cognitive skills in older adults.

KEYWORDS

older adults, music-based intervention, computerized, auditory processing (AP), cognition, working memory, intervention

1. Introduction

Technology has widely influenced health care with respect to evaluation, diagnosis, intervention as well as rehabilitation. Currently, technology augments clinical observations of physicians and aids in precision of diagnosis (Van Os et al., 2013; Dunkel et al., 2021; Tilli, 2021). The impact of the pandemic has been extensively felt in the follow up to hospitals for reviews, therapy, long duration interventions, etc. It has especially impacted the older population and people with poor health conditions due to their co-morbidities.

Hearing loss is becoming so widespread that the World Health Organization (2019) predicts that one in every ten people in the world will have disabling hearing loss by 2050. They also estimate that currently, more than 25% of people over the age of 60 years are impacted by disabling hearing loss. Auditory processing deficits are commonly observed in older adults even in the absence of any hearing loss (Füllgrabe et al., 2015; Sardone et al., 2019). Auditory processing deficits refer to difficulties in processing sounds in order to make sense of the auditory information. Normal auditory processing occurs due to the collective functioning of different mechanisms such as auditory discrimination, temporal processing (acoustic timing related information), and binaural processing (combining acoustic information from both ears) (American Speech-Language-Hearing Association, 2005). In older adults, auditory processing deficits are commonly exhibited as difficulty understanding speech in difficult listening situations such as talking over the phone, noisy environments (Rodriguez et al., 1990; Boboshko et al., 2018), difficulty maintaining attention to spoken information (Wayne et al., 2016), and asking for frequent repetitions (American Speech-Language-Hearing Association, n.d.), among others. Rehabilitation for people with auditory processing deficits involves working on the affected deficit or auditory process, environmental modifications to help improve the ease of listening, and providing assistive listening devices to enhance the acoustic quality.

Music has been repeatedly shown to have positive effects on different physical and mental aspects of the human body such as pain relief (Kumar et al., 2014), quality of sleep (Deshmukh et al., 2009), improved concentration, reduction in anxiety (Sung et al., 2010), and reduction in psychiatric symptoms (Lyu et al., 2018; Schroeder et al., 2018; Tsoi et al., 2018). Likewise, music training also has a positive impact on cognitive processes such as attention and memory, and other auditory functions such as listening in challenging situations (Besson et al., 2011; Strait and Kraus, 2011b). Patel (2011) proposed the OPERA hypothesis which also suggests that music benefits the ability of speech processing. The hypothesis states that music contributes to adaptive plasticity in the neural networks responsible for speech understanding when five conditions are met: there is **O**verlap in the neural networks between speech and music, music requires a higher precision of **P**rocessing from the neural networks, music elicits positive **E**motions, musical activities are **R**epetitive, and musical activities require high levels of **A**ttention. Patel (2014) also proposed the extended OPERA hypothesis, incorporating the effects of non-linguistic music on cognitive processes such as auditory attention and working memory. It was proposed that these cognitive processes could also benefit from musical training if the following conditions were met: there existed shared neural pathways, music placed higher demands on the cognitive processes than speech, and if music and the cognitive processes were combined with emotion, repetition, and attention. Hence, employing a music-based intervention to improve auditory processing and cognitive skills may be considered suitable.

The contribution of technology in the area of rehabilitation of individuals with hearing difficulties has been enormous. Wireless technology including Bluetooth and near frequency magnetic induction are being incorporated in amplification devices and hearing solutions (Kim and Kim, 2014). Hearing devices such as cochlear implants, hearing aids, and assistive listening devices benefit from the prime technological advances made by the industry and contribute to better quality of life of people with hearing difficulties (Hansen et al., 2019; Corey and Singer, 2021; Fabry and Bhowmik, 2021). For children with auditory processing deficits, there are computerized intervention programs such as Earobics (Concepts, 1997). Computerized interventions for auditory processing have also been studied and shown to be effective in older adults (Vaidyanath, 2015). Other commonly used computerized auditory training programs include the Computerized Learning Exercises for Aural Rehabilitation, clEAR (Tye-Murray et al., 2012), and Listening and Communication Enhancement, LACE (Sweetow and Sabes, 2007). Softwares such as FastForWord (Scientific Learning Corporation, 1998), Escuta Ativa, and others[1] have also been used for the pediatric population.

The current study was an attempt to combine technology and Carnatic Classical music, a genre very common in the Southern part of India, into a music-based module to improve auditory processing and cognitive abilities. Carnatic music was chosen, as most older listeners in this region are exposed to and are interested in this music genre. The study aimed to evaluate the effectiveness of a short-term computerized intervention module which delivers music-based stimuli and its influence on auditory processing and working memory in this population. This program, if found effective, can be adopted as a clinical protocol for rehabilitation of older adults with auditory processing and working memory deficits. In the context of the ongoing pandemic, this program can also be delivered to vulnerable populations in the safety and comfort of their homes, while utilizing the advantage that technology offers.

2. Materials and methods

This pilot study examines the effect of a computerized music-based intervention module on the auditory processing and cognitive skills of older adults. Approval from the Institutional Ethics Committee was obtained prior to the initiation of the study.

2.1. Development and administration of the computerized music-based intervention module

The musical style employed in the current study was Carnatic music, a style of Indian Classical music which emerged in Southern India and is being performed globally. Scales consisting of 8 notes are called *Raagaas* in Indian classical music. Carnatic music has 72 such fundamental scales or *raagaas*. These raagaas have three notes in common—first (Sa), fifth (Pa), and eighth (Sa of the higher octave). In this study, six such scales consisting of 8 notes were chosen—*Maayaamaalavagowla*, *Hemavathi*, *Vaachaspathi*, *Gaurimanohari*, *Thodi,* and *Kaamavardhini*. The latter five scales differed from the first by five, four, three, two notes, and one note respectively (Table 1). Each of the scales were played at a base scale

1 CTS Informatics. Escuta Ativa [Computer Software].

of D, and in tempos ranging from 120 beats per minute (BPM) to 480 BPM, with intervals of 60 BPM. This yielded a total of 7 tempos. A total of 42 scales were obtained (6 scales in 7 tempos). The music for the stimuli was played by a professional violinist trained in Carnatic music and recorded in a professional recording studio. All the stimuli were content validated by a professional musician with teaching and performing experience of more than 2 decades.

The interactive intervention module comprised of two types—note discrimination and tempo discrimination. The note discrimination task was constructed on three levels with increasing complexity, ranging from simple to a more challenging task. The first and second levels employed the twelve single notes, and the task was to discriminate between them. For generating the stimuli, scales played at a tempo of 120 BPM were chosen. The twelve octave notes were isolated using the Adobe Audition CS5.5 software. Each note was about 0.5 second long at this tempo.

The first level of note discrimination included pairs of single notes that differed from each other by 10, 9, 8, or 7 semitones, i.e., having large frequency differences, and easier to discriminate between (e.g., Ni1-Ri1 and Dha2-Ri1). Catch trials (pairs with the same note) were also included. This yielded 14 combinations and 10 catch trials. Participants were instructed to indicate if the two stimuli were 'Same' or 'Different'.

The second level included single notes that differed from each other by 6, 5, and 4 semitones. The notes were arranged in sets of three, with one of the notes differing from the other two presented (e.g., Ri1-Pa-Pa and Ni2-Ma2-Ni2). The second level had 21 combinations. Participants indicated which one of the three stimuli were different. The first and second levels of note training involved only single notes.

The third level of note discrimination included discrimination of entire 8-note scales (Sa-Ri-Ga-Ma-Pa-Dha-Ni-Sa) in the 240 BPM tempo. This tempo was chosen out of the seven tempos recorded, as it was neither too slow nor too fast to perceive. One scale was kept constant (*Maayaamaalavagowla*) and repeated twice, and the third stimuli was any one of the other five scales. For example, the first and second stimuli may be *Maayaamaalavagowla,* and the third *Thodi*, in which case the correct response would be '3'. The easiest to discriminate between were *Maayaamaalavagowla* and *Hemavathi*, as they differed by five notes, and the most difficult to discriminate between were *Maayaamaalavagowla* and *Kaamavardhini*, as they differed by only one note. The position of *Maayaamaalavagowla* among the three stimuli was randomly assigned by the software for each trial.

Tempo discrimination employed 8-note scales as stimuli, similar to the stimuli used in level 3 of note discrimination. This was constructed on two levels. The first level (Figure 1) involved two 8-note scales in which the tempo difference between each scale was large: 240 BPM and 300 BPM. For example, 8-note scales with tempos of 120 BPM and 360 BPM were paired, or tempos of 240 BPM and 480 BPM were paired, for discrimination. This level included all six *raagaas* and yielded 30 combinations and 33 catch trials. Participants were instructed to indicate whether the two were 'Same' or 'Different'. The second level (Figure 2) comprised of 8-note scales differing from each other by tempos of 120 BPM and 180 BPM. Three such scales were presented in which two were of the same tempo and one had a different tempo (e.g., 120 BPM-300 BPM-300 BPM and 360 BPM-240 BPM-360 BPM). The tempo differences between the scales were smaller than the first level and had 54 combinations. Participants had to indicate which one of the three scales presented was different.

The intervention module was structured using Apex (4.1.2), a software commonly employed in psychoacoustical experiments, developed by Francart et al. (2008). The inter-stimulus-interval was maintained at 500 ms. The subsequent set of stimuli was presented only after a response was obtained from the participants, allowing them adequate time to reflect on the differences between the stimuli presented. The participants were instructed to provide a response even if they were unsure whether their response was correct. Visual feedback was provided immediately after each response from the participant. The software flashed a green colored 'thumbs up' picture for a correct response and a red colored 'thumbs down' picture for a wrong response. The experiments in this software were defined using the XML format.

2.2. Participants

Eight right-handed participants (four females & four males) between the ages of 56 and 79 years (mean 63.87 years) were recruited for the study. Written informed consent was obtained from each participant prior to the commencement of the study. All the participants were native Tamil (South-Indian language) speakers and had an education of 10th grade or higher. Their hearing thresholds were screened using pure-tone audiometry to ensure audibility of 25 dB HL or better in both ears between 250 Hz and 2,000 Hz. Montreal Cognitive Assessment-Tamil (MoCA-Tamil) was administered to confirm normal cognitive functioning (score of $\geq$26). None of the participants had any neurological, psychiatric, or otologic conditions at the time of study or prior to the study. None of the participants were formally trained in any style of music.

2.3. Procedure

All assessments and training were carried out in a quiet room with ambient noise levels <40 dBA, as measured using the NIOSH SLM app (version 1.2.4.60) on an Iphone SE 2020 device. The assessment of auditory processing and cognitive skills were carried out twice before initiation of the intervention and once after completion. The two pre-intervention assessments were carried out with a gap of 2 to 4 weeks between them to ensure consistency of the baseline measurement.

The auditory processing tests assessed binaural integration (Dichotic Digit Test in Tamil), speech perception in quiet and noise (Tamil Matrix Sentence Test), temporal resolution (Gap-In-Noise), temporal patterning (Duration and Pitch Pattern tests), and temporal fine structure perception of the participants. Prior to administration of these tests, practice trials were provided to familiarize the participants with the task and the response expected.

The Dichotic Digit Test in Tamil, DDT-T (Sudarsonam and Vaidyanath, 2019) consisted of 30 trials in which two digits were

TABLE 1 Twelve octave notes and their nomenclature in Indian and Western Classical music.

Indian nomenclature	Western nomenclature	*Maayaamaalavagowla*	*Hemavathi* (5 notes different)	*Vaachaspathi* (4 notes different)	*Gaurimanohari* (3 notes different)	*Thodi* (2 notes different)	*Kaamavardhini* (1 note different)
Sa	Unison	**X**	**X**	**X**	**X**	**X**	**X**
Ri1	Minor second	**X**				X	X
Ri2	Major second		X	X	X		
Ga1	Minor third		X		X	X	
Ga2	Major third	**X**		X			X
Ma1	Perfect fourth	**X**			X	X	
Ma2	Augmented fourth		X	X			X
Pa	Perfect fifth	**X**	**X**	**X**	**X**	**X**	**X**
Dha1	Minor sixth	**X**				X	X
Dha2	Major sixth		X	X	X		
Ni1	Minor seventh		X	X		X	
Ni2	Major seventh	**X**			X		X
Sa (Next octave)	Octave	**X**	**X**	**X**	**X**	**X**	**X**

Each note differs from the subsequent note by a semitone. The scales chosen in this study and their differences from *Maayaamaalavagowla* have outlined. X denotes the notes comprised in each of the 6 scales chosen.

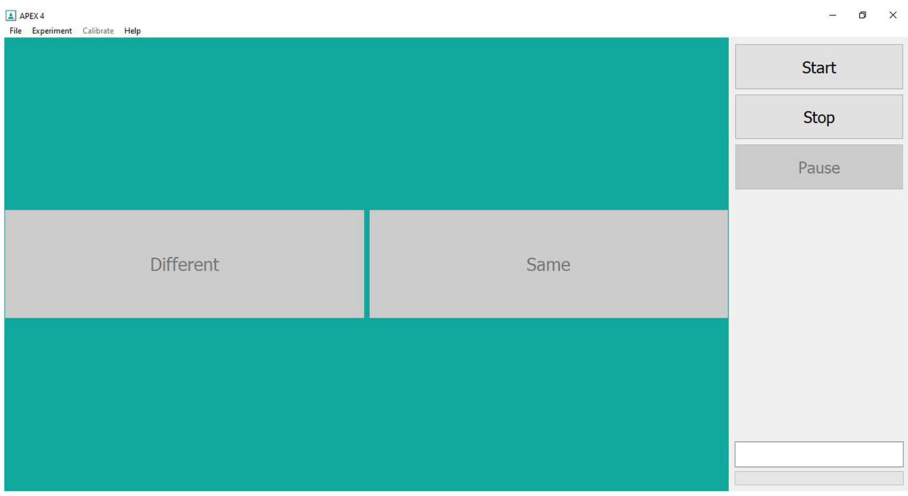

FIGURE 1
Screenshot of Level 1 of tempo and note discrimination module of the Computerized music-based training program on the Apex 4.1.2 software.

presented to the right ear and two to the left ear simultaneously. Participants were instructed to repeat all four digits, irrespective of the order or ear of presentation. A total score out of 30 was obtained each for single correct scores in the left and right ears, and double correct scores (correct responses from both ears). Norms developed by Sudarsonam and Vaidyanath (2019) in normal hearing young adults were used in this study.

Speech perception in quiet was assessed using Tamil Matrix Sentence Test, TMST (Krishnamoorthy and Vaidyanath, 2018) consisting of 24 lists with 10 sentences and with 5 words in each sentence. These sentences were derived from a base matrix of 50 words with 10 alternates for each word. The sentences were constructed to follow the grammatical structure of Tamil language. For example, the English translation of a Tamil Matrix sentence

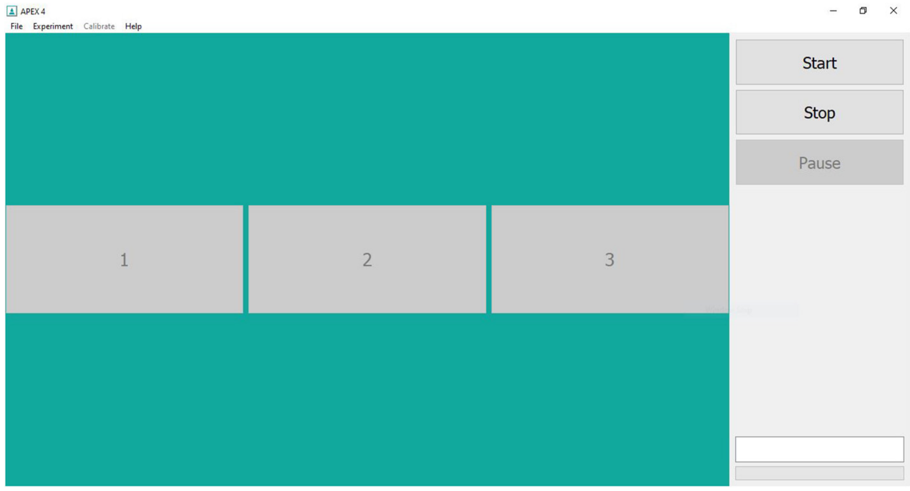

FIGURE 2
Screenshot of Levels 2 and 3 of note discrimination and level 2 of tempo discrimination modules of the Computerized music-based training program on the Apex 4.1.2 software.

is *Vasu bought ten wrong bags*. This prevented the participants from deducing entire sentences from context. For assessing speech perception in noise, these sentences were combined with multi-talker speech babble, developed by Gnanasekar and Vaidyanath (2019), at three Signal-noise ratios (SNRs): −5 dB SNR (speech 5 dB lower than noise), 0 dB SNR (speech and noise at equal levels), and +5 dB SNR (speech 5 dB higher than noise). Participants were instructed to repeat as many words of the sentence as they could. One list was presented per ear, for each condition of sentence recognition, and a total score out of 50 was obtained.

The Gap-In-Noise test, GIN (Musiek et al., 2005) consisted of 6s noise tokens in which small intervals of silences were embedded. GIN was administered separately in each ear. The participants were instructed to listen carefully and indicate when they heard a gap in the noise. This test yielded an approximate threshold of gap detected in ms (A.th) and a percentage correct score. Norms for the GIN test in Indian population, reported by Aravindkumar et al. (2012), were considered in the current study to judge whether a person passed or failed on the test.

The Duration Pattern Test, DPT (Musiek et al., 1990) and the Pitch Pattern Test, PPT (Musiek and Pinheiro, 1987) assessed temporal patterning skills separately in each ear. DPT consisted of two pure-tones of different lengths. These were arranged in sets of three (e.g., long-long-short, or short-long-short) and the participants were instructed to report the pattern heard during each trial. PPT involved two pure tones of high and low pitches. They were arranged in a similar manner as DPT and the participants were to report the pattern heard (e.g., high-low-low, or low-high-low). Scores out of 30 were recorded for both DPT and PPT. Cut off scores obtained by Gauri (2003) were used for DPT and the norms given by Tiwari (2003) were considered for PPT.

TFS1 (Moore and Sek, 2009) was a software developed to assess monaural temporal fine structure perception. Two fundamental frequencies (F0s) were chosen – 222 Hz and 94 Hz, similar to Moore et al. (2012), and were used to test Temporal Fine Structure perception abilities in the two ears separately. The software presented two stimuli and the participants were instructed to indicate which of the stimuli they found to be fluctuating. The software employed an adaptive procedure, which, if the participant was able to complete, yielded a result in Hz. This was the minimum frequency fluctuation that was perceived by the participant. If they were unable to complete the adaptive procedure, the software provided a score out of 40. For the purpose of statistical analysis, this was converted to Hz as delineated by Moore et al. (2012) and using the detectability index d' standard tables provided by Hacker and Ratcliff (1979). The norms obtained by Moore et al. (2012) were used in the current study.

The cognitive tests were carried out using the Smriti Shravan software (version 3.0) developed at the All India Institute of Speech and Hearing, Mysuru, India. This software is a customized platform to assess auditory and visual memory. The cognitive tests included forward digit span, backward digit span, running span, N-back, and operation span tests which assessed auditory attention, sequencing, and working memory. These tests were administered binaurally. The stimuli for the cognitive tests were the Tamil digits used in the TMST, which were fed into the software. In the forward digit span test, a series of digits were presented, and the participants were instructed to repeat the digits in the same order as was presented. In the backward span, the participants were instructed to repeat the digits presented in the reverse order of presentation. Initially, two numbers were presented and if repeated correctly, the software presented three numbers and so on. A one-up one-down adaptive procedure was followed and eight reversals were employed. The final span obtained was an average of the final three reversals.

In the N-back test, a 2-back task was carried out. This entailed the participants to repeat the penultimate number heard of the set of numbers presented. In the running span, 2-span was employed, and the participants were asked to repeat the final two numbers presented. Scores out of 20 were obtained for both N-back and running span tests. Operation span alternated presentations of a simple arithmetic equation (timed to complete within 3 seconds) and a number displayed on the screen. The participant had to

10.3389/fcomp.2023.1158046

indicate 'true' or 'false' to the solution for the equation and then memorize and repeat the numbers displayed, in the correct order. The software yielded a percentage score for the arithmetic task (secondary task), an All-or-None Score Weighted (ANSW) indicating the order and correctness of the response, and a Partial Credit Score Weighted (PCSW) indicating whether the correct numbers were repeated at all.

In addition to the tests of auditory processing and working memory, the Screening checklist for Auditory Processing in Adults, SCAP-A (Vaidyanath and Yathiraj, 2014) and Noise Exposure Questionnaire (Johnson et al., 2017) were also administered. A score of ≥4 out of 12 indicated a risk of auditory processing deficit on the SCAP-A. None of the participants were found to have any history of noise exposure in the past 12 months as noted on the Noise Exposure Questionnaire.

Following the second baseline assessment, the participants began the intervention module. The intervention was provided for 10 sessions (half an hour duration) over a period of 3 weeks. Each session consisted of 15 min each of note and tempo discrimination. The software provided the participants with feedback about the correctness of each response. Each level was repeated until the participants obtained at least 80% score, after which they progressed to the next level. If there was no improvement in performance in any of the levels during three consecutive sessions, training was discontinued. The number of sessions of training administered ranged from five to ten for the eight participants reported in this study (Table 2). Immediately after the intervention concluded, a post-intervention assessment of auditory processing and cognitive skills was carried out.

All the assessment and training stimuli were administered using calibrated Sennheiser HD280 Pro supra-aural headphones from a laptop with Intel(R) Core (TM) i5-6200U CPU processor @ 2.30 GHz and 8 GB RAM at an intensity level of 70 dB SPL. The pre-intervention and post-intervention assessment values were tabulated on a Microsoft Excel 365 worksheet. Statistical analysis was carried out using IBM SPSS Statistics 21 software. A Shapiro-wilk test was carried out to assess normality among the obtained data. As the data were found to be not normally distributed, and the sample size was also small, non-parametric tests (Wilcoxon Signed Ranks Test) were carried out to compare the performances of the participants before and after the intervention.

3. Results

The study examined the effects of a computerized music-based intervention module on the auditory processing and working memory abilities of older adults. This was a pilot study to an

TABLE 2 Session wise performance of the participants on note and tempo discrimination modules of the music-based intervention.

Session	Module	Participants (Performance in %)							
		1	2	3	4	5	6	7	8
1	Note	99.53 (1)	96.75 (1)	68.98 (1)	92.12 (1)	95.37 (1)	98.61 (1)	98.14 (1)	95.83 (1)
	Tempo	98.48 (1)	100.00 (1)	98.48 (1)	88.63 (1)	94.69 (1)	100.00 (1)	96.21 (1)	100.00 (1)
2	Note	93.51 (2)	89.88 (2)	65.74 (1)	91.07 (2)	83.33 (2)	94.64 (2)	97.02 (2)	94.64 (2)
	Tempo	98.80 (2)	100.00 (2)	83.33 (2)	86.11 (2)	82.40 (2)	98.14 (2)	91.66 (2)	97.22 (2)
3	Note	79.00 (3)	51.00 (3)	89.81 (1)	50.00 (3)	63.00 (3)	89.00 (3)	65.00 (3)	54.00 (3)
	Tempo	97.22 (2)	96.29 (2)	83.33 (2)	78.70 (2)	80.55 (2)	97.22 (2)	98.14 (2)	90.74 (2)
4	Note	78.00 (3)	59.00 (3)	50.00 (2)	53.00 (3)	65.00 (3)	90.00 (3)	80.00 (3)	80.00 (3)
	Tempo	92.59 (2)	98.14 (2)	93.51 (2)	71.29 (2)	79.62 (2)	98.14 (2)	94.44 (2)	95.37 (2)
5	Note	93.00 (3)	67.00 (3)	40.00 (2)	54.00 (3)	71.00 (3)	98.00 (3)	80.00 (3)	82.00 (3)
	Tempo	100.00 (2)	100.00 (2)	92.59 (2)	84.25 (2)	85.18 (2)	99.07 (2)	99.07 (2)	100.00 (2)
6	Note	88.00 (3)	82.00 (3)	32.00 (2)		73.00 (3)		91.00 (3)	80.00 (3)
	Tempo	96.29 (2)	100.00 (2)	89.81 (2)		74.07 (2)		100.00 (2)	100.00 (2)
7	Note	82.00 (3)	87.00 (3)			65.00 (3)			
	Tempo	98.14 (2)	98.14 (2)			90.74 (2)			
8	Note		93.00 (3)			75.00 (3)			
	Tempo		98.14 (2)			90.74 (2)			
9	Note					85.00 (3)			
	Tempo					84.25 (2)			
10	Note					73.00 (3)			
	Tempo					100.00 (2)			

Training level during each session is mentioned in brackets.

TABLE 3 Median and quartiles (Q1 & Q3) for scores obtained on the tests of auditory processing and working memory for the two pre-intervention assessments and Wilcoxon Signed Ranks test results comparing the two pre-intervention assessments.

Assessment			Pre-Intervention 1			Pre-Intervention 2			Wilcoxon Signed Ranks test	
			Median	Q1	Q3	Median	Q1	Q3	Z	p
DDT (max 30)	Single correct right		29.00	28.63	29.88	29.00	27.50	29.50	1.27	.20
	Single correct left		25.25	24.13	28.25	23.50	22.63	26.75	1.80	.07
	Doubt correct		20.00	16.00	25.75	15.50	14.00	23.50	1.26	.21
Speech in noise (max 50)	Right	-5 dB SNR	21.50	17.50	23.00	24.50	17.75	29.00	1.56	.12
		0 dB SNR	43.50	38.25	45.00	40.50	37.00	46.00	1.19	.23
		5 dB SNR	46.00	43.50	47.50	45.50	42.75	48.75	0.10	.92
	Left	-5 dB SNR	20.00	17.00	23.00	19.00	15.75	29.00	0.34	.73
		0 dB SNR	42.00	37.50	46.25	41.50	38.50	43.00	0.63	.53
		5 dB SNR	46.00	44.50	46.75	44.50	44.00	46.00	1.55	.12
GIN	Right	A.th (ms)	10.00	8.50	12.00	10.00	8.50	10.00	1.34	.18
		% correct	38.30	30.83	51.20	40.80	38.30	49.95	1.35	.18
	Left	A.th (ms)	10.00	8.00	10.00	10.00	8.00	10.00	0.44	.65
		% correct	41.65	38.30	55.35	44.95	35.38	54.15	0.07	.94
DPT (max 30)	Right		26.00	22.25	27.00	27.50	26.00	28.00	1.91	.06
	Left		26.00	25.25	27.00	27.50	24.00	29.00	0.51	.61
PPT (max 30)	Right		25.50	16.00	29.25	27.00	17.75	28.75	1.51	.13
	Left		27.00	19.25	29.75	26.50	18.25	28.00	1.13	.26
TFS1	Right	222 Hz F0	166.50	166.50	166.50	166.50	166.50	166.50	1.00	.32
		94 Hz F0	70.50	70.50	70.50	70.50	70.50	70.50	0.00	1.00
	Left	222 Hz F0	166.50	166.50	166.50	166.50	166.50	166.50	1.00	.32
		94 Hz F0	70.50	70.50	70.50	70.50	70.50	70.50	1.00	.32
Forward span (max 9)			6.00	5.00	6.00	6.00	5.25	6.00	0.81	.41
Backward span (max 9)			4.50	4.00	6.00	5.00	4.25	5.00	0.37	.71
N-back (max 20)			19.00	16.25	20.00	19.00	18.00	19.75	0.95	.34
Running span (max 20)			19.00	17.25	20.00	19.00	18.25	19.00	0.00	1.00
Operation span	ANSW (max 1)		0.22	0.00	0.71	0.31	0.00	0.68	0.00	1.00
	PCSW (max 1)		0.81	0.63	0.96	0.84	0.65	0.94	0.74	.46
	Secondary task (max 100)		75.00	36.11	91.66	83.34	40.28	94.43	1.21	.22

ongoing research work and hence the results of a limited sample have been reported. The Wilcoxon Signed Ranks Test was carried out between the pre-intervention baseline 1 and baseline 2, in order to verify that there was no change in the performances of the participants before the intervention was provided. No significant differences were found between the two baseline assessments in any of the tests of auditory processing and working memory (Table 3). Hence, for further analyses, the pre-intervention baseline 1 was considered for comparison with the post-intervention assessment.

The Wilcoxon Signed Ranks test comparing the difference in performance between the two ears showed a significantly better performance in the right ear compared to the left in DDT-T ($Z = -2.38$, $p = .017$). This is called the Right Ear Advantage and is commonly observed in right-handed individuals. Following the computer-based intervention, statistically significant differences were found between the pre and post intervention performances of DDT-T (right single correct and double correct scores). The right ear scores of all the participants had reached the norms while the left ear and double correct scores of two participants achieved norms.

With respect to auditory closure which was assessed using TMST in the presence of noise, performance in the right ear (−5 dB SNR and +5 dB SNR conditions) significantly improved post-intervention. There was no significant difference between the performance of the right and left ears.

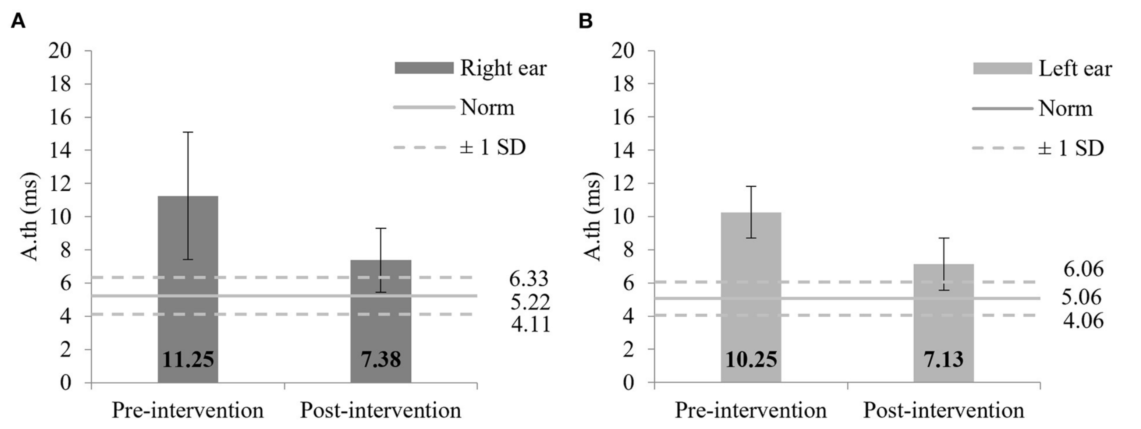

FIGURE 3
(A) Mean performance on GIN A.th pre and post intervention in the right ear; **(B)** Mean performance on GIN A.th pre and post intervention in the left ear. Error bars indicate standard deviations. Norms obtained by Aravindkumar et al. (2012) and ±1 SD in young normal hearing adults are shown.

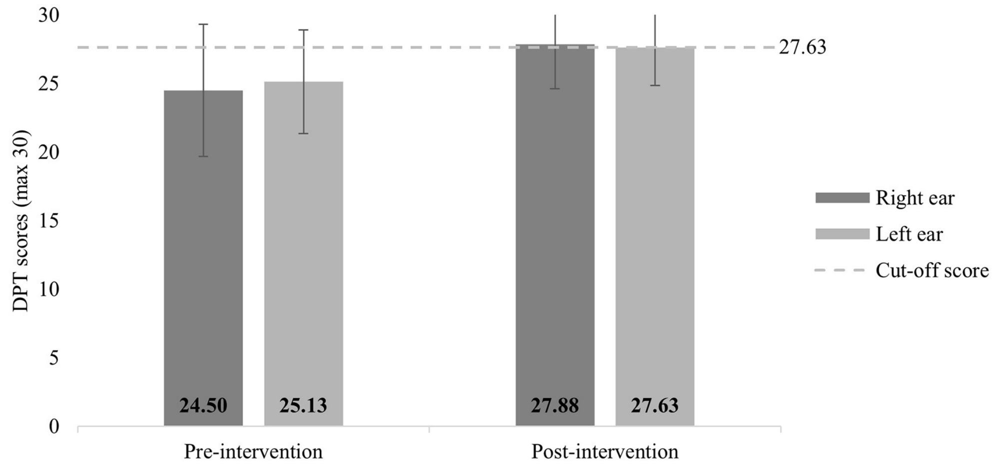

FIGURE 4
Mean performance on DPT pre and post intervention. Error bars indicate standard deviations. Cut-off scores obtained by Gauri (2003) for young normal hearing adults are depicted.

Temporal processing, as assessed using GIN, also showed a significant difference between the right and left ear percentage correct scores ($Z = -2.21$, $p = .027$), but this was not observed in the A.th. There was a significant improvement observed post-intervention in both ears' A.Th (Figure 3), with five participants performing normally. Percentage scores did not exhibit a significant difference pre and post intervention.

Temporal patterning which was evaluated using DPT and PPT showed a significant improvement in both ears post-intervention (Figures 4, 5). All but one participant had normal DPT scores in the right ear and left ear scores, post intervention. Six participants showed normal PPT scores in both ears following the intervention.

Temporal Fine Structure Perception which was evaluated using the TFS1 software employed two conditions—F0 of 222 Hz and 94 Hz. A significant improvement was found in the 222 Hz condition in both ears, with three participants achieving normal levels. Despite the 94 Hz condition not showing a statistically significant difference, two participants showed betterment, reaching the normal values following the intervention.

Among the cognitive tasks, the forward digit span task showed a significant difference post-intervention. Although all the tasks showed an improvement following the intervention, the difference was not statistically significant. The self-rating questionnaire of SCAP-A showed a significant improvement with reduction in scores, with all but one falling within the range denoting no risk of auditory processing deficit (Table 4).

4. Discussion

The current pilot study explored the use and effectiveness of a computerized music-based intervention module, in improving the auditory processing and working memory abilities in older

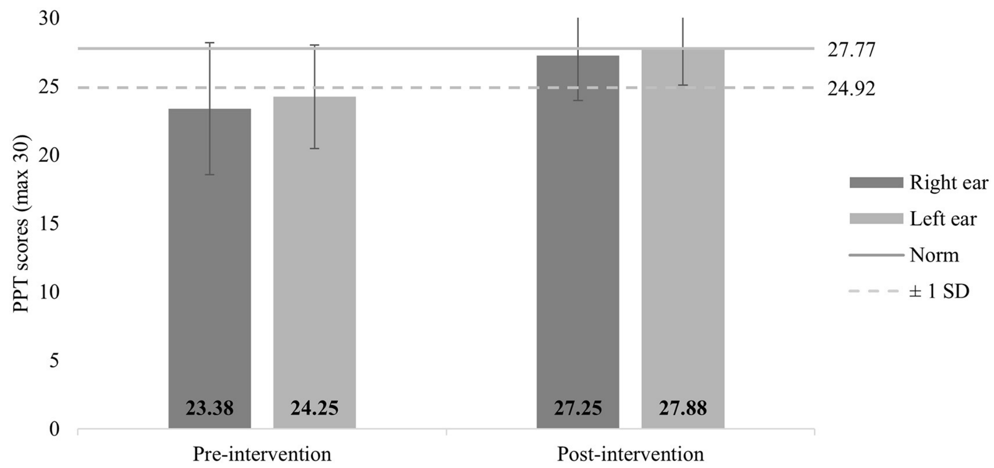

FIGURE 5
Mean performance on PPT pre and post intervention. Error bars indicate standard deviations. Norms obtained by Tiwari (2003) for young normal hearing adults are depicted.

adults. Two baseline assessments were carried out, before the short-term intervention was provided, to ensure that the auditory processing and working memory performances were comparable before the intervention. A post intervention assessment was carried out to observe if there was any improvement after the short-term intervention. The results revealed that all the auditory processing and cognitive tests, especially those assessing temporal processing abilities such as GIN, DPT, PPT, and Temporal Fine Structure perception showed improvements after the short-term music-based intervention.

It was found that in DDT-T, a difference between performances in right and left ears were present. This observation is known as the right ear advantage. The DDT-T is a test in which both ears receive different stimuli at the same time (dichotic). This ear difference can be explained by the majority of auditory neural fibers crossing over to the to the contralateral cerebral hemisphere. Therefore, the right ear is more efficient in perceiving verbal stimuli in right handed individuals who have dominant left cerebral hemispheres (Kimura, 1961). In the current study, all eight participants were right-handed, which explains the presence of right ear advantage. In addition, binaural integration performance was observed to have improved following the intervention. The sustained and focused attention for about half an hour during each session may have contributed to increased attention to the auditory stimuli presented to each ear, resulting in better DDT-T scores.

A significant improvement in speech perception in noise was observed in addition to all the tests assessing temporal processing such as GIN, DPT, and PPT. Given the contribution of temporal processing to speech perception in quiet and noise (Phillips et al., 2000; Schneider and Pichora-Fuller, 2001; Snell et al., 2002; Vaidyanath, 2015; Roque et al., 2019), it is expected that with a larger sample size, the effects on the speech perception scores will be larger. Participants also had better temporal fine structure perception following the intervention. However, greater improvement was observed in the 222 Hz F0 condition than the 94 Hz F0 condition. This was in agreement with the observations of Moore et al. (2012), as perception of temporal fine structure worsens with reduction in F0 (Moore and Sek, 2009). Pre-intervention, only two participants were able to complete the adaptive procedure in at least one ear, in either of the F0 conditions. Post-intervention, five participants were able to complete adaptive procedure in at least one F0 condition, in at least one ear. Of these, four had thresholds within normal range. These changes could be due to the manner in which the intervention module was constructed. The module incorporated temporal and frequency perception tasks extensively, beginning from simple discrimination and increasing in complexity. Level 2 and 3 of the intervention module required finer temporal and frequency discrimination between musical notes. The repeated presentations of the musical clips, which had minute differences in frequency and temporal aspects, may have resulted in the improvements in these auditory skills. A study conducted by Jain et al. (2015) that employed a computerized short term music training for normal hearing young adults showed improved speech perception in noise. A similar short term computerized music based training showed improvements in psychoacoustic skills such as frequency, intensity, and temporal perception as well, but these were not significant (Jain et al., 2014). Both training programs were short-term and involved Carnatic music. Music has long been recommended for maintaining and improving auditory skills (Strait and Kraus, 2011a,b; White-Schwoch et al., 2013), and has shown positive results even if it is a short-duration training (Lappe et al., 2011).

A significant change in forward digit span was observed, while minor individual improvements were observed in the other working memory tasks, although they did not reach significance. This could be because the N-back (2-back) and running span (2-span) tasks did not place adequate cognitive demands for the older adults. Possibly, if 3-back and 3-span were assessed, it may have necessitated a greater cognitive effort and could have judged the effect of the intervention on these processes better. Nevertheless, on the 2-back and 2-span tasks, the participants

TABLE 4 Median and quartiles (Q1 & Q3) for scores obtained on the tests of auditory processing and working memory across pre-intervention 1 and post-intervention assessments and Wilcoxon Signed Ranks test comparing the pre and post intervention performances.

			Pre-Intervention 1			Post-Intervention			Wilcoxon signed ranks test	
			Median	Q1	Q3	Median	Q1	Q3	Z	p
DDT (max 30)	Single correct right		29.00	28.63	29.88	30.00	29.50	30.00	2.04	.04*
	Single correct left		25.25	24.13	28.25	27.25	26.00	29.00	1.94	.05
	Doubt correct		20.00	16.00	25.75	25.00	20.25	28.25	2.11	.03*
Speech in noise (max 50)	Right	-5 dB SNR	21.50	17.50	23.00	25.00	22.25	30.75	2.37	.02*
		0 dB SNR	43.50	38.25	45.00	42.50	40.50	45.25	0.25	.80
		5 dB SNR	46.00	43.50	47.50	47.00	46.00	48.00	2.41	.02*
	Left	-5 dB SNR	20.00	17.00	23.00	26.50	20.50	29.75	1.86	.06
		0 dB SNR	42.00	37.50	46.25	43.00	42.25	46.25	1.49	.14
		5 dB SNR	46.00	44.50	46.75	48.00	46.00	48.75	1.58	.11
GIN	Right	A.th (ms)	10.00	8.50	12.00	7.00	6.00	9.50	2.38	.02*
		% correct	38.30	30.83	51.20	53.30	40.40	58.75	1.75	.08
	Left	A.th (ms)	10.00	8.00	10.00	6.00	6.00	9.50	2.03	.04*
		% correct	41.65	38.30	55.35	58.30	41.63	61.60	1.52	.13
DPT (max 30)	Right		26.00	22.25	27.00	29.00	28.00	29.75	2.37	.02*
	Left		26.00	25.25	27.00	28.00	28.00	29.00	2.53	.01*
PPT (max 30)	Right		25.50	16.00	29.25	28.50	26.50	29.75	2.10	.04*
	Left		27.00	19.25	29.75	29.50	27.25	30.00	2.12	.03*
TFS1	Right	222 Hz	166.50	166.50	166.50	97.33	41.90	164.96	2.20	.03*
		94 Hz	70.50	70.50	70.50	70.50	68.54	70.50	1.34	.18
	Left	222 Hz	166.50	166.50	166.50	126.14	44.98	135.28	2.53	.01*
		94 Hz	70.50	70.50	70.50	70.50	60.59	70.50	0.53	.59
Forward span (max 9)			6.00	5.00	6.00	7.00	5.25	7.00	2.07	.04*
Backward span (max 9)			4.50	4.00	6.00	5.00	5.00	5.00	0.44	.65
N-back (max 20)			19.00	16.25	20.00	20.00	19.00	20.00	1.72	.08
Running span (max 20)			19.00	17.25	20.00	19.00	18.00	20.00	0.37	.71
Operation span	ANSW (max 1)		0.22	0.00	0.71	0.50	0.37	0.63	1.62	.10
	PCSW (max 1)		0.81	0.63	0.96	0.81	0.65	0.83	0.18	.85
	Secondary task (max 100)		75.00	36.11	91.66	80.56	61.11	83.33	1.37	.17
SCAP-A (max 12)			5.00	4.00	6.75	4.00	3.25	4.00	2.41	.02*

*$p < 0.05$.

reported lesser fatigue and required lesser number of pauses between tasks, post-intervention.

Pre-intervention, five participants had SCAP-A scores >4, which indicated self-perceived auditory processing difficulties. Post-intervention, only one participant had a score >4. This was also found to be statistically significant and demonstrated that the participants experienced improved auditory processing in their daily lives after undergoing the computerized music-based intervention. The uninterrupted attention that the intervention module entailed over a period of a little more than half an hour per session may have contributed to improvement in the working memory tasks, and to the cognitive abilities required to perform the auditory processing tasks as well.

The improvements observed following the music based intervention is also supported by the OPERA hypothesis (Patel, 2011), according to which music may help improve speech perception if five conditions are fulfilled. The current protocol satisfies four of the conditions: **O**verlap of neural pathways between music and speech, **P**recision required to discriminate between the stimuli presented, **R**epetition of each level of the intervention until the participant reaches a score of 80%, and **A**ttention to the minute differences between the stimuli to achieve the 80% score. As this study made use of small clips of non-linguistic instrumental music, which was unfamiliar for the participants, the **E**motion criterion of the OPERA hypothesis may not be fulfilled. However, the extended OPERA hypothesis (Patel, 2014) suggests that instrumental music, due to having no linguistic component, does place a higher level of demand on auditory memory.

The advantages that this intervention module had over conventional auditory training was the usage of music and delivering it in a standardized and controlled manner using technology. For assessing the auditory processing and working memory skills of the participants, all the tests were administered in a structured manner by making use of calibrated output through headphones and computer softwares. This standard was maintained during the intervention notwithstanding the use of music, which is a diverse art form and includes multiple genres being widely practiced globally, such as classical, fusion, folk, and so on. Even in Indian Classical music, there is the Carnatic style and the Hindustani styles, both varying in theory and presentation. This control and standard could be achieved by incorporating the chosen style of music (Carnatic classical) into a software, Apex, which lent accuracy to factors such as intensity of presentation, duration of the intervention, and documentation of the participants' progress. The software recorded the responses of the participants and presented a percentage correct score at the end of each session.

A minor limitation of this study is that, although the handedness of the participants was verbally recorded, it was not assessed using a published tool. Future studies can incorporate a formal handedness assessment tool, which may aid in the interpretation of the binaural listening abilities (DDT-T). In addition, the results reported here for a limited sample size may be confirmed by the results in a larger sample of older listeners with auditory processing deficits, to be able to draw better conclusions.

In conclusion, this preliminary study found that a computerized short-term intervention protocol that makes use of Carnatic musical stimuli resulted in improvements in auditory temporal processing skills, which led to better speech perception in noise. If found effective on a larger population, the intervention protocol can be delivered to older and vulnerable individuals at their convenience, given that it is designed to be delivered using a computer software at a calibrated output. This may lead to better rehabilitation schedules and clinical resources being allocated and utilized. Prospective studies can explore the effectiveness of this computerized music-based intervention protocol on children with auditory processing disorders.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Ethics Committee, Sri Ramachandra Institute of Higher Education and Research. The patients/participants provided their written informed consent to participate in this study.

Author contributions

VR and RV designed the study with inputs from AU and SV. VR and RV and analyzed the data. VR carried out data collection and wrote the initial draft of the manuscript with extensive inputs from RV. All the authors have contributed to and reviewed the manuscript.

Funding

VR was funded by the Sri Ramachandra Founder-Chancellor Shri. N. P. V. Ramaswamy Udayar Research Fellowship.

Acknowledgments

The authors acknowledge the Faculty of Audiology and Speech Language Pathology at Sri Ramachandra Institute of Higher Education and Research, Chennai, India for lending their support and equipment for completing the study. All participants are also acknowledged for their participation.

Conflict of interest

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

Publisher's note

References

American Speech-Language-Hearing Association (2005). *(Central) Auditory Processing Disorders-The Role of the Audiologist.* [Position Statement], Available online at: www.asha.org/policy (accessed January 13, 2023).

American Speech-Language-Hearing Association (n.d.). *Central Auditory Processing Disorder. (Practice Portal).* Available online at: http://www.asha.org/Practice-Portal/Clinical-Topics/Central-Auditory-Processing-Disorder/ (accessed November, 20, 2022).

Aravindkumar, R., Shivashankar, N., Satishchandra, P., Sinha, S., Saini, J., and Subbakrishna, D. K. (2012). Temporal resolution deficits in patients with refractory complex partial seizures and mesial temporal sclerosis (MTS). *Epilepsy Behav.* 24, 126–130. doi: 10.1016/j.yebeh.2012.03.004

Besson, M., Chobert, J., and Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Front. Psychol.* 2, 94. doi: 10.3389/fpsyg.2011.00094

Boboshko, M., Zhilinskaya, E., and Maltseva, N. (2018). "Characteristics of Hearing in Elderly People," in *Gerontology*, eds. G., D'Onofrio, D., Sancarlo, and A., Greco (London: IntechOpen). doi: 10.5772/intechopen.75435

Concepts, C. J. C. (1997). *Earobics: Auditory development and phonics program.* Computer software.

Corey, R. M., and Singer, A. C., (2021). "Adaptive Binaural Filtering for a Multiple-Talker Listening System Using Remote and On-Ear Microphones," in *2021 IEEE Workshop on Alications of Signal Processing to Audio and Acoustics (WASPAA)* 1–5. doi: 10.1109/WASPAA52581.2021.9632703

Deshmukh, A. D., Sarvaiya, A. A., Seethalakshmi, R., and Nayak, A. S. (2009). Effect of Indian classical music on quality of sleep in depressed patients: A randomized controlled trial. *Nordic J. Music Ther.* 18, 70–78. doi: 10.1080/08098130802697269

Dunkel, L., Fernandez-Luque, L., Loche, S., and Savage, M. O. (2021). Digital technologies to improve the precision of paediatric growth disorder diagnosis and management. *Growth Hormone IGF Res.* 59, 1–8. doi: 10.1016/j.ghir.2021.101408

Fabry, D. A., and Bhowmik, A. K. (2021). Improving speech understanding and monitoring health with hearing aids using artificial intelligence and embedded sensors. *Semin Hear.* 42, 295–308. doi: 10.1055/s-0041-1735136

Francart, T., van Wieringen, A., and Wouters, J. (2008). APEX 3: a multi-purpose test platform for auditory psychophysical experiments. *J. Neurosci. Methods* 172, 283–293. doi: 10.1016/j.jneumeth.2008.04.020

Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6, 1–25. doi: 10.3389/fnagi.2014.00347

Gauri, D. T. (2003). *Development of Norms on Duration Pattern Test.* Mysore: All India Institute of Speech and Hearing. Available online at: http://203.129.241.86:8080/xmlui/handle/123456789/4479 (accessed May 8, 2023).

Gnanasekar, S., and Vaidyanath, R. (2019). Perception of tamil mono-syllabic and bi-syllabic words in multi-talker speech babble by young adults with normal hearing. *J. Audiol. Otol.* 23, 181–186. doi: 10.7874/jao.2018.00465

Hacker, M. J., and Ratcliff, R. (1979). A revisted table of d′ for M-alternative forced choice. *Percept. Psychophy.* 26, 168–170. doi: 10.3758/BF03208311

Hansen, J. H. L., Ali, H., Saba, J. N., Ram, C. M. C., Mamun, N., Ghosh, R., et al. (2019). "CCi-MOBILE: Design and Evaluation of a Cochlear Implant and Hearing Aid Research Platform for Speech Scientists and Engineers," in *IEEE EMBS International Conference on Biomedical & Health Informatics (BHI).* doi: 10.1109/BHI.2019.8834652

Jain, C., Mohamed, H., and Kumar, A. U. (2014). Short-term musical training and pyschoacoustical abilities. *Audiol. Res.* 102, 40–45. doi: 10.4081/audiores.2014.102

Jain, C., Mohamed, H., and Kumar, A. U. (2015). The effect of short-term musical training on speech perception in noise. *Audiol. Res.* 111, 5–8. doi: 10.4081/audiores.2015.111

Johnson, T. A., Cooper, S., Stamper, G. C., and Chertoff, M. (2017). Noise exposure questionnaire: a tool for quantifying annual noise exposure. *J. Am. Acad. Audiol.* 28, 14–35. doi: 10.3766/jaaa.15070

Kim, J. S., and Kim, C. H. (2014). A review of assistive listening device and digital wireless technology for hearing instruments. *Korean J. Audiol.* 18, 105–111. doi: 10.7874/kja.2014.18.3.105

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian J. Psychol.* 15, 166–171. doi: 10.1037/h0083219

Krishnamoorthy, T., and Vaidyanath, R. (2018). *Development and validation of Tamil Matrix Sentence Test.* Unpublished Masters dissertation, Sri Ramachandra Medical College and Research Institute.

Kumar, T. S., Muthuraman, M., and Krishnakumar, R. (2014). Effect of the raga ananda bhairavi in post operative pain relief management. *Indian J. Surg.* 76, 363–370. doi: 10.1007/s12262-012-0705-3

Lappe, C., Trainor, L. J., Herholz, S. C., and Pantev, C. (2011). Cortical plasticity induced by short-term multimodal musical rhythm training. *PLoS ONE* 6, 1–8. doi: 10.1371/journal.pone.0021493

Lyu, J., Zhang, J., Mu, H., Li, W., Mei, C., Qian, X., et al. (2018). The effects of music therapy on cognition, psychiatric symptoms, and activities of daily living in patients with Alzheimer's disease. *J. Alzheimer's Dis.* 64, 1347–1358. doi: 10.3233/JAD-180183

Moore, B. C. J., and Sek, A. (2009). Development of a fast method for determining sensitivity to temporal fine structure. *Int. J. Audiol.* 48, 161–171. doi: 10.1080/14992020802475235

Moore, B. C. J., Vickers, D. A., and Mehta, A. (2012). The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. *Int. J. Audiol.* 51, 715–721. doi: 10.3109/14992027.2012.690079

Musiek, F. E., Baran, J. A., and Pinheiro, M. L. (1990). Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. *Audiology* 29, 304–313. doi: 10.3109/00206099009072861

Musiek, F. E., and Pinheiro, M. L. (1987). Frequency Patterns in Cochlear, Brainstem, and Cerebral Lesions: Reconnaissance mélodique dans les lésions cochléaires, bulbaires et corticales. *Audiology* 26, 79–88. doi: 10.3109/00206098709078409

Musiek, F. E., Shinn, J. B., Jirsa, R., Bamiou, D.-E., Baran, J. A., and Zaida, E. (2005). GIN (Gaps-In-Noise) test performance in subjects with confirmed central auditory nervous system involvement. *Ear Hear.* 26, 608–618. doi: 10.1097/01.aud.0000188069.80699.41

Patel, A. (2011). Why would musical training benefit the neural encoding of speech? The OPERA Hypothesis. *Front. Psychol.* 2, 1–14. doi: 10.3389/fpsyg.2011.00142

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hear. Res.* 308, 98–108. doi: 10.1016/j.heares.2013.08.011

Phillips, S. L., Gordon-Salant, S., Fitzgibbons, P. J., and Yeni-Komshian, G. (2000). Frequency and temporal resolution in elderly listeners with good and poor word recognition. *J. Speech Lang. Hear. Res.* 43, 217–228. doi: 10.1044/jslhr.4301.217

Rodriguez, G. P., DiSarno, N. J., and Hardiman, C. J. (1990). Central auditory processing in normal-hearing elderly adults. *Audiology* 29, 85–92. doi: 10.3109/00206099009081649

Roque, L., Karawani, H., Gordon-Salant, S., and Anderson, S. (2019). Effects of age, cognition, and neural encoding on the perception of temporal speech cues. *Front. Neurosci.* 13, 1–15. doi: 10.3389/fnins.2019.00749

Sardone, R., Battista, P., Panza, F., Lozupone, M., Griseta, C., Castellana, F., et al. (2019). The age-related central auditory processing disorder: silent impairment of the cognitive ear. *Front. Neurosci.* 13, 619. doi: 10.3389/fnins.2019.00619

Schneider, B. A., and Pichora-Fuller, M. K. (2001). Age-related changes in temporal processing: implications for speech perception. *Semin. Hear.* 22, 227–240. doi: 10.1055/s-2001-15628

Schroeder, R. W., Martin, P. K., Marsh, C., Carr, S., Richardson, T., Kaur, J., et al. (2018). An individualized music-based intervention for acute neuropsychiatric symptoms in hospitalized older adults with cognitive impairment: a prospective, controlled, nonrandomized trial. *Gerontol. Geriatr. Med.* 4, 1–9. doi: 10.1177/2333721418783121

Scientific Learning Corporation, B. C. A. (1998). *FastForWord.* Computer Software.

Snell, K. B., Mapes, F. M., Hickman, E. D., and Frisina, D. R. (2002). Word recognition in competing babble and the effects of age, temporal processing, and absolute sensitivity. *J. Acoust. Soc. Am.* 112, 720–727. doi: 10.1121/1.1487841

Strait, D., and Kraus, N. (2011a). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front. Psychol.* 2, 113. doi: 10.3389/fpsyg.2011.00113

Strait, D., and Kraus, N. (2011b). Playing music for a smarter ear: cognitive, perceptual and neurobiological evidence. *Music Percept.* 29, 133–146. doi: 10.1525/mp.2011.29.2.133

Sudarsonam, R., and Vaidyanath, R. (2019). *Development of Dichotic Digit Test in Tamil.* Unpublished Masters dissertation, Sri Ramachandra Institute of Higher Education and Research.

Sung, H.-C., Chang, A. M., and Lee, W.-L. (2010). A preferred music listening intervention to reduce anxiety in older adults with dementia in nursing homes. *J. Clin. Nurs.* 19, 1056–1064. doi: 10.1111/j.1365-2702.2009.03016.x

Sweetow, R. W., and Sabes, J. H. (2007). Listening and communication enhancement (LACE). *Semin. Hear.* 28, 133–141. doi: 10.1055/s-2007-973439

Tilli, T. M. (2021). Precision medicine: technological impact into breast cancer diagnosis, treatment and decision making. *J. Person. Med.* 11, 1348. doi: 10.3390/jpm11121348

Tiwari, S. (2003). *Maturational Effect of Pitch Pattern Sequence Test.* Mysore: All India Institute of Speech and Hearing. Available online at: http://203.129.241.86:8080/xmlui/handle/123456789/4490 (accessed May 8, 2023).

Tsoi, K. K. F., Chan, J. Y. C., Ng, Y. M., Lee, M. M. Y., Kwok, T. C. Y., and Wong, S. Y. S. (2018). Receptive music therapy is more effective than interactive music therapy to relieve behavioral and psychological symptoms of dementia: a systematic review and meta-analysis. *J. Am. Med. Directors Assoc.* 19, 568–576.e3. doi: 10.1016/j.jamda.2017.12.009

Tye-Murray, N., Sommers, M. S., Mauzé, E., Schroy, C., Barcroft, J., and Spehar, B. (2012). Using patient perceptions of relative benefit and enjoyment to assess auditory training. *J. Am. Acad. Audiol.* 23, 623–634. doi: 10.3766/jaaa.23.8.7

Vaidyanath, R. (2015). *Effect of temporal processing training in older adults with temporal processing deficits*. Unpublished Doctoral Thesis, University of Mysore.

Vaidyanath, R., and Yathiraj, A. (2014). Screening checklist for auditory processing in adults (SCAP-A): Development and preliminary findings. *J. Hear. Sci.* 4, 33–43. doi: 10.17430/890788

Van Os, J., Delespaul, P., Wigman, J., Myin-Germeys, I., and Wichers, M. (2013). Beyond DSM and ICD: introducing "precision diagnosis" for psychiatry using momentary assessment technology. *World Psychiat.* 12, 113–117. doi: 10.1002/wps.20046

Wayne, R. V., Hamilton, C., Jones Huyck, J., and Johnsrude, I. S. (2016). Working memory training and speech in noise comprehension in older adults. *Front. Aging Neurosci.* 8, 1–15. doi: 10.3389/fnagi.2016.00049

White-Schwoch, T., Carr, K. W., Anderson, S., Strait, D. L., and Kraus, N. (2013). Older adults benefit from music training early in life: biological evidence for long-term training-driven plasticity. *J. Neurosci.* 33, 17667–17674. doi: 10.1523/JNEUROSCI.2560-13.2013

World Health Organization (2019). *Deafness and hearing loss*. World Health Organization. Available online at: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss (accessed July 25, 2019).

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 20 June 2023
DOI 10.3389/fcomp.2023.1111806

OPEN ACCESS

EDITED BY
Kjetil Falkenberg,
Royal Institute of Technology, Sweden

REVIEWED BY
Simon Order,
Murdoch University, Australia
Graham Sewell,
The University of Melbourne, Australia

*CORRESPONDENCE
Amanda E. Krause
✉ amanda.krause1@jcu.edu.au

RECEIVED 30 November 2022
ACCEPTED 30 May 2023
PUBLISHED 20 June 2023

CITATION
Krause AE and Fletcher H (2023) The purpose of radio and how it supports older adults' wellbeing. *Front. Comput. Sci.* 5:1111806. doi: 10.3389/fcomp.2023.1111806

The purpose of radio and how it supports older adults' wellbeing

Amanda E. Krause[1]* and Heather Fletcher[2]

[1]Department of Psychology, James Cook University, Townsville, QLD, Australia, [2]Faculty of Fine Arts and Music, The University of Melbourne, Melbourne, VIC, Australia

In Australia today, radio continues to draw large audiences, with high engagement among older adults. This research investigated how radio personnel and listeners regard the purpose of radio, and further how engaging with radio is perceived to influence listener wellbeing. Thematic analysis of semi-structured interviews with radio personnel ($N = 16$) and focus groups with older adult listeners ($N = 32$) suggest that the purpose of radio is to stay informed (e.g., news and information), for entertainment (e.g., music), and for perceived social purposes (e.g., communion, social connection, company, and companionship). Findings indicate there are implicit and explicit ways in which radio facilitates the wellbeing of their listenership. Explicitly, radio promotes mental health through broadcasts and programming, as well as exploiting the medium of radio as a public service for the community to call and rely on. Participants implicitly indicated that radio acts as a surrogate friend in their home; someone to keep them company and encourage connection to their greater community. Broadly, perceived relationships with radio programs and individual presenters, built and sustained over time through repeating listening, underpin the radio's ability to support listener wellbeing. These findings have implications for broadcasting practices as well as future work concerning how the radio might be used as a widely accessible tool for promoting quality of older life.

KEYWORDS

radio, broadcasting, media, wellbeing, social surrogacy

1. Introduction

Research indicates high engagement in music listening by people of all ages; yet studies have tended to focus understanding the listening behaviors of adolescents and young adults (e.g., Albarran et al., 2007; Ferguson et al., 2007; McClung et al., 2007) and on digital technologies (e.g., Krause and North, 2016). Therefore, less is known about older adults' everyday radio listening practices and how engaging with the radio might have benefits for people's wellbeing (Krause, 2020).

In Australia today, radio continues to draw large audiences—Australian Communications Media Authority (2020) findings indicate that 95% of Australians have a radio and that 78% of folks had listened to FM radio in the past seven days when surveyed. They also reported that Australians 45 and older are more likely to listen to radio than those aged between 18 and 44 (Australian Communications Media Authority, 2020). Additional figures suggest that community radio reaches over 5.1 million Australians weekly (Community Broadcasting Association of Australia, 2022) and that commercial radio reaches 68% of Australians 65 years of age and older (Shepherd, 2022). Average weekly listening hours are 15 and nearly 13 for community and commercial radio, respectively (Community Broadcasting Association of Australia, 2022; Shepherd, 2022). Indeed, radio listening times has been reported to have increased due to COVID-19 (Hasnain et al., 2022); however, even pre-COVID-19 statistics indicated high engagement by older adults (e.g., 29% of ∼10.9 million Australians listening to commercial radio are 55 years or older—Community Broadcasting Association of Australia, 2019).

The radio is an important medium of communication in Australia, transcending geographical and social boundaries (Foxwell, 2012; Meadows, 2013; Oliveira, 2013; Watson, 2016), to offer news and information, music programming, and opportunities for social exchange (Bednarek, 2014; Ames, 2016; Ewart and Ames, 2016). In Australia, listeners have many stations to choose from: in addition to the Australian Broadcasting Corporation (ABC) and Special Broadcasting Service (SBS) as public broadcasters, there are more than 300 commercial AM and FM stations and more than 450 community stations (www.infrastructure.gov.au). Via on-air offerings, the radio informs, educates, and empowers listening audiences (Watson, 2013). Moreover, radio stations, and community stations in particular, are able to offer diverse, and tailored content for different communities (Meadows and Foxwell, 2011; Foxwell, 2012; Hasnain et al., 2022). This includes broadcasting targeted to older adults: Golden Days Radio in Melbourne, Australia, for instance, markets itself as a "senior citizens' radio station", playing music from "the 1920s to the 1950s" (Foxwell, 2012, p. 170). It is perhaps the ability for local radio, and especially community radio, to customize their approach to meet the needs of their target community that promotes strong listener engagement. Indeed, stations can reflect the cultural climate of the community, provide easy access to citizen participation (Siemering, 2000), and facilitate community rights (Moffat et al., 2023).

Previous researchers have highlighted that the radio can broadcast health messages, thereby playing a role in health promotion and community wellbeing (Forde et al., 2009; Ewart, 2011; Smith et al., 2011). It assists further during community emergencies, providing crisis communication during natural disasters as well as political crises (North and Dearman, 2010; Hugelius et al., 2019; Rodero, 2020; Laskar and Bhattacharyya, 2021). Moreover, community radio is considered to be a more trustworthy medium than other technologies and mainstream media (Guo, 2015). Additionally, there is some evidence that engaging in the radio supports individual wellbeing (Krause, 2020; Hasnain et al., 2022). Community radio station volunteers, for example, gain a sense of purpose and identity by volunteering (Order and O'Mahony, 2017), are able to express creativity and enable maintenance of language and culture (Krause et al., 2020), and experience a sense of belonging, thereby promoting social relationships and relationships (Vuuren, 2002; Maina, 2013; Vidal, 2019). The radio promotes social connections between listeners, presenters, and the wider community (Meadows and Foxwell, 2011; Oliveira, 2013). For instance, listeners can become more active in their communities (Milan, 2008) and attend social events advertised on air (Keough, 2010). Ewart (2011) suggested that the radio might provide companionship for isolated and disconnected individuals, a finding that Krause (2020) discussed in terms of social surrogacy. Media can function as social surrogates, such that when individuals engage with media such as music listening and the radio, their experiences engender feelings of connection and empathy (Schäfer and Eerola, 2020; Schäfer et al., 2020).

Regarding older adult listeners in particular, Krause (2020) found that engaging with radio was a common everyday activity with strong, individual preferences and practices. Regardless of how older adults engaged with the radio (e.g., listening at specific times vs. constantly having the radio on), participants highlighted positive outcomes to radio listening which appeared to influence mental health and wellbeing, particularly mood regulation, company and comfort, relaxation, passing the time, and reminiscence. Moreover, these benefits mirrored those established in research on music listening (e.g., for a review see Krause et al., 2018), suggesting that the findings concerning the wellbeing benefits of listening to music might also apply to listening to the radio (Krause, 2020, p. 8). Yet, in addition, some participants pointed to the voice as contributing to their listening experiences—that "hearing people speaking" promoted feelings of company and companionship (Krause, 2020). The role of the voice calls into question the role of the presenter regarding how radio engagement might promote listener wellbeing.

While Krause's (2020) findings are useful to begin to consider how radio engagement might promote wellbeing in older age, the study is limited by its small, exploratory scope. Thus, there is a need for a more explicit consideration of the role of radio in promoting wellbeing for older adults. Moreover, given the findings concerning companionship and community amongst listeners and presenters, Krause (2020) suggested that future research might explore the perspectives of both listeners and radio personnel.

Therefore, this research examined the perceived purpose of radio and considered whether (and how) radio stations and radio presenters play a role in promoting the wellbeing of older listeners. The first research question investigated what radio personnel and older adult listeners considered as the purpose of radio. We anticipated that people would suggest various purposes, perhaps underpinned by their own listening preferences and practices. In line with prior work, we predicted that people might elaborate on providing information and educating audiences (Watson, 2013), opportunities for dialogue and social exchange (Bednarek, 2014; Ewart and Ames, 2016), and entertainment (e.g., through preferred programming, such as music—Krause, 2020). We did not know if people would refer to health promotion or wellbeing, although we anticipated that ideas of community and companionship (Krause, 2020) might also arise.

The second research question probed the purpose of the radio with specific regard to the role the radio might play in promoting listener wellbeing. While the presented research was exploratory in nature, we anticipated that people's suggestions may align with previous work on the experiences of individuals, such that some of the outcomes highlighted by Krause's (2020) participants might arise. For instance, while some wellbeing benefits of radio listening mirrored those concerning music listening (e.g., mood regulation, company and comfort, relaxation, passing the time, and reminiscence), the fact that Krause's (2020) participants pointed to the voice as contributing to their listening experiences suggests the possibility for additional benefits to be identified. Additionally, we anticipated that people might mention the capacity radio has for large-scale communication or its role for the community, such as radio's use during crises (North and Dearman, 2010; Rodero, 2020) or health promotion (Smith et al., 2011).

2. Method

2.1. Design

A qualitative enquiry approach was used. Individual, semi-structured interviews with radio personnel and semi-structured focus groups with older radio listeners were conducted. In both the interviews and focus groups, a conversational style was employed to explore the participants' experiences (Bhattacharya, 2017). The James Cook University Human Research Ethics Committee approved this research (ID: H8022).

2.2. Participants

Individuals volunteered to participate in the study. Recruitment techniques involved direct invitations and snowball sampling.

2.2.1. Radio personnel

Sixteen individuals involved in presenting live, music-based or talk-based radio programming at one of six public and community radio stations broadcasting in Melbourne, Australia aged 30–81 (M = 60.80, Mdn = 61, SD = 15.29) were interviewed. Of these radio personnel, six (37.50%) identified as female and ten (62.50%) as male. We targeted these six radio stations based on previous research in which older adults residing around Melbourne shared their station preferences (Krause, 2020).

2.2.2. Listeners

A total of 32 older adults took part in the listener focus groups. This included 25 (78.10%) individuals who identified as female and seven (21.90%) who identified as males. Of these 32 participants, 28 reported being between 64 and 82 years of age (M = 72.46, Mdn = 72.50, SD = 5.73). The majority (26 of 32) of these older adults reported listening to the radio daily; and, on average, the sample listened to 4.48 h of radio daily (Mdn = 4, SD = 4.81). When asked about preferred stations, these participants listed between one and five stations (M/Mdn = 3). Participants referenced the stations targeted in this study in terms of the radio personnel participants as well as additional local, national, and international stations broadcasting both music-based and talk-based programming. The majority of stations were public and community radio stations (e.g., the local and national news stations and classical music stations were most common), though some commercial stations were also cited.

2.3. Procedure

Participants received the participant information sheet via email and provided written consent prior to the interview or focus group commencing. On the consent form, they were asked to report their gender, age, how often they listened to the radio (1 = *Never*, 5 = *Daily*), the average number of hours listened per day, and which radio station(s) they listen to (free-text response). Consent was reaffirmed verbally at the beginning of each interview and focus group. Interviews were individually scheduled to suit the radio personnel and were conducted online via Zoom (lasting between 11 and 53 min). Listeners were invited to select a focus group that was scheduled at a convenient time for them. Focus groups were done in person and online, lasting between 12 and 45 min.

2.4. Materials

We posed two key questions to both the radio personnel and older adult listeners during the conversations. These were: (1) What do you see as the purpose of radio and (2) (How) Does radio play a role in promoting listener wellbeing? As such, the participants were asked to reflect on their own engagement with the radio—be it presenting or listening. The resulting conversations formed the data collected and analyzed.

2.5. Data analysis

Individuals participated as part of a wider study concerning older adults' radio engagement. The present research details the data concerning the purpose of the radio and its potential role in promoting listener wellbeing. All interviews and focus group conversations were audio-recorded; verbatim transcriptions were created from these recordings. A thematic analysis was used to analyze the data. Braun and Clarke (2006, 2013) six steps were followed, such that codes were generated after time was spent becoming familiar with the data. Using a reflexive and recursive approach (Braun and Clarke, 2019), initial codes were developed by considering semantically similar content and implicit concepts within participant responses. Themes were then identified and reviewed by considering how related codes clustered together.

We also acknowledge that our experiences and knowledge are brought into the analysis process. The second author, who led the analysis, volunteers for a community radio station. She used her lived experience of presenting on-air "knowingly" when analyzing the data (Braun and Clarke, 2013; Hemming et al., 2021). The first author served as a critical friend (Sparkes and Smith, 2014), drawing on her research experience when working with the second author to refine the higher-order themes and sub-themes to best represent the data with respect to the guiding research questions. Lastly, the final themes were labeled and reported. To uphold participant confidentiality, data extracts drawn from the corpus are simply labeled as coming from either radio personnel or listeners.

3. Results

In examining the perceived purpose of radio and the role it might play in facilitating listener wellbeing, responses to the latter were more implicit than explicit. In other words, participants' responses regarding their perceptions of the purpose of radio indicated that, in a roundabout way, radio-listening supported their wellbeing. Given this strong evidence of implicit wellbeing support, the results are considered in tandem and focus more on the purpose of radio whilst identifying how these responses speak to wellbeing.

3.1. Radio serves multiple purposes

Both listeners and radio personnel noted that radio has multiple purposes. In separate interviews, one listener succinctly said, "*it's information, entertainment, and connection, and companionship*" and one radio presenter equally stated, it's "*information, entertainment, companionship… I think it's an even mix of those three things, and you have to be ready to provide all of them whenever they're needed*". This multi-purpose aspect of radio appears to be determined by when and why listeners engage with it as they consume radio in various ways at different times of the day. For example, a male listener said, "*I like to hear the news when I get up in the morning… I do it partly for company. I listen to the radio when I'm cooking a meal, sometimes to music and sometimes conversation. I like radio mainly for news and also for music*". Some listeners talked about their complete reliance on radio for their information and entertainment needs, as one woman stated, "*basically I rely on radio for everything and yes, it's my total source of news, music and entertainment. So everything, it's got everything I need*".

Radio presenters identified themselves as responsible for facilitating the diverse medium of radio by delivering informative content in an engaging way or educating listeners about the music they are playing, for example. One executive producer of a radio program said radio's purpose is "*to inform and entertain. It's this duality that exists within a public broadcaster… so we try to create immersive content that people can kind of lose themselves in*". One presenter even highlighted the purpose of radio on a larger scale, believing that radio has an ethical responsibility to inform and connect its listeners: "*Hopefully we are in the business of building a better world and making a better generation for the next part of the world's history than the one that we're involved in at the moment*".

3.1.1. Information

Consuming radio for the purpose of staying informed was a highlighted theme among radio listeners. Some were specific about the information they were interested in, for example as one woman said, "*I like politics. I like listening to the news to understand what about the political events and everything that's going on in the country*", whereas some listened more broadly and regularly to stay informed, "*I do listen to radio to get information, to get more updates and also listening to radio again is so portable and easy…*". One woman spoke about radio's origins as a source information, "*the main purpose in the first place was when it started off was communication and that was for news to tell the people what was going on in the world. I think it's changed over the years*". Both men and women talked about listening for mental stimulation, both from an educative standpoint, "*I like to think that I'm feeding my mind with something that's worthwhile*", but also to maintain focus, such as when driving: "*If I listen to something like Richard Fidler [an Australian radio host] that grabs my attention, then my mind doesn't wander and I don't feel sleepy if I'm driving long distances*". Radio personnel identified the ethical responsibility they have to do rigorous journalism and provide accurate news as a public service. As one presenter acknowledged, "*we need to be able to provide authentic information, not shy away from things, but at the same time, acknowledge that life is a big mixture of things. And what's more we're going to do all those things for free*". This was also reflected in listener feedback where one woman commented that radio has a responsibility to deliver, "*intelligent, well-based information. If they're communicating effectively with a lot of people… it's so important*".

3.1.2. Entertainment

Listeners also considered radio to be a source of entertainment, as one noted: "*It's more entertainment than just straight-out communication and news*". Another listener simply stated, "*I do listen to radio to get entertainment, music, and trying… to keep myself more entertained*". Music was frequently referred to when discussing entertainment. While listening for news or information was often identified as more active, listening to music was considered both active and passive depending on what the participant was doing. One woman said, "*I do find that music is a good distraction if you are doing something like housework. I listen to the music and I'm entirely listening to that and doing other things at the same time. So, it's good that way, I get things done like ironing - I listen to the radio. So, it gets you through boring tasks*". Some people talked about how music listening influenced their mood, as one woman said, "*I love classical music, it's the music that enhances my mood*". Another woman identified that she listened to music for the sake of changing her mood for the better, "*Sometimes when you're in a low mood, you want to listen to music so that you can pull yourself [up]*". Radio personnel appear to be aware of the impact music has on the mood of their listeners, as one station secretary noted, "*people are feeling a bit down and then you've got some nice piece of music playing, that sort of immediately lifts their spirits which music I think is all about*". A few presenters on music stations highlighted how they take into account their listenership when programming to facilitate not only their entertainment, but also their education of music. One presenter acknowledged the power music has to influence people and create inner-calm as well as promoting comfort:

> Its first purpose is to entertain and to be able to be enjoyed by listeners. In that way, if you draw out from that the notion of that entertainment leading to insight and understanding and depth of awareness, that's what serious music does and so that then has a wider purpose of bringing healthy influences into society and making people feel more at ease with themselves and comfortable in their relationships with other people.

3.1.3. Social support

Radio was also considered to be important for broader social reasons, specifically communion, social connection, company, and companionship. These themes appeared to have significant implications for listener wellbeing as participants frequently referred to them when discussing their "relationship" with presenters, programs, and the radio. With regards to communion, radio personnel, in particular, considered this to be the purpose of community radio, as one station manager stated:

> The purpose of community radio broadly is to provide a voice, a forum, an outlet, a means of communication for those elements of whatever field that you happen to be a part of… being community radio, it's our duty, it's our responsibility to be tied into the local community, and that's what we try to do as much as possible.

Radio presenters also perceived radio as a medium for facilitating social connection, and in that manner the listeners were part of their program. According to one male presenter, "*for our listeners and a lot of listeners, it has always been a place where they can go to feel connected… you almost become part of their show*". This is reflected in listener feedback where being informed about their community and beyond is a strong motivation to tune in. As one man said, "*I think overall, it connects people… I think that connection is something that radio, especially in the early days did, that just broadened people's worlds enormously*". Some female listeners spoke about listening to expand on their views of the world: "*It gives you a wider view because you do become a little narrower as we age on our own, so it keeps us in touch*", while others identified how radio can elicit an almost empathetic response to the community: "*Good radio makes you think. And it makes you think in a social context too, not just in terms of your own little being*". Consideration was also made of the limited mobility of some older adults, and how radio provides a vital service for those "*who can't get out much to connect with other people*". One listener felt connected simply by knowing that other people were listening to the same program she was at any given time, "*it's that common thing that you're not listening to something like a CD which only you are listening to, you're actually receiving the same information or music as a lot of other people are at the same time*". Indeed, real-time engagement with radio was highlighted by several listeners as different from their experience of listening to podcasts or CDs.

Most of the listeners considered the radio to be a form of company. This was evident in how often they listened to the radio and why they would tune in. Some listeners, most of whom were women, have the radio on all day, as one listener stated, "*first thing I do is put the radio on*". In fact, some listeners talked about using multiple radios at home. This facilitated hearing the radio throughout the house ("*I used to have maybe four or five radios in the house all tuned to the same station so if I walked from one room to the other, I could still hear the same music*") as well as tuning into different stations ("*I have it on in three rooms in the house and I'll have it on whenever I'm in a room*"). In this regard, listeners acknowledged the radio as "*background noise*" that could provide both comfort and stimulus. According to one man, "*so that I'm not alone with my thoughts as I'm loading the dishwasher or cooking dinner or doing the housework*".

Further on background sound, multiple listeners also pointed out that they live alone, in some cases due to the bereavement of a partner or pet, and so they found the radio could "*fill the silence*" in their homes. One woman said, "*I listen to the radio for company. My partner died six years ago and we used to listen to the radio together and now it's always very quiet, so the radio is on as my company. Yeah, so I don't feel alone*". Notably, in these comments listeners did not focus on the nature of the radio content, but rather that having the radio on simply was a contrast to the quietude of living alone. For example, one woman stated, "*if I'm around the house, I like to have it on. It's sort of a bit of noise, nice noise too. It can get a bit quiet otherwise*". This remedy for silence was not just during the day but also overnight, as one woman commented, "*if I'm in bed I'm trying to go to sleep, or I've woken up and can't go back to sleep the radio is the perfect way to go back to sleep… rather than silence, I can't go to sleep in silence*".

Some listeners referred to the radio in a manner congruent with having an actual living being who was in their home. According to one woman, "*I have it on from seven o'clock in the morning until I go out. If I go out, I turn it off so that I don't feel as though there's someone here when I get home*". In contrast, another woman liked to leave the radio on when she went out so that, "*when I come back and go upstairs, it's like a welcome*". The perception that radio provides company was also tied to the presenter who could be perceived as an actual person to engage with: as one listener stated, "*with a lot of very old people it's another voice in the house. It makes them feel that there's somebody else there they can communicate with*". Interestingly, radio personnel seemed to understand how listeners consume radio and perceived it as company. This was either demonstrated through generalized statements: "*People are listening to the radio because they're either traveling or moving about and you're there as company*", or specifically in reference to receiving listener feedback: "*You have the old ladies calling saying, 'I'm all alone and you keep me company'*".

The variety and consistency of presenters and programs also provide a degree of companionship. From a broad perspective, one woman said, "*it's like having a friend come in all throughout the day, a different friend at different times of the day*". Radio personnel share a similar perspective: according to one presenter, "*we have to accept that we are perhaps their best friend for that particular – for the duration of our program. I think we should treat them like that with respect and so on*". With regards to tuning into specific programs at specific times, participants would talk about presenters they liked and the impact of knowing when they would be on the air: "*it's just like a known friend. I know what it's going to be like it if I turn it on. I know what to expect*". Sometimes this is specific to the presenter, where listeners connect with a particular presenter and even feel like they know them. For this reason, changing presenters on radio programs can impact listeners, as one woman stated "*it annoys me when they change their announcers. Because it's like, hey, I like that woman. Why did you get rid of her? She was my friend*". This is also the case with changing or canceling a program, "*I think you get this attachment to those programs too. When they're gone you find it very hard to reconnect with it*". Radio personnel appeared to be aware of the importance of consistency in what they deliver and how it impacts listeners. According to one station manager, during COVID-19 their station endeavored to:

> …keep everybody on air, or keep the voices that people knew on air so that that regularity of people's lives, even that small part of hearing the same person at the same time on the same day was maintained. Because it meant - it really did, the feedback was unequivocal - it really meant so much to people that we were still there in the same format, so that sense of regularity was there in people's lives at a time when so much changing and there was so much uncertainty and fear.

3.1.4. Explicit wellbeing support

As can be seen in the above themes, the influence of radio listening on mental health was often implicit, particularly with regards to eliciting a perceived sense of connection and providing companionship. Indeed, some radio personnel identified that "engaging" radio could facilitate wellbeing in a more roundabout way. According to one station manager, "*I don't know that [wellbeing] is something that we actively think about in terms - I think that's something that I'm stoked our content does and I think is a really great outcome of the radio*". However, as anticipated, participants also acknowledged that radio explicitly supported wellbeing through broadcasting information about the community. This was expressed, for example, with regard to disasters: one listener stated, "*with the Tongan earthquake, it was the AM radio people that got that information first*", and another commented: "*also now in Australia with the fires and floods and so on, the radio is the key*". Some listeners noted that the content of some broadcasts and news events could be challenging for listeners and believed that presenters were delivering this information sensitively while sometimes providing contact details to services such as Lifeline for people who required further support. As one woman said, "*even the way things are sensitively talked about now which address things like family violence and PTSD and those kinds of issues is now - I think it's a regulated requirement that the helplines are mentioned of course*". Some female listeners also indicated a preference for hearing the news rather than watching the news on TV for their own mental health:

> Female 1: I've actually got this thing that I don't want to see the news... I don't want to see the visuals. I find it too distressing. It's probably my age, I think. I just don't want to see what we could all see on TV.
>
> Female 2: That's the thing about the radio. You can actually use your imagination. That's right.

Moreover, radio explicitly supported wellbeing through specific programming such as having psychologists on air doing segments about mental health. As one presenter noted, "*We do lots of reports on wellbeing. For example, we've been doing for the last six, seven weeks a segment with a psychologist and that is for the focus was COVID and Australia being in lockdown... we definitely try to do lots of health-related topics and psychological topics*".

Lastly, some listeners talked about how radio provided a service for people who had problems such as getting access, needing assistance, or seeking clarification, as one listener pointed out: "*it becomes like a community service then for people who are feeling they're just banging their head against a brick wall and they think, right, I'm going to ring the radio station, I might get some action that way*". Another listener commented that some radio programs even reach out either offering this service or generating awareness, "*They say, look, if you've got something that is of concern to you and people around you in your area, please let us know and we'll try and do something on it*". Radio personnel similarly acknowledged that listeners being able to call into a radio station was meaningful for this reason, but also because it provided an opportunity to connect with someone, as one community station volunteer identified, "*I took three calls between 6:15 and 6:35 from a 92-year-old who was obviously on his own and just wanted to talk. You think, well that's wellbeing. So partly it's the music, partly it's the fact that he can ring up and talk to someone*".

4. Discussion

In examining the perceived purpose of radio and its potential to promote wellbeing, the primary purposes that both listeners and radio personnel identified aligned our prediction (RQ1) as well as with prior research categorizing radio's offerings as news and information, entertainment, and social exchange opportunities (e.g., Bednarek, 2014; Ames, 2016). Additionally, however, the present study's top-level themes (information, entertainment, and social support) concerning radio's purpose underpin the discussions of how radio might promote listener wellbeing both explicitly and implicitly (RQ2). Listeners and personnel also had similar answers which indicated a consensus from these two perspectives.

Participants identified that radio is multi-purposed, such that listeners said radio offered different things at different times. Moreover, radio simultaneously achieves these purposes of informing, entertaining, and stimulating one's social needs without them being mutually exclusive. When asked if or how radio supports wellbeing, participants spoke about its role in delivering broadcasts during times of crisis and programs specific to health. While this mirrors prior work illustrating the use of the radio during crises (e.g., North and Dearman, 2010; Rodero, 2020; Hasnain et al., 2022) as well as its ability to assist with health promotion (e.g., Smith et al., 2011); there was additional strong evidence that radio implicitly supports mental and social wellbeing based on how listeners and personnel discussed their relationship with the radio and their perception of its purpose. For example, listeners conveyed multiple positive wellbeing outcomes while acknowledging how important radio was to their daily experience ("*it's a vital part of my day. I'd be lost without it*").

The implicit ways that radio can support wellbeing is especially apparent in the identified sub-themes of radio's perceived social purpose: communion, social connection, company, and companionship. The potential for these outcomes is key to the contribution that this research makes in understanding how engaging with radio can support wellbeing. While previous work has highlighted the potential for radio to develop social connections between listeners, presenters, and the wider community (e.g., Meadows and Foxwell, 2011; Oliveira, 2013), the present findings also indicate that radio listening may go further to offset feelings of isolation. For example, listeners identified using the radio to create company and fill the silence while home alone, demonstrating that solitary listening may have benefits similar to that of social engagement. Further, the present study's sub-theme of companionship highlighted the importance of having consistent presenters and programs. Indeed, our listener-participants often referred to presenters as "friends"—a fact not lost on radio personnel, who aim to create relationships with their listeners, and to present in ways such that listeners "*become part of their show*". Radio's ability to create a sense of company and companionship aligns with the notion that radio can function as a social surrogate (Krause, 2020)—not only for isolated and disconnected individuals (Ewart, 2011) but for any and all listening. Thus, while researchers have remarked that audience participation is key to radio engagement (e.g., Moffat et al., 2023), radio's ability to create community has benefits beyond promoting engagement, including those related to wellbeing. As Guo (2015) and others have argued, the media landscape has become increasingly individualized. However, radio's perceived social purpose—in its ability to draw people together—underpins its continued relevance as a media technology.

While the present study extends previous work because it included both listener and personnel perspectives, it is not without limitations. For example, we focused on traditional radio broadcasting without a dedicated focus on podcasting. With increasing technologies to access and listen to radio, questions have been raised concerning what is classified as radio (Freire, 2007). Thus, additional research that considers on-demand catch up radio and podcasting will also be beneficial to understanding both how and why people engage with radio as well as the outcomes and benefits from doing so. Moreover, our participants did not overtly speak to commercial radio. Additional work might explicitly compare different radio formats.

While we focused on exploring how the radio might support wellbeing, no overt measures of perceived wellbeing were used. Future research is needed to explicitly measure wellbeing, and longitudinal studies would be well-placed to examine radio engagement and wellbeing patterns over time. Additionally, this convenience sample of older adults was highly engaged in radio listening such that it was interwoven into their daily lives. Therefore, the present study did not include the perspectives of older adults who rarely or never listen to radio. Designing an intervention or randomized control trial to further examine the impact of radio listening on perceived wellbeing is recommended.

Acknowledging these limitations, the present findings that radio explicitly and implicitly supports listener wellbeing have implications for broadcasting practices. For example, when looking at the role of radio during COVID-19, Hasnain et al. (2022) found that while radio stations sought to air accurate health messaging, presenters were also conscious to avoid "information fatigue". They recognized the importance of maintaining radio as a source of entertainment and not just news. Further, the continued regularity of broadcasting notably assisted people in a time of much uncertainty (Hasnain et al., 2022). These elements are congruent with findings in this present study that listeners identify radio as being multi-purposed and they develop attachments to certain presenters and programs. Indeed, our findings indicate that the regularity of programming must certainly assist with its ability to provide company and companionship, underpinning social surrogacy.

The findings also have implications concerning non-pharmacological and arts-based ways to support wellbeing and quality of later life. Alongside the increasing evidence for how music and the arts can facilitate wellbeing (e.g., Fraser et al., 2015; Cann, 2017), the present study adds evidence that radio also provides that support as a source of stimulation, entertainment, and social connection. The low-cost, accessibility of the radio as well as the variety of content means that it is well-placed to contribute to the promotion of wellbeing. Though not considered a "new" technology, people can use both analog and digital technology to listen to the radio. In fact, older adults report using all manner of digital devices including mobile telephones, tablets, and the television (Krause, 2020) to explore AM, FM, and internet radio possibilities. Comments from participants in our study confirmed the use of varied technology to listen to the radio ("*I have an app on my phone… there are so many different ways you can listen these days*").

This versatility in technological access may very well underpin why the radio is still so popular and relevant as a listening technology with older adults. Moreover, this versatility is important to note regarding how it can facilitate wellbeing, because previous research has shown that the provision of music listening technologies does not necessarily equate to their use (e.g., Krause and Davidson, 2021). As Krause and Davidson (2021) found, a barrier to older adults making use of listening technologies can be self-efficacy, or the belief in their capability or confidence in doing something. However, the radio is less likely possess this barrier as people can use technology they have and are familiar with (analog or digital) to tune in. This flexibility coupled with radio's ability to reach people in regional and remote locations, transmit information during crises, and provide feelings of community and connection beyond entertainment mean that the radio offers people a very important medium for supporting wellbeing.

Data availability statement

The datasets presented in this article are not readily available because the ethics approval for this project did not permit the sharing of collected data. Requests to access the datasets should be directed to AK, amanda.krause1@jcu.edu.au.

Ethics statement

The studies involving human participants were reviewed and approved by the James Cook University Human Ethics Research Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AK developed the study, obtained funding, and gained ethical approval for the research. AK and HF collected the data. HF conducted the data analysis with input from AK. All authors collaborated to draft and revise the manuscript. All authors approved the final version of the manuscript.

Funding

This research received funding from the Australian Association of Gerontology Research Trust's (AAG Research Trust) Hal Kendig Research Development Program in 2019. The funder was not involved in the study design, collection, analysis, interpretation of data, and writing of this article or the decision to submit it for publication.

Acknowledgments

The authors express sincere gratitude to all the participants who have participated in this research.

Conflict of interest

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

Publisher's note

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

References

Albarran, A. B., Anderson, T., Bejar, L. G., Bussart, A. L., Daggett, E., Gibson, S., et al. (2007). "What happened to our audience?" Radio and new technology uses and gratifications among young adult users. *J. Radio Stud.* 14, 92–101. doi: 10.1080/10955040701583171

Ames, K. (2016). Talk vs chat-based radio: a case for distinction. *Radio J. Int. Stud. Broadcast Audio Media* 14, 177–191. doi: 10.1386/rjao.14.2.177_1

Australian Communications and Media Authority (2020). *Trends in Viewing and Listening Behaviour: ACMA Consumer Survey 2020*. ACMA. Available online at: https://www.acma.gov.au/sites/default/files/2020-11/Trends-in-viewing-and-listening-behaviour_ACMA-consumer-survey-2020.pdf (accessed June 6, 2023).

Bednarek, M. (2014). Involvement in Australian talkback radio—a corpus linguistic investigation. *Aust. J. Ling.* 34, 4–23. doi: 10.1080/07268602.2014.875453

Bhattacharya, K. (2017). *Fundamentals of Qualitative Research: A Practical Guide*. Routledge.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Braun, V., and Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners*. London: Sage.

Braun, V., and Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qual. Res. Sport Exercise Health* 11, 589–597. doi: 10.1080/2159676X.2019.1628806

Cann, P. L. (2017). Arts and cultural activity: a vital part of the health and care system. *Australas. J. Ageing* 36, 89–95. doi: 10.1111/ajag.12438

Community Broadcasting Association of Australia (2019). *Community Radio National Listener Survey 2019: Summary Report of Findings Community Broadcasting Association of Australia*. Available online at: https://www.cbaa.org.au/broadcasters/get-data-national-listener-survey-station-census (accessed June 6, 2023).

Community Broadcasting Association of Australia (2022). *Community Radio Listener Survey*. Available online at: https://www.cbaa.org.au/broadcasters/get-data-national-listener-survey-station-census/national-listener-survey-fact-sheets (accessed June 6, 2023).

Ewart, J. (2011). Therapist, companion, and friend: the under-appreciated role of talkback radio in Australia. *J. Radio Audio Media* 18, 231–245. doi: 10.1080/19376529.2011.615775

Ewart, J., and Ames, K. (2016). Talking text: exploring SMS and e-mail use by Australian talkback radio listeners. *Radio J. Int. Stud. Broadcast Audio Media* 14, 91–107. doi: 10.1386/rjao.14.1.91_1

Ferguson, D. A., Greer, C. F., and Reardon, M. E. (2007). Uses and gratifications of MP3 players by college students: are iPods more popular than radio? *J. Radio Stud.* 14, 102–121. doi: 10.1080/10955040701583197

Forde, S., Foxwell, K., and Meadows, M. (2009). *Developing Dialogues: Indigenous and Ethnic Community Media Broadcasting in Australia*. Intellect. Available online at: https://www.intellectbooks.com/developing-dialogues (accessed June 6, 2023).

Foxwell, K. (2012). Community radio in an Australian city: the Melbourne experience. *Radio J. Int. Stud. Broadcast Audio Media* 10, 161–172. doi: 10.1386/rjao.10.2.161_1

Fraser, K. D., O'Rourke, H. M., Wiens, H., Lai, J., Howell, C., and Brett-MacLean, P. (2015). A scoping review of research on the arts, aging, and quality of life. *Gerontologist* 55, 719–729. doi: 10.1093/geront/gnv027

Freire, A. M. (2007). Remediating radio: audio streaming, music recommendation and the discourse of radioness. *Radio J. Int. Stud. Broadcast Audio Media* 5, 97–112. doi: 10.1386/rajo.5.2-3.97_1

Guo, L. (2015). Exploring the link between community radio and the community: a study of audience participation in alternative media practices. *Commun. Cult. Critique* 10, 112–130. doi: 10.1111/cccr.12141

Hasnain, A., Krause, A. E., Hajek, J., Lloyd-Smith, A., and Lori, L. (2022). Broadcasting during COVID-19: community language radio and listener wellbeing. *Radio J. Int. Stud. Broadcast Audio Media* 20, 227–247. doi: 10.1386/rjao_00061_1

Hemming, L., Pratt, D., Bhatti, P., Shaw, J., and Haddock, G. (2021). Involving an individual with lived-experience in a co-analysis of qualitative data. *Health Expect.* 24, 766–775. doi: 10.1111/hex.13188

Hugelius, K., Adams, M., and Romo-Murphy, E. (2019). The power of radio to promote health and resilience in natural disasters: a review. *Int. J. Environ. Res. Public Health* 16, : 2526. doi: 10.3390/ijerph16142526

Keough, S. B. (2010). The importance of place in community radio broadcasting: a case study of WDVX, Knoxville, Tennessee. *J. Cult. Geogr.* 27, 77–98. doi: 10.1080/08873631003593265

Krause, A. E. (2020). The role and impact of radio listening practices in older adults' everyday lives. *Front. Psychol.* 11, 603446. doi: 10.3389/fpsyg.2020.603446

Krause, A. E., and Davidson, J. W. (2021). A qualitative exploration of aged-care residents' everyday music listening practices and how these may support psychosocial well-being. *Front. Psychol.* 12, 585557. doi: 10.3389/fpsyg.2021.585557

Krause, A. E., Davidson, J. W., and North, A. C. (2018). Musical activity and well-being: A new quantitative measurement instrument. *Music Percept.* 35, 454–474. doi: 10.1525/MP/2018.35.4.454

Krause, A. E., Lloyd-Smith, A., and Hajek, J. (2020). The role of community language radio for understanding creativity and wellbeing in migrant communities in Australia. *Int. J. Wellbeing* 10, 83–99. doi: 10.5502/ijw.v10i5.1495

Krause, A. E., and North, A. C. (2016). Music listening in everyday life: devices, selection methods, and digital technology. *Psychol. Music* 44, 129–147. doi: 10.1177/0305735614559065

Laskar, K. A., and Bhattacharyya, B. (2021). Community radio stations' production responses to COVID-19 pandemic in India. *Media Asia* 48, 243–257. doi: 10.1080/01296612.2021.1970421

Maina, S. N. (2013). Communication strategies employed by Kenya Community Radio in fostering social cohesion and integration in Kenya. *Int. J. Soc. Sci. Entrepreneurship* 1, 688–713. Available online at: https://www.ijsse.org/articles/ijsse_v1_i5_688_713.pdf

McClung, S., Pompper, D., and Kinnally, W. (2007). The functions of radio for teens: where radio fits among youth media choices. *Atl. J. Commun.* 15, 103–119. doi: 10.1080/15456870701215842

Meadows, M. (2013). Putting the citizen back into journalism. *Journalism* 14, 43–60. doi: 10.1177/1464884912442293

Meadows, M., and Foxwell, K. (2011). Community broadcasting and mental health: the role of local radio and television in enhancing emotional and social well-being. *Radio J. Int. Stud. Broadcast Audio Media* 9, 89–106. doi: 10.1386/rjao.9.2.89_1

Milan, S. (2008). What makes you happy? Insights into feelings and muses of community radio practitioners. *Westminster Papers Commun. Culture* 1, 25–43. doi: 10.16997/wpcc.49

Moffat, B., Osunkunle, O. O., Mabika, M., and Moffat, F. (2023). "The importance of audience participation in the success of community radio," in *African Language Media*, eds P. Mpofu, I. Fadipe, and T. Tshabangu (Abingdon: Taylor & Francis), 179–192.

North, L., and Dearman, P. (2010). The rhetoric of "community": ABC local radio's coverage of the 2009 Victorian bushfires. *Media Int. Aust.* 137, 80–89. doi: 10.1177/1329878X1013700110

Oliveira, M. (2013). "Sounds and identity: the role of radio in community building," in *Radio: Community, Challenges, Aesthetics*, ed G. Stachyra (Lublin: Maria Curie-Sklodowska University Press), 177–188.

Order, S., and O'Mahony, L. (2017). Building a purposeful identity in the older adult volunteer space: a case study of community radio station 6RPH (Radio Print-Handicapped). *Commun. Res. Pract.* 3, 31–44. doi: 10.1080/22041451.2017.1271971

Rodero, E. (2020). Radio: the medium that best copes in crises. Listening habits, consumption, and perception of radio listeners during the lockdown by the Covid-19. *El profesional de la información* 29, e290306. doi: 10.3145/epi.2020.may.06

Schäfer, K., and Eerola, T. (2020). How listening to music and engagement with other media provide a sense of belonging: an exploratory study of social surrogacy. *Psychol. Music* 48, 232–251. doi: 10.1177/0305735618795036

Schäfer, K., Saarikallio, S., and Eerola, T. (2020). Music may reduce loneliness and act as social surrogate for a friend: evidence from an experimental listening study. *Music Sci.* 3, 1–16. doi: 10.1177/20592043209 35709

Shepherd, E. (2022). *Commercial Radio Audience up 1.3% to 11.2 Million in the Five Metro Capitals for 2021*. Available online at: https://mumbrella.com.au/commercial-radio-audience-up-1-3-to-11-2-million-in-the-five-metro-capitals-for-2021-722661 (accessed June 6, 2023).

Siemering, W. (2000). Radio, democracy and development: evolving models of community radio. *J. Radio Stud.* 7, 373–378. doi: 10.1207/s15506843jrs0702_10

Smith, M. L., Menn, M., and McKyer, E. L. J. (2011). Effectiveness of the radio as a health information source. *J. Radio Audio Media* 18, 196–211. doi: 10.1080/19376529.2011.615776

Sparkes, A.C., and Smith, B. (2014). *Qualitative Research Methods in Sport, Exercise and Health: From Process to Product*. Abingdon, VA: Routledge.

Vidal, M. M. (2019). Radio as a sociability space. *J. Radio Audio Media* 26, 75–78. doi: 10.1080/19376529.2019.1564995

Vuuren, K. V. (2002). Beyond the studio: a case study of community radio and social capital. *Media Int. Aust.* 103, 94–108. doi: 10.1177/1329878X0210300113

Watson, I. (2013). Connecting, informing and empowering our communities: remote indigenous radio in the Northern peninsula area. *Media Int. Aust.* 148, 59–69. doi: 10.1177/1329878X1314800107

Watson, I. (2016). No more 'voices from down south': parallel voices in remote Queensland indigenous communities. *Radio J. Int. Stud. Broadcast Audio Media* 14, 137–151. doi: 10.1386/rjao.14.1.137_1

TYPE Original Research
PUBLISHED 19 July 2023
DOI 10.3389/fcomp.2023.1165442

OPEN ACCESS

EDITED BY
Roberto Therón,
University of Salamanca, Spain

REVIEWED BY
Balandino Di Donato,
Edinburgh Napier University, United Kingdom
James Dooley,
The Open University, United Kingdom

*CORRESPONDENCE
Hans Lindetorp
✉ hans.lindetorp@kmh.se

RECEIVED 14 February 2023
ACCEPTED 03 July 2023
PUBLISHED 19 July 2023

CITATION
Lindetorp H, Svahn M, Hölling J, Falkenberg K and Frid E (2023) Collaborative music-making: special educational needs school assistants as facilitators in performances with accessible digital musical instruments.
Front. Comput. Sci. 5:1165442.
doi: 10.3389/fcomp.2023.1165442

Collaborative music-making: special educational needs school assistants as facilitators in performances with accessible digital musical instruments

Hans Lindetorp[1,2]*, Maria Svahn[2], Josefine Hölling[2], Kjetil Falkenberg[2] and Emma Frid[2,3]

[1]Department of Music and Media Production, KMH Royal College of Music, Stockholm, Sweden, [2]Media Technology and Interaction Design, KTH Royal Institute of Technology, Stockholm, Sweden, [3]Science and Technology of Music and Sound (STMS), Institute for Research and Coordination in Acoustics/Music (IRCAM), Paris, France

The field of research dedicated to Accessible Digital Musical Instruments (ADMIs) is growing and there is an increased interest in promoting diversity and inclusion in music-making. We have designed a novel system built into previously tested ADMIs that aims at involving assistants, students with Profound and Multiple Learning Disabilities (PMLD), and a professional musician in playing music together. In this study the system is evaluated in a workshop setting using quantitative as well as qualitative methods. One of the main findings was that the sounds from the ADMIs added to the musical context without making errors that impacted the music negatively even when the assistants mentioned experiencing a split between attending to different tasks, and a feeling of insecurity toward their musical contribution. We discuss the results in terms of how we perceive them as drivers or barriers toward reaching our overarching goal of organizing a joint concert that brings together students from the SEN school with students from a music school with a specific focus on traditional orchestral instruments. Our study highlights how a system of networked and synchronized ADMIs could be conceptualized to include assistants more actively in collaborative music-making, as well as design considerations that support them as facilitators.

KEYWORDS

collaborative music-making, Web Audio, assistants as facilitators, accessible digital musical instruments, interactive musical systems

1. Introduction

The field of research dedicated to Accessible Digital Musical Instruments (ADMIs), i.e., *accessible musical control interfaces used in electronic music, inclusive music practice, and music therapy settings* (see Frid, 2019a), is growing. With an increased interest in exploring how different accessible music technologies can be used to promote diversity and inclusion in music-making and musicking (Small, 1998) comes a need to understand, design, and possibly also evaluate such systems. Although attempts have been made to introduce design principles and classification methods based on different use cases (Frid, 2019b; Davanzo and Avanzini, 2020; Harrison, 2020) and to investigate the potential of existing ecological frameworks in design and evaluation of ADMIs (e.g. the Human Activity Assistive Technology and the Matching Person and Technology frameworks deployed by Lucas et al. 2021), there is still no commonly accepted or established evaluation methodology for ADMIs. The topic of evaluation has been widely debated in the fields of New Interfaces for Musical Expression

(NIME) and Digital Musical Instruments (DMI) over the years (see e.g. Greenberg and Buxton, 2008; Johnston, 2011; Barbosa et al., 2015). Lately, researchers have emphasized the importance of considering the sociocultural context of the musical interaction when trying to understand what makes "a good musical instrument" (see e.g. Jack et al., 2020; Rodger et al., 2020). In this paper, we describe the context-specific design and evaluation of a master controller instrument developed for assistants at a Special Educational Needs (SEN) School. More specifically, we explore how such an interface should be designed to support the assistants' roles as facilitators in a musical performance with a group of students with Profound and Multiple Learning Disabilities (PMLD)[1,2] who are playing ADMIs together with a professional musician.

The field of Digital Musical Instruments (DMIs) has historically largely focused on borrowing tools and frameworks from Human-Computer Interaction (HCI) (see notable work on evaluation by Wanderley and Orio, 2002, for example). Such approaches often build on the idea that DMIs can be described as "devices" with certain properties that can be evaluated from a "usability" and "accessibility" perspective. Employing HCI concepts directly to musical instruments is, however, something that should be done with care, since this might lead to an oversimplification of the complexity of musical interactions. While the framework proposed by Wanderley and Orio (2002) indeed is useful in certain contexts, it also has drawbacks. The reduction of musical interaction to simple tasks may compromise the authenticity of the interaction; since musical interactions involve creative and affective aspects, they cannot simply be described as tasks with completion rates (Stowell et al., 2009). Frameworks that emphasize the different perspectives in instrument design, such as that of the listener, the performer, vs. the instrument constructor, among others (Kvifte and Jensenius, 2006, see also O'Modhrain, 2011) have been proposed to address (some of) these issues. The need to move away from task-based views of musical performance (see for example El-Shimy and Cooperstock, 2016) and instead take a more holistic view of music activities has been stressed. Rodger et al. (2020) argue that the functional properties of an instrument can only be meaningfully understood relative to the capabilities of a specific musician at a specific period in their musical development. In other words, it is difficult to fully comprehend what a musician does with an instrument if disconnected from the immediate and extended sociocultural context. Building on this notion, they propose to view instruments as "processes" rather than "devices", and musicians as "agents" in "musical ecologies", rather than "users".[3] Jack et al. (2020) propose, based on the idea of "performance eco-systems" discussed by Waters (2007), to view DMIs as situated and ecologically valid artifacts which should be evaluated using qualitative and reflective processes focusing on sociocultural phenomena rather than first-wave HCI techniques. Similarly, Lucas et al. (2021) explored ecological perspectives of human activity in the use of DMIs and assistive technology. To conclude, numerous researchers have stressed the importance of shifting focus from the technical aspects of DMIs to the sociocultural contexts in which musical interactions are taking place.

In the current paper we describe an exploratory study focused on a specific scenario in which a group of students with Profound and Multiple Learning Disabilities (PMLD) at a Special Educational Needs (SEN) school in Sweden[4] interacted with a set of ADMIs – the *Funki instruments* – in collaborative music-making sessions together with their assistants and a professional musician (an accordionist improvising over a four-chord progression). The Funki instruments are tangible ADMIs that allow for different modes of interaction to create sound by physically engaging with sliders, buttons, and touch pads, embedded in wooden boxes (Figure 1). In this paper, we use the term "assistant" to include any personnel attending these sessions, such as teachers, teaching assistants, or carers. Our work is part of a longer study with the overall goal of organizing a joint concert that brings together students from the SEN school with students without disability enrolled in a music school focused on traditional orchestral instruments. The music-making aspects that we focus on in this study are primarily those of a goal-oriented composition with a relatively well-defined structure. Our research is largely inspired by the "Able Orchestra",[5] a project based on the principle of enabling people to create and perform music on equal terms, regardless of their physical dexterity or musical experience. The Able Orchestra allows musicians using technology to join acoustic instrumentalists and orchestral players to create new work in an ensemble (Orchestras Live, 2023).

To reach the above-discussed goal, multiple challenges need to be addressed. Besides aspects such as instrument design and the sociocultural setting of a joint concert, one challenge is that the music to be performed, i.e., the composition, will have musical boundaries in terms of tonality and chord changes. The students from the music school, which is focused on traditional orchestral instruments, read scores and are trained to follow a shared tempo. However, neither the students nor the assistants at the SEN school have this type of musical training. The students at the SEN school also have multifunctional physical and intellectual challenges which make it difficult for them to play notes at pre-defined fixed timings. Overall, it should be noted that the musical activities held at the two schools differ a lot in their focus; music sessions at the school dedicated to traditional orchestral instruments generally put emphasis on performance aspects, while music sessions at the SEN school focus more on aspects such as active participation, personal development, and social aspects of music-making. To explore how we could merge these two musical practices into one musical performance context, we designed a dedicated system. The design of the system was based on allowing an assistant to

1 In this paper, we use Person/People First Language (PFL) when writing about disability, as opposed to Identity First Language (IFL) (see Ferrigon and Tucker, 2019).

2 PMLD is commonly used to describe a person with severe learning disabilities who most likely has other complex disabilities and health conditions (Bellamy et al., 2010), although there is no single universally agreed definition of the term, and research has highlighted that no definition can fully articulate the complexities associated with it (Bellamy et al., 2010).

3 An ecology in this context corresponds to a system compromising an agent and environment.

4 Dibber Rullen, see https://dibber.se/skola/rullens-sarskola/om-oss/.

5 See for example https://youtu.be/pf8k3uA3dxM.

FIGURE 1
The three Funki instruments used in the study, with different controllers. From the left: two touch-sensitive pads, two momentary buttons, and two sliders. The black area is the backside of the loudspeaker.

control certain aspects of the music, i.e., the chord changes, and to communicate with a professional musician who improvised over a pre-defined chord progression. We organized music-making sessions at the SEN school to explore how a master controller instrument developed specifically for the assistants should be designed to support their role as facilitators in a performance context. The aim of the music-making sessions was to explore the interplay between assistants and the professional musician that may arise in this context, and what is required for such a session to be successful.

Although student assistants have existed as a support function for students in the Swedish school system since the 1960s, the profession has not received much attention in the Scandinavian educational literature, and research on this role is scarce, with very few official statistics describing the group (Östlund, 2017). Overall, literature exploring the advantages vs. disadvantages of including assistants in SEN settings appears to be somewhat inconclusive, with work pointing toward both potential benefits and risks. For example, a study on how help is provided for pupils with physical disabilities and how school assistants influence their participation published by Hemmingsson et al. (2003) suggested that the assistants could both facilitate and hinder student participation. More specifically, the study highlighted the dilemma of closeness and distance; an overly distant approach to the pupil can result in an experience of alienation, but a too-close relationship may result in the student not participating in the challenges faced by other pupils, leading to less independence. Recommendations on how to enhance the efficacy and practices of teaching assistants are discussed in the review of studies on inclusive classrooms presented by Sharma and Salend (2016).

The ADMIs used in this study, the Funki Instruments, have previously been used in experiments in the same SEN school (without a professional musician) for a total of eight half-hour music-making sessions (see Svahn et al., 2021). Findings from this work suggested that the instruments provided a foundation that allowed the students to play together in a band; different instruments allowed students with different abilities and needs to take part, and each student could find some instrument that they liked. However, the results also indicated that musical interactions and group dynamics were highly dependent on the participation of session assistants. In other words, it is important to consider the role of assistants in the music-making taking place at the school, and further research is needed to understand how the assistants' roles as facilitators can be supported by providing tools that may enable more active participation. It should be noted that the assistants at the specific SEN school described in this paper have little or no music training. They also have limited experience with music technology overall. The assistants often have a lot of responsibilities in everyday situations, and therefore they usually have little headroom to learn or take on new tasks. In the previous study, Svahn et al. (2021) concluded that it is important that the assistants understand both the student's needs as well as their interaction with the musical instrument, in order to be able to successfully engage the students in musical activities. The added master instrument described in the current paper should not only be easy to play and allow the assistant to support the students in their music-making; it should also complement the other instruments with sounds that contribute to the musical context. By designing a master instrument according to these constraints, we aim to support the assistants to become more actively involved in the music-making process, thereby also facilitating the students' performances on their instruments.

The current work expands on the previous study by Svahn et al. (2021) by embedding WebAudioXML (Lindetorp and Falkenberg, 2022) into the previous versions of the Funki instruments. This allows for expanded communication between the instruments to let interactions from multiple users control sounds according to mapping rules. WebAudioXML can through these mappings be used to connect the Funki instruments to a master instrument controlling musical aspects such as tonality, rhythmic density, and structure of the composition. In the configuration used for this particular study, we focused especially on tonality. The master instrument intended to be used by the assistants could both trigger sounds in itself and control the pitches of the students' instruments. The metaphor that inspired us during this design process was that of a child strumming guitar strings, with a parent simultaneously changing the chords by moving the fingers on the fretboard. Building on this metaphor, the master instrument acted as a harmony controller for the three student instruments; they all changed pitches according to the rules set by the assistant.

With the above-described setup, the students played along with a professional musician in a performance repeated three times, each with a new assistant. This setting allowed us to explore aspects of collaboration in the music-making taking place between the participants. The sessions were video recorded and the assistants' controller events were tracked and analyzed. This was followed by semi-structured interviews with the assistants and the musician, focusing on how the setup supported inclusive music-making in the group and the assistant's role in this context. Our findings highlight how ADMI systems could be conceptualized and designed to include assistants more actively in collaborative music-making involving a professional musician, and design considerations that may support the assistants' facilitator role in such contexts.

2. Related work

As previously mentioned, this paper builds on work by Svahn et al. (2021), in which a set of four collaborative AMDIs for students with Profound and Multiple Learning Disabilities (PMLD) were developed and tested with a group of students at a Swedish SEN school. Students with PMLD may express themselves using several different communication methods, depending on what is most efficient for them at the time. Pre-verbal students who do not have verbal communication skills yet can use bodily gestures, nonverbal sounds, pointing, and facial expressions to express themselves. To support communication with students with PMLD, alternative methods for augmented communication, such as PODD (Pragmatic Organization Dynamic Display) (Light, 1988; Porter and Cafiero, 2009) can be used. In a related previous study, we invited a group of students from the same SEN school to take part in a 1.5-year-long project focused on musical interfaces and musical haptics (Frid et al., 2022). In this work, we assessed how the students could be involved in the customization and evaluation of the design of a multisensory music experience intended for a large-scale ADMI using a *Participatory Design with Proxies (PDwP)* methodology. The proxies in this context were teaching assistants and a teacher working with the students, as well as parents. The PDwP method enabled the inclusion of input from different stakeholders providing valuable insights and feedback to augment direct input from the students. Findings highlighted accessibility limitations of the musical interface as well as the importance of using a multifaceted variety of methods to arrive at more informed conclusions when applying a PDwP methodology with pre-verbal students. In the current study, we explore the role of the teaching assistants further, using similar techniques as the ones employed in the previous study, e.g., *stimulated recall* of videos collected during music-making sessions with the students (see Calderhead, 1981).

To the authors knowledge, little previous work has focused specifically on design considerations for ADMIs that support assistants when musically engaging in ensemble play together with students with PMLD. Despite a long research tradition focused on collaborative music-making using DMIs and NIMEs (see e.g. "Multi-User Instruments" by Jordà, 2005; "Orchestras of Digital Musical Instruments" by Berthaut and Dahl, 2015; and the dimension space for evaluating collaborative musical performance systems by Hattwick and Wanderley, 2012) findings from the systematic review of ADMIs published by Frid (2019a) revealed that few of the surveyed instruments were designed for ensemble settings. Relevant work in this context includes, for example, an exploration of the context and design of collaborative musical experiences for novices published by Blaine and Fels (2003). A discussion of accessibility challenges that people with disabilities face when making music in groups was presented in work by Steinmeier et al. (2022), together with design suggestions. 18 design considerations for DMIs used in SEN settings were proposed by Ward et al. (2017) and a set of design considerations based on interviews with SEN teachers in Germany was recently published by Förster (2022). Access barriers in SEN settings have also been discussed (Davis et al., 2019). Farrimond et al. (2011) mention three access barriers for music technology in special educational and disabled music settings in the UK: a need for specialist training (this was also mentioned by Welch et al., 2016 and Davis et al., 2019), resources, and a fear and dislike or indifference to technology. For example, Davis described issues when asking non-musically trained staff to come up with and deliver music-based activities.

A study on facilitator involvement in a music technology project where participants with complex disabilities[6] use technology to assist them in music performance was described in Dickens et al. (2018). In this paper, the term "facilitated performance" was used to describe the practice of musical performance involving performers with complex disabilities who are supported by musical experts and other facilitators. Findings suggested that including facilitators in the design of DMIs could allow for improved accessibility for users with complex disabilities, highlighting the importance of the facilitator role in adapting and supporting the use of technology in such contexts. The work also revealed that the social relationships between performer and facilitator were paramount for the success of the project, and that participatory design is a strong design methodology for facilitated performance. The authors describe that facilitation can take many forms throughout the process of supporting interactions with DMIs. This includes *prompting*, in which participants are encouraged to play and structure performances, for example using tools to inform about the current section of the piece or queuing gestures like pointing at performers when it is their time to play; *demonstration*, in which demonstrative behavior (including mimetic gestures) is used repeatedly to reinforce understanding and provide reminders; and *assisting*, for example involving holding equipment in place and making sure that the equipment is positioned correctly. Interfaces developed for the facilitator should focus on providing a simplified baseline level that can be

6 Here defined as "conditions that affect both cognitive and motor abilities of an individual".

built on, to expand the interface for those with greater technical knowledge. Dickens et al. (2018) suggested that facilitators can be considered gatekeepers to musical activities for performers with complex disabilities since they possess a multitude of knowledge around music performance and technologies and thus are most equipped to communicate this knowledge to the performer. In the current work, we explore the assistants' roles as gatekeepers in musical activities for students with PMLD. However, the assistants described in our work are somewhat different from the facilitators in Dickens et al. (2018), in the sense that the assistants have little to no prior background in musical performance.

3. Materials and methods

The current study focuses on the design and evaluation of a system developed to support the assistants roles as facilitators in music-making sessions with students with PMLD (playing ADMIs) and a professional musician (playing the accordion). In this section, we describe the technical setup, how the instruments communicate with each other, the musical preconditions, and how the user interactions affect the sound.

3.1. Hardware

The setup consisted of four ADMIs: three existing Funki instruments and one new master instrument to be used by an assistant. The prototyping, development, and affordances of the earlier versions of the student instruments are further described in Svahn et al. (2021) and have in accordance with a continued participatory design approach been updated for the purpose of this study, based on findings from long-term use in the SEN school's daily practice. The student instruments are made of wood (30 × 20 × 20 cm) and can be placed on a table, on a wheelchair, or on your lap. Each instrument has an integrated loudspeaker, a microcomputer, and two interactive controllers of the same type (10 cm in size), which can be used to trigger sounds. The controllers are of three different types: trim potentiometer sliders, discrete momentary push-buttons, and touch-sensitive pads made using conductive paint. The wooden boxes are painted in white color, and the controllers are made with different distinct colors (Figure 1). The master instrument differs from the student instruments in several ways. Most importantly, it is not embedded into a wooden box, but designed as a Graphical User Interface (GUI) that runs on a computer. As such, it can be played using the space-bar on the keyboard to trigger a sound, and to control the selection of sounds for the other instruments.

3.2. Software

All instruments are connected to the master instrument, which has several important functions: it stores the samples and audio configurations for the different instruments, it receives the control data from all instruments, it determines how the assistant can control the composition, and finally, it sends out the appropriate sound to each integrated loudspeaker (Figure 2). The master instrument computer runs a server built with Node.js, Socket.IO and Express.js, which gets the control data from all connected instruments through Open Sound Control (OSC)[7]. The instruments send OSC messages continuously to update the system about their current state and the master instrument responds by generating and returning the sounds to the loudspeakers in the corresponding instruments. The logic that controls how the sounds shall respond to the interactions is built using web technologies and is configured using WebAudioXML.

3.3. Composition

The sounds generated by the instruments are sampled elements of a musical composition. The first author (HL) produced a collection of 42 audio files to respond to various combinations of interactions with the instruments. The overall goal of the sound design was to give the students a natural response to their actions and also to limit the way the sounds could be combined, to ensure that the result would always stay within a defined tonality. This was achieved by creating a collection of sounds for each instrument that could be combined in predefined ways. The master instrument triggers a strummed harp cord, cycling through a chord progression in E major (E, C#m7, F#m7, B11). At the same time, the master instrument will also set the sound of the other instruments to match the chord that is currently playing. The instruments trigger sounds that are composed and arranged to complement each other in terms of pitch and timbre and are selected according to the actions on the instrument, the current chord, and a random factor that adds variation to the performance. A detailed description of the composition represented as a score is displayed in Figure 3.

3.4. Mapping strategies

The instruments use different gestures to trigger sounds. The final mapping was informed by previous experiments with the Funki instruments conducted with the same user group (Svahn et al., 2021). To make the interactions intuitive, we tried to match the sounds with the actions as closely as possible to acoustic instruments. For example, the instrument with two sliders uses the metaphor of a bow and a string to make the instrument sound only when the user is actively moving the slider. It maps the speed of the slider movement to control the volume of the sound (no movement results in silence). For this instrument, we used a collection of six samples (three for each slider). The second and the fourth chord used the same samples. One slider plays notes in a high octave and the other plays notes one octave higher. The instrument with two buttons triggers samples with a decay similar to the response of depressing a piano key or plucking a guitar string. For this instrument, we use a collection of 20 acoustic guitar samples; each button iterates through a sample set with five different pitches, one set for the E and C#m7 chords, and one set for the F#m7

7 See: Node.js (https://nodejs.org/), Socket.IO (https://socket.io), Express.js (https://expressjs.com/) and Open Sound Control (https://opensoundcontrol.stanford.edu/).

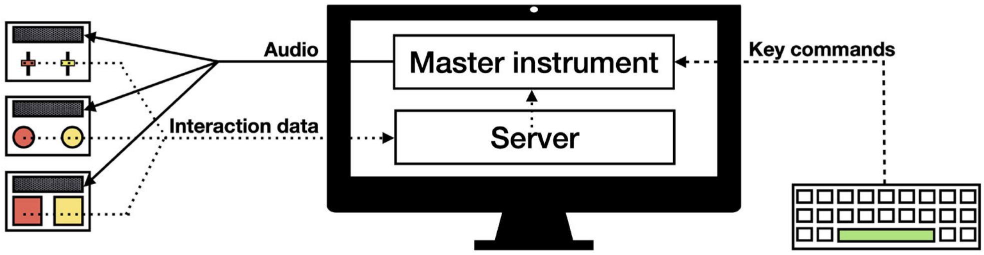

FIGURE 2
The system with three instruments sending control data (OSC) to a Node.js-server. The data is forwarded to the master instrument (in a web client). The instrument data is combined with the master instrument interactions to determine the audio output sent to the loudspeakers of respective instruments.

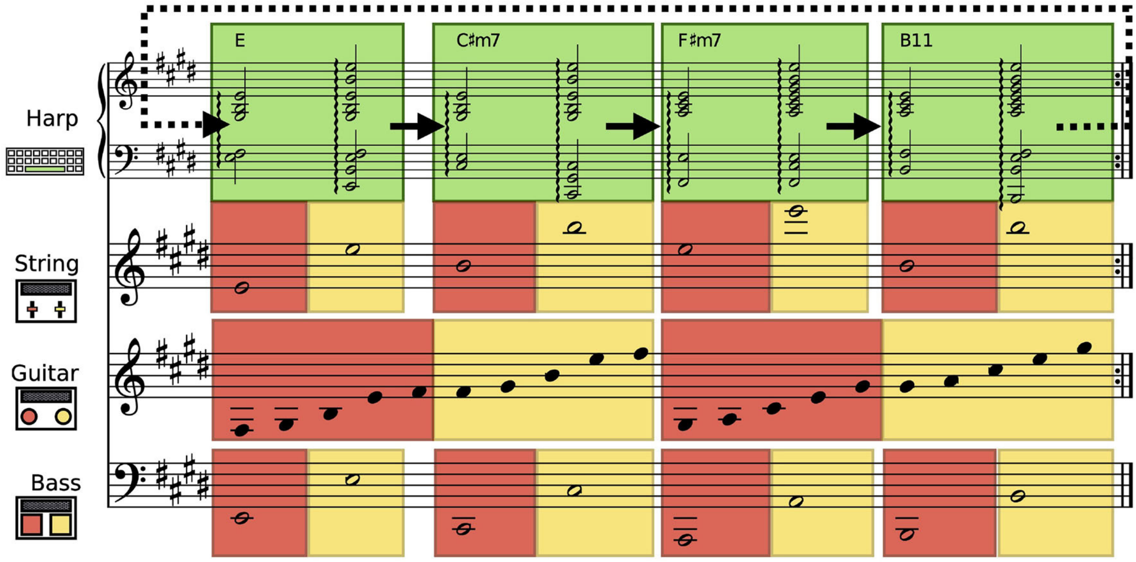

FIGURE 3
All components of the composition represented as a score. Each green box represents a chord controlled by the master instrument, which plays one of two variations for each chord. The red and yellow boxes indicate the different notes for the two controllers on each instrument. The guitar has the same sequence of notes for the first two and the second two chords, respectively.

and B11 chords. One button triggers notes in a medium octave and the other button triggers notes one octave higher. Finally, the instrument with two touch-sensitive pads turns on and off the volume of its sound in an organ-like manner. It is built using eight loops playing samples of synth bass notes. The pads have one set of samples each, with one sample for each chord. One pad triggers pitches in a low octave, and the other pad triggers pitches one octave below. Several testing sessions were devoted to fine-tuning the interactions/gestures to the sounds and deciding how the instruments should respond to the control data from the master instrument.

3.5. Workshop

The evaluation of the system was performed in the form of a workshop with three students (2F, ages=10 & 16 yrs and 1M, age=11 yrs) with Profound and Multiple Learning Disabilities. The students who participated are pre-verbal, have multifunctional physical challenges, varying motor skills, moderate to severe intellectual challenges, and use wheelchairs. We repeated the music-making sessions three times, with the same students, but with three different assistants (3F, ages=25, 47, 52 yrs). The students kept the same instrument in all three sessions. One of the authors (MS) was assisting two of the students while the assistant played the master instrument. Each assistant participated once, i.e., in one session each. Apart from managing the master instrument, the assistants also assisted one of the students. A professional musician (M, age=37 yrs) participated in all three sessions, playing the accordion together with the students and the assistant. The following sections describe these sessions, as well as the data collected, in detail. The first author (HL) was present during all sessions, to observe the interplay in the room, and to provide help with the technical setup if required.

3.5.1. Procedure

Each assistant spent 15 min with the group (three students, one musician, as well as the first and second author). The assistants first got a short introduction to the system by the first author. This introduction focused on the functionality of the master instrument. The assistant and the musician were instructed to communicate, with or without words, about when the assistant should press the key on the master instrument in order to move the composition forward, see Sections 4.1.1, 4.1.2, 4.1.3. The musician then indicated that the music would start. This was followed by an improvised session shaped by the creative possibilities and limitations of the composition outlined in Section 3.3. When the assistant pressed the key on the master instrument, the music advanced to the next chord. This also triggered a sample with the sound of a harp arpeggio and updated the selection of possible samples for the other instruments to harmonize with the currently selected chord. The musician was asked to identify strategies for communicating his intentions to the assistant. He was free to test different approaches during the three different sessions. Video and audio were captured during the three sessions.

3.5.2. Logging of data

The videos were manually tagged by the first author, who is a professional musician and lecturer in music, to identify master instrument onsets. The events were tracked using a MIDI keyboard and a Digital Audio Workstation with video support (we used Logic Pro[8]). The events were exported as a standard MIDI file and converted to JSON data using Tone.js.[9]

3.5.3. Interviews

In a second meeting organized after the workshop, two of the authors (HL, KF) performed a semi-structured interview with the assistants and the musician. These interviews were based on a stimulated recall methodology (Calderhead, 1981), meaning that the assistants were shown replays of video clips from the sessions to stimulate a commentary upon the thought process at that time. The clips were selected by the first author with a focus on when the music started or stopped, if the tempo changed, or if anything special happened that affected the communication between the participants or the music. Each interview session included 2–3 video clips showing instances of communication for a total duration of a couple of minutes. The assistants only watched clips from their own sessions. They were then asked about what happened in the video at that specific time. In addition, the assistants were also asked questions about (1) their musical background, (2) how they communicated with the musician, and (3) how they understood when to press the keyboard on the master instrument. We also asked all participants about suggested improvements. This discussion reflected ideas both for the instruments, the composition, and the way the sessions were structured and organized. All interviews were held in Swedish and the average duration was 18 min. They were recorded and automatically transcribed using the built-in transcription feature in Microsoft Word.

3.5.4. Analysis

The quantitative data (recorded MIDI onsets) was analyzed to calculate chord durations and the deviation of the time between onsets. Deviation here refers to the difference between subsequent chord duration values; this is a measure of how much the assistant deviated from a steady tempo for the chord changes. The MIDI onsets were visualized with positions and durations using the "arrange view" of the Digital Audio Workstation and the graphs showing the deviation for each onset were produced to align the values with the corresponding MIDI onset. Statistical analysis was performed to identify significant differences between assistants.

A thematic analysis of the material collected during the interview sessions was performed, following the procedure outlined by Braun and Clarke (2006). The transcription was saved in a spreadsheet format for easier processing. First, one of the authors (KF) watched the recordings and read the transcriptions, generating a set of initial codes for tagging the responses and collating these into themes. Then, two other authors (HL, EF) systematically tagged the material, adding new codes when relevant, collated the codes into themes, and then reviewed and revised the themes. In the final step, all three authors involved in the tagging searched the material for recurrent themes, discussed the material, and revised and refined the themes to a set that was coherent with the study's objective.

3.5.5. Ethics statement

The research procedure described in this paper was reviewed by the Swedish Ethical Review Authority (application No. 2021-06307-01). The study was carried out in accordance with the declaration of Helsinki. We followed informed consent rules and guidelines for ethical research practices.[10] All parents gave written informed consent before participation and agreed to the data being collected as described in the consent form. It was important to make sure that all students gave consent to participate at all times; this was ensured through direct communication between students and the assistants. The assistants and the musician filled out a consent form and gave written informed consent prior to participation. Management of datasets that include personal information of study participants was compliant with the General Data Protection Regulation (GDPR). Procedures for registration and storage of personal data (including sensitive personal data, Swedish: känsliga personuppgifter) were reviewed and approved by KTHs data protection officer (dataskyddsombud@kth.se) and KTHs Research Data Team. None of the video material is published as Supplementary Material. The full Ethics Approval can be obtained from the Swedish Ethical Review Authority or by emailing the first author. The Ethics Approval report includes all consent forms and information for research subjects (Swedish:

8 "Apple Logic Pro", see https://www.apple.com/logic-pro/, accessed May 31, 2023.

9 "Parse a MIDI file into a Tone.js-friendly JSON format", see http://tonejs.github.io/Midi/, accessed January 31, 2023.

10 APA Ethical Principles of Psychologists and Code of Conduct: https://www.apa.org/ethics/code, CODEX: https://www.codex.uu.se/?languageId=1, ALLEA: https://allea.org/code-of-conduct/, SATORI: https://satoriproject.eu/framework/section-1/.

personuppgiftsinformation), as well as the Data Management Plan (DMP) approved by KTH officials.

4. Results

The data from the study is both quantitative (tracked events in the form of MIDI data) and qualitative (transcriptions of the interviews). The data are presented separately in the sections below.

4.1. Tracked events

Boxplots of the chord durations for respective assistants are displayed in Figure 4. Since neither the chord duration nor the absolute deviation data met the assumptions required for a One Way Between Groups ANOVA, Kruskal-Wallis Tests were used as omnibus tests to explore potential differences between assistants. The Kruskal-Wallis test revealed a statistically significant difference in chord durations across the three assistants, $\chi^2(2) = 131.01, p < 0.001^{***}$. *Post-hoc* comparisons using Mann-Whitney *U* tests revealed significant differences in duration between Assistant 1 (A1) and Assistant 3 (A3), with higher mean rank[11] for A1 than for A3 ($MR_{A1} = 229.9$, $MR_{A3} = 117.5$, $U = 8786$, $p < 0.001^{***}$); and between Assistant 2 (A2) and A3, with higher mean rank for A2 than for A3 ($MR_{A2} = 258.5$, $MR_{A3} = 117.5$, $U = 9342$, $p < 0.001^{***}$). No significant difference between A1 and A2 could be observed.

A Kruskal-Wallis test revealed a statistically significant difference in absolute deviation (the difference between subsequent chord duration values, i.e., how much the assistant deviated from a steady tempo for the chord change) across the three assistants, $\chi^2(2) = 118.73, p < 0.001^{***}$. *Post-hoc* comparisons using Mann-Whitney U tests revealed a significant difference in absolute deviation between A1 and A3, with higher median for A1 than for A3 ($Mdn_{A1} = 1.50\,s$, $Mdn_{A3} = 0.14\,s$, $U = 8775$, $p < 0.001^{***}$); and between A2 and A3, with higher median for A2 than for A3 ($Mdn_{A2} = 1.34\,s$, $Mdn_{A3} = 0.14\,s$, $U = 8617$, $p < 0.001^{***}$). No significant difference between A1 and A2 could be observed.

Overall, the tracked event data suggested that although the setup and overall procedure were equivalent for the three assistants (sessions), some tendencies toward different musical strategies could be observed. The following sections highlight the unique qualities of the respective session.

4.1.1. Assistant 1

Before the music started, Assistant 1 (A1) was instructed that "you and the musician will talk and he will show you when to press the key". During the first 2 min, A1 was pressing the key in a manner that resulted in different chord durations and the musician was following the chord changes when they happened. After two minutes, the musician started playing a steady pulse, but this did not prompt A1 to synchronize with the pulse. The music stopped after 5 min and 12 s when the musician made eye contact with the assistant to signal that the music had come to an end.

11 We report mean ranks instead of medians since the assumption of equal distribution shape was not met for chord duration; the Mann Whitney U thus explores differences in distributions, not differences in medians (which is the case for the absolute deviation parameter discussed below).

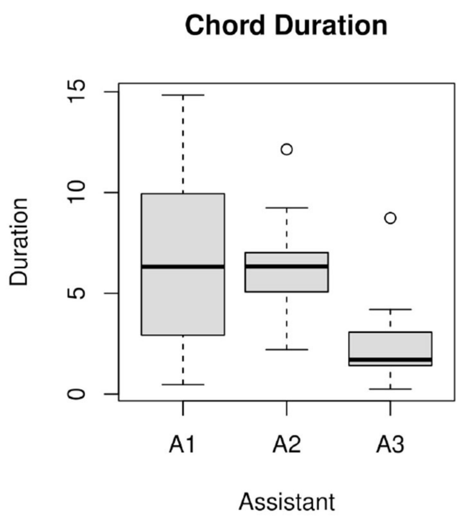

FIGURE 4
Boxplots of chord durations for assistants A1-A3. Whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box.

Analysis of tracked chord onsets revealed that A1 pressed the key 48 times, playing rubato without a steady pulse throughout the session. After one minute the chord durations increased (Figure 5). This corresponds to when A1 started assisting the student more actively. The mean ranks of chord duration, as well as the median of the absolute deviation, were significantly higher for A1 compared to those for A3. Moreover, 12 chords were longer than 10 s, corresponding to instances when the assistant focused completely on the student.

4.1.2. Assistant 2

Before the music started, Assistant 2 (A2) was given the following instruction "You and the musician will agree on when you shall press the key". After that, she waited for the musician to nod before pressing the key, typically one second after the nod. After 30 s, the musician started playing at a steady pulse, but this did not make A2 synchronize with the beat. The music stopped after 4 min and 30 s, when the musician indicated an ending through eye contact, playing slower and slower.

Analysis of tracked onsets revealed that A2 pressed the key 44 times and played rubato without a steady pulse throughout the session. The mean ranks of chord duration, as well as the median of the absolute deviation, were significantly higher for A2 compared to those of A3. Only one chord duration was longer than 10 s (around 2:35–2:45) (Figure 6). This event correspond to when A2 was assisting the student more actively.

4.1.3. Assistant 3

Before the music started, Assistant 3 (A3) was instructed that "The two of you will find a way of playing together so it sounds like music" and then musician told her to start. During the first

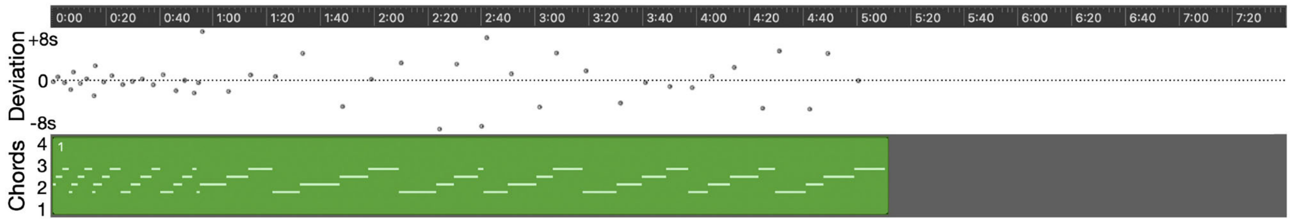

FIGURE 5
Tracked key events for assistant A1. The X-axis represents time and the Y-axis represents, top (white): the deviation (in seconds) between the onsets for each chord, compared to the previous chord, and bottom (green): the duration of each chord in the sequence.

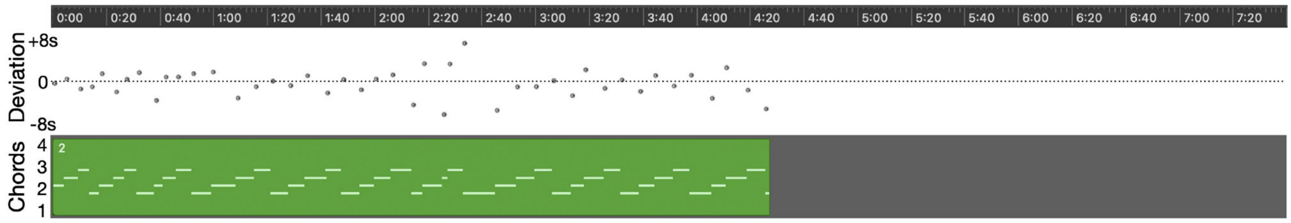

FIGURE 6
Tracked key events for assistant A2. The X-axis represents time and the Y-axis represents, top (white): the deviation (in s) between the onsets for each chord, compared to the previous chord, and bottom (green): the duration of each chord in the sequence.

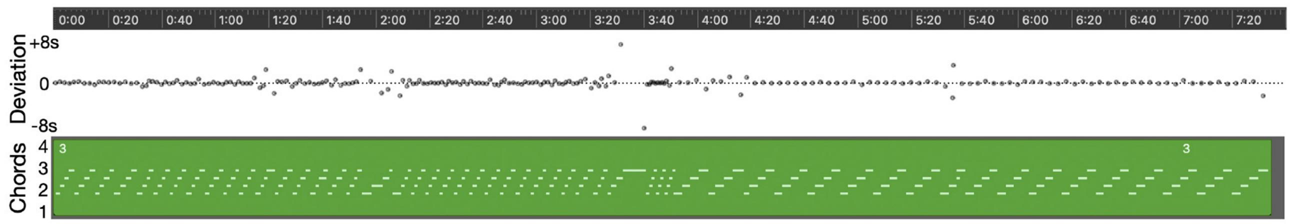

FIGURE 7
Tracked key events for assistant A3. The X-axis represents time and the Y-axis represents, top (white): the deviation (in s) between the onsets for each chord, compared to the previous chord, and bottom (green): the duration of each chord in the sequence.

minute, A3 took the lead and the musician followed. Then the musician started playing at a steady pulse, but this did not make A3 synchronize with the beat. After two minutes, A3 started assisting the student, who had put the instrument cable in her mouth. This caused the chord duration to increase for a while. After 3 minutes and 30 s, A3 and the musician made a pause. The musician then instructed A3 to follow his tempo, indicated by counting to four. A3 responded by pressing the key on every beat. The musician then instructed A3 to only press the key on the first beat in every bar. This caused A3 to immediately increase the chord duration to the length of four beats. The music stopped after 7 min and 35 s when the musician indicated an ending using eye contact, and by playing slower and slower.

The tracking of the onsets revealed that A3 pressed the key 217 times and kept a relatively steady pulse throughout the session, except for the instances described above. After that the musician had given a clear tempo, A3 kept the tempo even during periods when she needed to assist the student. The outliers in the recorded interval data are caused by external factors, like a double trigger and the pause when the musician gave the second instruction (Figure 7). The mean ranks of chord duration, as well as the median of the absolute deviation, were significantly lower for A3 compared to those of A1 and A2.

4.2. Interviews

Themes deemed to be coherent with the study's objective, identified by all three authors, are presented and described in the following sections.

4.2.1. Positive feedback

In general, the assistants all provided positive feedback during the interviews. They said that "it was a fun activity" (A1) and that the "musical result was nice" (A3). The assistants also described the students' experience saying that they "were engaged and enjoyed the workshop" (A1), "did really well" (A2), "were relaxed and focused" (A1), and "liked the instruments" (A1). One of the assistants also expressed that "a live musician brings an atmosphere and a musical flow that the group can join".

4.2.2. The assistants' musical background

With "musical background" we refer to any prior experience of musical activities like singing, dancing, or playing an instrument. All three assistants were asked about their musical background and the answers indicate a range of different backgrounds, spanning from no musical training (apart from music classes at elementary school) (A1), to "I sang in a choir and have always been very interested in music even if I have never given it a chance." (A3). The musician was asked about which one of the assistants he thought had the most musical training and he guessed (correctly) that it was A3.

4.2.3. The physical interface

With "physical interface" we refer to all hardware used in the sessions. The assistants made a few comments about the instruments, the most common one being that "the students were used to the instruments" (A1, A2). These comments also revealed that there are other aspects of the instruments that might attract more attention from the students, apart from their ability to create sounds. For example, one student showed interest in the instrument by placing the instrument cable in her mouth.

4.2.4. Sound and music

With "sound and music" we refer to all aspects of the sound of the instruments and the composition as a whole. The assistants made comments about the master instrument, for example suggesting that it "adds a layer of structure to the music" (A1). A1 said that the sound of the instruments does not necessarily affect the students' playing experience, but A2 thought that a new version with new sounds would probably stimulate curiosity among the students. The interview also revealed that it was hard for the assistants to identify and remember the different sounds of the different instruments. The musician, on the other hand, commented on the harmonic structure of the composition and that it required a steady pulse and a fixed chord length to make sense. He also mentioned that the volume of the instruments was too weak to match the sound of the accordion and that it would be interesting to explore how the instruments could work with other genres of music (for example, he mentioned that the gestures could affect a soundscape sound in various ways).

4.2.5. Musical interplay/collaboration

With "musical interplay and collaboration" we refer to any aspects related to playing music together, using the Funki instruments. The assistants were generally very positive about participating in the music-making sessions and playing together with a professional musician. A1 said that "It is a very nice feeling to have a setting like a band that plays with an external musician that plays continuously" and A3 commented that "My instrument made me more engaged and gave my student more space to play on her instrument, but yet it felt like we were doing it together". A comment that highlighted the need for focusing on more than one thing (on both the student, the musician, and on the master instrument) was made by A2, who said: "My student needs eye contact and the confirmation that we're doing this together". A3 also expressed that "He [the musician] told me how to count and then he followed me really well" indicating that she felt she was in control of the progression of the music. The musician thought "it was hard when it felt like the assistant (A3) was not comfortable with me" but also said that "When I and the assistant play well together, it starts becoming fun for me and I started to get impulses from what the students did". He expressed that the composition was dependent on a steady beat and that he "would have wanted to play the drums instead [of the accordion] to be able to control the pulse".

4.2.6. The role of the musician

A1 and A2 expected the musician to lead and expressed that "we had good eye contact" (A1) and that "the musician was very clear" (A2). Even if the video shows that it was the musician that made a count-in to indicate the tempo, A3 said that "we counted to four, I believe, and he followed me". A2 expressed that she thought the students "listened a lot to the musician". A3 speculated about the potential advantages of playing with a musician, commenting: "If the students would notice when a musician imitates them, I think it will encourage them to keep on playing". The musician commented on his role that "It is a quite demanding task for me as a musician to choose a strategy [depending on the assistant]".

4.2.7. The role of the assistant

The assistants generally describe their role as helping or encouraging the student to play and commented that "I use my hand to guide her to play on the instrument" but A1 also reflected on how she guided the student saying that "I should maybe have let her take more initiatives on her own". Both A1 and A3 pointed out that imitation is important and A3 said that "When she [my student] saw that I did my thing, she also started doing her thing".

4.2.8. The role of the students

The assistants described both the actions of listening to the music and playing music when talking about the students. For example, A2 said: "I can see that she is very focused and interested. She likes to explore." A1 also described: "I think they listened a lot to the musician". A3 mentioned: "She was active and pressed the buttons. She once unplugged the instrument cable but we got it sorted." The assistants also pointed out that the situation with a new person and a new instrument in the room, compared to earlier music-making sessions with the Funki instruments when no strangers had participated, requires a lot of effort from the students. A1 commented "I think she was so fascinated by the instrument cables and the accordion so she forgot to play on her instrument".

She also pointed out that "repetition is very important for our students". One factor that was mentioned in the interviews was that the students' current condition and energy levels can significantly affect their participation and experience overall.

4.2.9. Timing

The composition could be played freely or with a steady pulse. A3 expressed that "It's fun and there is a reward when you keep the beat" and that "It feels like my student also got more active when we started to play in tempo". A1 pointed out that "it is easier to get into the rhythm when you are relaxed" and A3 said that "it was quite simple to keep the beat".

4.2.10. Nonverbal communication

With nonverbal communication we refer to any form of communication using eye contact, nodding the head, using body gestures, or changing the way the musician played on his instrument (through musical gestures). When the assistants were commenting on the nonverbal communication, they focused on the communication between the musician and themselves. For example, A3 described that "he [the musician] sought eye contact". A2 described that "it felt like we had good communication". The assistants also commented on how they interpreted the students' engagement through their body language. A3 described: "I think I saw that they were active when they moved their hand or arm". Although the assistants overall made positive comments about the musicians' non-verbal communication, the musician himself described that "It was obvious when I saw the video with A1 that my body language is very unclear".

4.2.11. Uncomfortable/confused

This theme refers to comments that are related to the problem of not being confident or comfortable with the situation and musical task. For example, A1 and the musician reflected on the need to be relaxed to perform well in the session. A1 said that "I was a bit nervous in the beginning but [then] I understood what to do". She continued: "I [then] became more relaxed and could get more into the rhythm". The musician emphasized the importance of the assistant being comfortable with the situation, in order for the communication to work.

4.2.12. Suggested improvements

All respondents emphasized that it is important that the instruments get a more separated and well-defined sound. The musician also mentioned that it would be interesting to try out "different genres and letting the instruments control different aspects of the composition". The assistants pointed out the importance of having repeated sessions with the same musician, to create a familiar situation for the students. They also suggested that they could have a similar instrument to the student instrument, to better use imitation as a strategy to guide the students. Finally, the musician reflected on how he gave instructions to the assistants and suggested that he should "lead more clearly" by "setting a tempo by counting in".

5. Discussion

The results presented in Section 4 shed light on many aspects that are important to consider when it comes to the assistants' role as facilitators in music-making sessions with students with PMLD involving a professional musician. The quantitative data from the tracked events highlight how the three assistants understand the task differently, and also how different strategies used by the musician, and perhaps also prior musical training, might affect the results. The qualitative data from the interviews add another perspective that further describes how the assistants and the musician experienced the sessions, with hints about aspects that may require improvement. In this section, we discuss the results in terms of how we perceive them as drivers or barriers toward reaching our overarching goal of organizing a joint concert that brings together students from the SEN school with students from a music school with a specific focus on traditional orchestral instruments.

5.1. Drivers

A common theme that all assistants and the musician emphasize is that the activity to play together using the different instruments, as well as the roles assigned in the particular musical interaction explored in this study, was fun and engaging for everyone involved. This is arguably an important quality of both the setting and the system that shall be valued and fostered in future projects. In the current study, the main contributing factors to this positive experience were likely that the students were familiar with the instruments since before, that the assistants could relax (at least after a couple of minutes of training), and the fact that they all got to play together with a live musician. That being said, it is worth evaluating how the different aspects of the setup affect the participants differently and how the roles, as well as the musical responsibilities, can be shared between the participants.

The tracked event data overlaps well with what could be observed in the video from the respective session, as well as the comments from the interviews about the communication between the assistants and the musician. The assistants generally thought that the musician was clear in his body language and that they understood what to do, after a while. A3 differs by having a much more stable pulse than the others, both when playing freely and when the musician instructed her to follow his tempo. It would be interesting to further investigate what factors contribute to this and to what degree prior musical training plays an important role. The musician's reflection that he would probably count-in more directly if he was to repeat the session suggests that he wanted the music to have a steady pulse. This presence of a steady beat was something that also was appreciated by A3 (once it happened).

When asked about improvements, the assistants mentioned "repetition" as an important driver for the students. This is well-known when it comes to learning to play a traditional musical instrument, but it might easily be overseen in a research setting. Having repeated rehearsals would require more logistics and put certain pressure on the personnel at the school, but it would most certainly also greatly affect the outcome, compared to a session that happens only once (without the chance to practice or improve over time).

Another suggestion from the assistants that remains unexplored is to provide instruments in pairs, so that the assistant can demonstrate to the student how to play, using the same gestures and the same output sounds as the students. Such an approach would make sense given the educational setting; the assistants often demonstrate actions using their hands, moving the students' hands to teach them how to perform specific tasks, thereby encouraging interaction.

Computational systems of musical expression always involve the establishment of a stratum that provides certain affordances to the musician, while simultaneously posing constraints (Magnusson, 2010). In this study focused on the Funki instruments, affordances have to do more with usability,[12] whereas constraints define the limits of the musical expression; the mapping can be viewed as a compositional process that engenders a structure of constraints. The Funki instruments allow for music-making with specific *stylistic constraints*, within the limits of a specific genre (Pearce and Wiggins, 2002). There are also *internal constraints* imposed by the system, affecting the logical possibilities of how the music can progress. Finally, there are *external constraints*, i.e., the need to be sure that the instruments are physically possible to play for the students and the teaching assistant. The composition used in this study was based on an E major pentatonic scale and a matching four-chord progression. In other words, the configuration of the Funki instruments did not allow you to play the "wrong note" or notes out of tune; there are fixed notes and there is no possibility to explore tones in between. There are advantages but also potential pitfalls with such a design (see e.g., the discussion about allowing for maximum participation within defined bounds vs. granting greater individual choice – with the risk of increased frustration—presented in Wright and Dooley, 2019). All elements in the composition worked well together even when the assistant and the musician did not plan the performance in advance or synchronize while playing. This compositional approach made it easy for the musician to improvise alongside the sounds from the other instruments, which in turn might have had an impact on the positive responses from the assistants. A strategy that includes a more predefined and score-based composition would likely introduce a new set of drivers and barriers for the participants. Future studies on different compositional approaches (improvisational vs. goal-oriented) and different strategies used for mapping interactions to sounds would be required to identify how different strategies could impact the music-making sessions and how this—in turn—would affect the level of engagement of all participants, in particular the students with PMLD.

5.2. Barriers

A lot of comments from the assistants focused on barriers to a successful session related to the role of the assistant and the students. Particularly interesting in this context is, arguably, that factors that can be drivers may at the same time become barriers. For example, the presence of a professional musician (playing live) was generally appreciated, but at the same time, this new person captured the students' attention so they forgot to play themselves. The assistants also mentioned that the instruments had properties (apart from the sounds) that made them attractive to the students, e.g., one of the students was at times more occupied by the instrument cables than the act of playing the instrument (using button presses). It would maybe be a good idea to reduce the number of distracting factors, for example by making the instruments wireless, in order to encourage the focus on creating sounds. However, considering that the cables were very interesting to some students, removing the cables completely might have the opposite effect, making the students less interested in the activity overall.

We could conclude that organizing repeated rehearsals probably would contribute to reducing various barriers. This could, for example, result in the students becoming even more familiar with the instruments and the sounds they produce, and allow them to become more used to the professional musician (as well as the acoustic instrument played by the musician). Repeated rehearsal and a sense of recognition of the music-making setting could in this sense perhaps help create a more safe setting for the students, which in turn would allow them to more comfortably engage in musical interactions.

Both the assistants and the musician agreed that unclear body language and lack of instructions create an insecure setting, resulting in the assistant easily getting nervous. This introduces barriers for all participants. These barriers, in combination with the driver of having a familiar context for the students, will be important factors to consider when planning the procedure for future workshops and new iterations of the study described in this paper.

One barrier that was mentioned, which could also be seen in the quantitative data, was that the assistants were somewhat split between the task to assist the student and to control the chord progression by hitting the space key on the master instrument. The main reason for this was that this musical interaction required the assistants to look at the musician while simultaneously maintaining eye contact with the student. A3 managed to synchronize with the musician without looking at him, but such an act requires some musical training. The system should ideally not require that the assistants have musical training to work successfully. A different approach that would be interesting to explore would be to let the assistant focus entirely on helping and encouraging the student to play an instrument, and to introduce another facilitator role; yet another trained musician that controls the master instrument and focuses on the synchronization with the professional musician.

Finally, one aspect that became apparent from this study was that the sound design of the instruments needs to be somewhat adjusted when combined with an acoustic instrument. The musician expressed that the volume on the instruments was too low to match his accordion and a general response was that it was hard to identify the sounds of the different instruments. One possible solution to this could be to use an external speaker that matches the sound of the particular instrument. If the role of the

12 The affordances and usability of the Funki instruments have been extensively explored in previous experiments (see e.g. Svahn et al., 2021).

instrument is to play a bass line, the speaker needs to be able to produce low frequencies with an amplitude matching that of an acoustic bass instrument, and if the sound contains a rich spectrum of frequencies, the speaker needs to be able to reproduce all of them with good separation.

5.3. Limitations

We acknowledge that the interactions taking place in the collaborative music-making that is described in this paper are complex to describe, with many different perspectives and factors that may influence one another. In this study, we did not gather data or analyze the interaction between students and their instruments in particular. Instead, the analysis presented in this paper is focusing on the communication between assistants and the professional musician; not on interactions between students and the musician. We have explored the student perspective more in detail in previous work (see e.g., Svahn et al., 2021; Frid et al., 2022). In addition, the Funki ADMIs have been previously validated, and their designs are not a topic of detailed examination in this paper. Although our current paper builds on several previous studies and a multi-year collaboration with the SEN school, this manuscript presents a small-scale study with 3 assistants in total. As such, the results should be interpreted with care, since the small sample size and the long-term collaboration with the assistants affects both generalizability and potential risk for positive bias.

6. Conclusion

This study aimed to explore how a system that was designed for collaborative music-making involving students with Profound and Multiple Learning Disabilities (PMLD) can support assistants in their facilitation of music performance with ADMIs. We carried out three workshop sessions with different assistants together with SEN school students and a professional musician to evaluate the setup using quantitative and qualitative methods. We could identify drivers and barriers for the assistants in using the system, as well as design considerations for similar and continued work focused on collaborative music-making with ADMIs in SEN schools.

One of the drivers that the assistants participating in the study mentioned was how they appreciated contributing to the actual music that was produced during the session. Considering that assistants may have little or no musical training or experience, a system for collaborative playing should thus allow for triggering musical events or changing the performance of the composition without making errors that impact the music. Furthermore, music composed for this setup should allow for unsynchronized, unplanned, and irregular actions from the assistant; else, if the music is more complex or have a goal-oriented structure, it might be better to let a musically trained person be responsible for synchronizing and aligning to a composition or plan.

Another driver mentioned was the possibilities given to allow students to learn by imitating assistants. The assistants stated that the students showed interest in the situation, which may have further positive chain effects such as motivation for repeated rehearsals and increased confidence in participating. To stimulate this, the ADMI design for the assistant should not only be similar to the ones used by the students – thereby promoting learning and enabling them to imitate gestures – but the sounds produced by the ADMIs also need to match each other in terms of volume, timbre, and frequency range, while simultaneously being easy to distinguish from one another, to support the identification of the different sound sources. This is especially important when an acoustic instrument is participating. The hands of the assistant should be in close proximity to the student's own hands during the musical interactions since this allows the assistant to guide the student and encourage participation, demonstrate gestures, and help out when the student gets distracted.

The seemingly biggest barrier for the assistants was an expressed concern about having to split attention between contributing to the music and attending to the student. ADMIs (as well as the composition, and the setup with the musician in the room) should be configured in a way that the assistants do not need to focus visually on the musician or the instrument's interface and thereby lose concentration on the student's gestures and actions. Ideally, the setup should allow the assistant to listen to the sound but maintain visual attention to the student. The main responsibility of the assistant naturally lies with the student and the system should not rely on the assistant's actions to produce a valuable musical output.

Using ADMIs together with a live professional musician and acoustic instruments also results in certain challenges in terms of sound design and composition. We believe that with the new possibilities for collaborations and meetings between different types of players, a system such as the one presented will inspire all participants as long as the resulting musical quality is not compromised. Within this format, even progress toward better artistic inclusion is imaginable.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Swedish Ethical Review Authority. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

MS, JH, and HL contributed to the conception and design of the hardware and software. HL contributed to the music composition, audio production, as well as tagging of audio data. MS, EF, HL, and KF contributed to the conception and design of the study. EF contributed to the literature study and performed the statistical analysis. EF, HL, and KF performed the thematic analysis and wrote sections of the manuscript. MS contributed to writing the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

Funding

This research was supported by a joint doctoral program between KMH and KTH and partly funded by the Swedish Research Council (Vetenskapsrådet), grant number 2020-00343. Funki Instruments are funded by KTH Innovation's Incubator program.

Acknowledgments

The authors wish to thank all participating students and personnel at Dibber Rullen in Stockholm, the accordion player Leif Ottosson, and all who have been involved in designing and testing the instruments.

Conflict of interest

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

Publisher's note

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

References

Barbosa, J., Malloch, J., Wanderley, M. M., and Huot, S. (2015). What does evaluation mean for the NIME community? in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)* (Baton Rouge, LA), 156–161.

Bellamy, G., Croot, L., Bush, A., Berry, H., and Smith, A. (2010). A study to define: profound and multiple learning disabilities (PMLD). *J. Intellect. Disabil.* 14, 221–235. doi: 10.1177/1744629510386290

Berthaut, F., and Dahl, L. (2015). BOEUF: a unified framework for modeling and designing digital orchestras, in *International Symposium on Computer Music Multidisciplinary Research.* Berlin: Springer, 153–166. doi: 10.1007/978-3-319-46282-0_10

Blaine, T., and Fels, S. (2003). Collaborative musical experiences for novices. *J. New Music Res.* 32, 411–428. doi: 10.1076/jnmr.32.4.411.18850

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Calderhead, J. (1981). Stimulated recall: A method for research on teaching. *Br. J. Educ. Psychol.* 51, 211–217. doi: 10.1111/j.2044-8279.1981.tb02474.x

Davanzo, N., and Avanzini, F. (2020). A dimension space for the evaluation of accessible digital musical instruments, in *International Conference on New Interfaces for Musical Expression*. Birmingham: Birmingham City University, 214–220. doi: 10.5220/0009816106200628

Davis, T., Pierson, D., and Bevan, A. (2019). Increasing access to music in SEN settings, in *Proceedings of the 16th Sound and Music Computing Conference (SMC 2019)* (Málaga), 281–286.

Dickens, A., Greenhalgh, C., and Koleva, B. (2018). Facilitating accessibility in performance: Participatory design for digital musical instruments. *J. Audio Eng. Soc.* 66, 211–219. doi: 10.17743/jaes.2018.0010

El-Shimy, D., and Cooperstock, J. R. (2016). User-driven techniques for the design and evaluation of new musical interfaces. *Comp. Music J.* 40, 35–46. doi: 10.1162/COMJ_a_00357

Farrimond, B., Gillard, D., Bott, D., and Lonie, D. (2011). Engagement with technology in special educational & disabled music settings, in Youth Music Report (London: National Foundation for Youth Music), 1–40.

Ferrigon, P., and Tucker, K. (2019). Person-first language vs. identity-first language: an examination of the gains and drawbacks of disability language in society. *J. Teach. Disabil. Stud.* (New York, NY) 70, 1.

Förster, A. (2022). Accessible digital musical instruments in special educational needs schools-design considerations based on 16 qualitative interviews with music teachers, in *International Journal of Human-Computer Interaction* (London: Taylor & Francis), 1–11.

Frid, E. (2019a). Accessible digital musical instrumentsa review of musical interfaces in inclusive music practice. *Multimodal. Technol.* 3, 57. doi: 10.3390/mti3030057

Frid, E. (2019b). *Diverse Sounds: Enabling Inclusive Sonic Interaction. PhD thesis.* Stockholm: KTH Royal Institute of Technology.

Frid, E., Panariello, C., and Nez-Pacheco, C. (2022). Customizing and evaluating accessible multisensory music experiences with pre-verbal children; a case study on the perception of musical haptics using participatory design with proxies. *Multimodal Technol. Interact.* 6, 55. doi: 10.3390/mti6070055

Greenberg, S., and Buxton, B. (2008). Usability evaluation considered harmful (some of the time), in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 111–120. doi: 10.1145/1357054.1357074

Harrison, J. (2020). *Instruments and Access: The Role of Instruments in Music and Disability. PhD thesis*. London: Queen Mary University of London.

Hattwick, I., and Wanderley, M. M. (2012). A dimension space for evaluating collaborative musical performance systems. in *Proceedings of the international conference on New Interfaces for Musical Expression* (Ann Arbor). doi: 10.5281/zenodo.1178281

Hemmingsson, H., Borell, L., and Gustavsson, A. (2003). Participation in school: School assistants creating opportunities and obstacles for pupils with disabilities. *OTJR.* 23, 88–98. doi: 10.1177/153944920302300302

Jack, R., Harrison, J., and McPherson, A. (2020). Digital Musical Instruments as Research Products, in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Birmingham), 446–451.

Johnston, A. J. (2011). Beyond evaluation: linking practice and theory in new musical interface design in *Proceedings of the International New Interfaces for Musical Expression Conference (NIME).* Oslo.

Jordà, S. (2005). Multi-user instruments: models, examples and promises, in *Proceedings of the 2005 Conference on New Interfaces for Musical Expression*, 23–26.

Kvifte, T., and Jensenius, A. R. (2006). *Towards a Coherent Terminology and Model of Instrument Description and Design*. Paris.

Light, J. (1988). Interaction involving individuals using augmentative and alternative communication systems: State of the art and future directions. *Augment. Alternat. Communicat.* 4, 66–82. doi: 10.1080/07434618812331274657

Lindetorp, H., and Falkenberg, K. (2022). Evaluating Web Audio for learning, accessibility, and distribution. *JAES.* 70, 951–961. doi: 10.17743/jaes.2022.0031

Lucas, A., Harrison, J., Schroeder, F., and Ortiz, M. (2021). Cross-pollinating ecological perspectives in ADMI design and evaluation, in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Shanghai, China.

Magnusson, T. (2010). Designing constraints: composing and performing with digital musical systems. *Comp. Music J.* 34, 62–73. doi: 10.1162/COMJ_a_00026

O'Modhrain, S. (2011). A framework for the evaluation of digital musical instruments. *Comp. Music J.* 35, 28–42. doi: 10.1162/COMJ_a_00038

Orchestras Live (2023). Able Orchestra.

Östlund, D. (2017). Elevassistenter: En möjlighet eller ett hinder för elevers inkludering och delaktighet. (Kristianstad) 37, 106–117.

Pearce, M., and Wiggins, G. A. (2002). Aspects of a cognitive theory of creativity in musical composition, in *Proceedings of the ECAI02 Workshop on Creative Systems.* Lyon.

Porter, G., and Cafiero, J. M. (2009). Pragmatic organization dynamic display (PODD) communication books: a promising practice for individuals with autism spectrum disorders. *Perspect. Augment. Altern. Commun.* 18, 121–129. doi: 10.1044/aac18.4.121

Rodger, M., Stapleton, P., Van Walstijn, M., Ortiz, M., and Pardue, L. (2020). What makes a good musical instrument? a matter of processes, ecologies and specificities, in *Proceedings of the International Conference on New Interfaces for Musical Expression.* Birmingham, UK: Birmingham City University, 405–410.

Sharma, U., and Salend, S. J. (2016). Teaching assistants in inclusive classrooms: A systematic analysis of the international research. *Aust. J. Teach. Educ. (Online).* 41, 118–134. doi: 10.14221/ajte.2016v41n8.7

Small, C. (1998). *Musicking: The Meanings of Performing and Listening.* Middletown, CT: Wesleyan University Press.

Steinmeier, C., Becking, D., and Kanders, M. (2022). The perfect musical instrument does not exist: Experience reports for the development of accessible nimes, in *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments* (Corfu), 507–513.

Stowell, D., Robertson, A., Bryan-Kinns, N., and Plumbley, M. D. (2009). Evaluation of live human-computer music-making: quantitative and qualitative approaches. *Int. J. Hum. Comput. Stud.* 67, 960–975. doi: 10.1016/j.ijhcs.2009.05.007

Svahn, M., Hölling, J., Curtsson, F., and Nokelainen, N. (2021). The Rullen band, in NIME 2021. in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Shanghai, China.

Wanderley, M. M., and Orio, N. (2002). Evaluation of input devices for musical expression: Borrowing tools from HCI. *Comp. Music J.* 26, 62–76. doi: 10.1162/014892602320582981

Ward, A., Woodbury, L., and Davis, T. (2017). *Design Considerations for Instruments for Users with Complex Needs in SEN Settings* (Copenhagen).

Waters, S. (2007). Performance ecosystems: ecological approaches to musical interaction, in *EMS: Electroacoustic Music Studies Network* (Leicester), 1–20.

Welch, G. F., Ockelford, A., Zimmermann, S.-A., Himonides, E., and Wilde, E. (2016). The provision of music in special education (PROMISE) 2015, in *Procedings of the International Perspectives on Research in Music Education.* Reading: International Music Education Research Centre (iMerc) Press.

Wright, J., and Dooley, J. (2019). On the inclusivity of constraint: Creative appropriation in instruments for neurodiverse children and young people, in *Proceedings of the New Interfaces for Musical Expression Conference (NIME).*

frontiers | Frontiers in Computer Science

TYPE Technology and Code
PUBLISHED 04 October 2023
DOI 10.3389/fcomp.2023.1113078

Check for updates

OPEN ACCESS

EDITED BY
Emma Margareta Frid,
Royal Institute of Technology, Sweden

REVIEWED BY
John Sullivan,
Université Paris-Saclay, France
Andrew McPherson,
Queen Mary University of London,
United Kingdom

*CORRESPONDENCE
Asha Ward
award@aub.ac.uk

RECEIVED 01 December 2022
ACCEPTED 12 September 2023
PUBLISHED 04 October 2023

CITATION
Ward A (2023) The development of a Modular Accessible Musical Instrument Technology Toolkit using action research.
Front. Comput. Sci. 5:1113078.
doi: 10.3389/fcomp.2023.1113078

The development of a Modular Accessible Musical Instrument Technology Toolkit using action research

Asha Ward*

Research and Development, Innovation Studio, Arts University Bournemouth, Bournemouth, United Kingdom

Within the field of digital musical instruments, there have been a growing number of technological developments aimed at addressing the issue of accessibility to music-making for disabled people. This study summarizes the development of one such technological system—The Modular Accessible Musical Instrument Technology Toolkit (MAMI Tech Toolkit). The four tools in the toolkit and accompanying software were developed over 5 years using an action research methodology. A range of stakeholders across four research sites were involved in the development. This study outlines the methodological process, the stakeholder involvement, and how the data were used to inform the design of the toolkit. The accessibility of the toolkit is also discussed alongside findings that have emerged from the process. This study adds to the established canon of research around accessible digital musical instruments by documenting the creation of an accessible toolkit grounded in both theory and practical application of third-wave human–computer interaction methods. This study contributes to the discourse around the use of participatory and iterative methods to explore issues with, and barriers to, active music-making with music technology. Outlined is the development of each of the novel tools in the toolkit, the functionality they offer, as well as the accessibility issues they address. The study advances knowledge around active music-making using music technology, as well as in working with diverse users to create these new types of systems.

KEYWORDS

accessible digital musical instruments (ADMIs), action research, DMI, music technology, special educational needs (SEN)

1. Introduction

The use of technology to facilitate accessibility to music-making has seen growing interest in the last two decades (Frid, 2018, 2019). The rise of new instruments for musical expression (NIMEs) including digital musical instruments (DMIs) and accessible digital musical instruments (ADMIs) has been concurrent with access to facilitating technologies. Cheaper microprocessors, such as Arduino (2015), Raspberry Pi (Raspberrypi.org, 2020), Bela Boards (Bela, 2023), and Teensy Boards (Stoffregen, 2014) alongside an increase in the availability of sensors and implementation with prototyping technologies such as 3D printing (Dabin et al., 2016), have heralded the development of many new music-making tools (Graham-Knight and Tzanetakis, 2015; Kirwan et al., 2015). Authors have covered a range of topics related to ADMIs, including design and development (Förster et al., 2020; Frid and Ilsar, 2021), targeting specific populations (Davanzo and Avanzini, 2020a), evaluation of NIMEs (Lucas et al., 2019; Davanzo and Avanzini, 2020b), evaluating the use of electronic music technologies (EMTs) (Krout, 2015), integration into specific settings (Förster, 2023), and discussion about customization, augmentation, and creation of tools for specific users (Larsen et al., 2016). The literature provides insights into the challenges and opportunities of creating ADMIs and offers recommendations for future research and development. These challenges have historically included fear, dislike, or indifference to technology (Farrimond et al., 2011); lack of confidence when putting technology into practice (Streeter, 2007); musical output that can be seen as uninspiring, artificial, and lacking expression (Misje, 2013); and impersonal and lacking sophistication (Streeter, 2007). Technology can be seen as a barrier when coupled with a lack of formal training and exposure (Magee, 2006) and the perceived need for insider knowledge upon use (Streeter, 2007). More current reviews of the perceptions of these types of tools conducted with electronic musicians indicate that many of these issues maintain currency. Research by Sullivan and Wanderley (2019) suggests that DMI durability, portability, and ease of use are of the greatest importance to performers who use DMIs. Frid and Ilsar (2021) found several frustrations in their survey of 118 electronic musicians (40.68% professional musicians and 10.17% identifying as living with a disability or access requirement) including software and hardware limitations, time-consuming processes, need for more or new interfaces, and physical space requirements for the equipment.

Presented here is the development of a technological system, the Modular Accessible Musical Instrument Technology Toolkit (MAMI Tech Toolkit), that is rooted in addressing some of these issues and leveraging the unique power of technology to provide an accessible way for users to actively participate in music-making. The research looked at barriers with current technology for music-making and focussed on creating easy-to-use, small form factor, wireless technology. This study has a particular focus on moving away from using a screen and into tangible, physical, flexible, and tailorable tools for active music-making. The contributions of this study center around the description of an action research methodology used to develop the toolkit. There is an emphasis throughout the study on the use of participatory and iterative methods to elicit current issues and to explore barriers with music technology. These issues and barriers then form the basis for developing novel tools to address these gaps in provision. These contributions advance knowledge around music technology in facilitating access to active music-making for diverse user groups.

2. Related work

2.1. New tools for music-making

New tools for music-making can be classified in a variety of ways. Wanderley (2001) proposed the classification of NIMEs in three tiers: instrument-like controllers—where the input device design tends to reproduce features of existing (acoustic) instruments in detail—as seen in developments such as Strummi (Harrison et al., 2019) and John Kelly's "Kellycaster" (Drake Music, 2016); augmented instruments—where the instrument is augmented by the addition of sensors to extend functionality—such as Electrumpet (Leeuw, 2021); and alternate controllers—whose design does not follow one of the established instruments—for example, the Hands (Waiswisz, 1985). These new tools can be configured in a variety of ways—as both controllers with processing being achieved via separate computers or with onboard processing and sound production—the latter being exemplified in the Theremini (Moog Music Inc, 2023). These systems can vary from the use of off-the-shelf components used as controllers (Ilsar and Kenning, 2020) to entirely bespoke assemblages of hardware and software with the ability to integrate (or not) with existing music technologies. They may be purely research artifacts or commercial products (Ward et al., 2019). Frid (2018) conducted an extensive review of ADMIs presented at the NIME, SMC, and ICMC conferences and identified seven control interface types including tangible controllers. Using these classification systems, the tools in the MAMI Tech Toolkit can be classified as tangible (Frid, 2018), touch-based, and alternate (Wanderley, 2001) controllers.

2.2. Commercially available technology

There are commercially available technologies designed for accessibility such as the Skoog (Skoogmusic, 2023), Soundbeam (Soundbeam, 2023), Alphasphere (AlphaSphere, 2023), and Musii (Musii Ltd, 2023). The commonality of these instruments is that they facilitate different modalities of interaction that move away from the traditional acoustic instrument paradigm—i.e., there are no strings to pluck, skins to percuss, reeds to blow, pipes to excite, or material to resonate—there are also no screens to interact with to play them. A more reductive categorization can be used to describe what is left as facilitating interactions via the use of touch or empty-handed gestures. One example of a touch-based controller is the Skoog. The Skoog is a cube-shaped controller with a "squishy" velocity-sensitive "skin" that when depressed can be used to control sounds as dictated within its software (requires computer/tablet connection), with the ability to control the sound produced as well as constraints around these sounds (i.e., scale and key). The Soundbeam is an empty-handed gestural interface that uses ultrasonic beams as "an invisible keyboard in space" (CENMAC, 2021) to trigger notes or samples when the beam is broken. The Soundbeam is a longstanding ADMI that has been in use since the late 1990s featuring integrated synthesis/processing—connecting to speakers/monitors for sound production. The Soundbeam has been a transformative tool in over 5,000 special education settings and adult day centers (Swingler, 1998). Another example is the Alphasphere, a spherical device 26 cm in diameter that houses 48 velocity-sensitive pads over the surface. It is a MIDI controller with proprietary software. The Musii is a 78 cm × 92 cm multisensory interactive inflatable featuring three "prongs" that change color and modulate sound when interacted with (pushed in or squeezed)—with the ability to change the sounds and lights remotely, as well as an integrated vibrating speaker.

2.3. Music-making apps

Apps have been used to facilitate access to music-making through the use of the touch screen and integrated sensors. Tablets (in particular iPads), which are commonly found in school settings, have been used to facilitate music therapy (Knight, 2013; Krout, 2014). The ThumbJam app (Sonosaurus, 2009) contains 40 sampled instruments, hundreds of scales with playing styles, arpeggiation, and a customisable graphical user interface (GUI) (Matthews, 2018). Also offered is the ability to loop; add effects; manipulate sonic content; create instruments; and record, import, and export data. Another example of a generic music app includes Orphion (Trump, 2016) which offers sound-making possibilities through interacting with the screen and/or the motion of the device. There are also tablet versions of software such as GarageBand that offer direct touch-based access to sound creation and manipulation via the GUI, providing a different mode of access than using a computer mouse and keyboard.

2.4. Organisations

Several organizations and charities provide much-needed help in navigating the use and integration of these tools. Charities such as Drake Music (Drake Music, 2023), OpenUp Music (OpenUp Music, 2023), and Heart n Soul (Heart n Soul, 2023) facilitate music-making through technology and inclusive practices for disabled musicians, often supporting the development and implementation of technology. One development that sprung from OpenUp Music is the Clarion (OpenOrchestras, 2023) which is an accessible music technology that can be integrated with eye gaze systems (detecting users' eye movements) or used on the iPad, PC, or Mac. Clarion features the ability to create a GUI to suit the user and customizable sounds that can be mapped in different ways to allow expression. This type of customizable and scalable mapping is a leap forward in moving from an instrument-first paradigm to a human-first approach in which instruments are built around the user, rather than the other way around.

2.5. Developing for and with disabled users

The toolkit was developed in line with the underlying principles of the social model of disability. Oliver (1996) states that "the individual (or medical) model locates the 'problem' of disability with the individual" (ibid, p. 32), whereas the social model of disability "does not deny the problem of disability but locates it squarely with society" (ibid, p. 32). These principles highlight the importance of asking people what they need and utilizing this to develop design directions (Skuse and Knotts, 2020). Paramount is ensuring what is centered is the perspectives of disabled people and of those around them who know them very well (albeit with an aside that there will always be issues with taking the opinions of others in proxy of the users themselves) and conducting "works that propose a position of pride against prejudice" (Ymous et al., 2020, p. 10). At the forefront of this research was the crucial aspect of making the research accessible to the participants both in choosing methods of gaining data, and then feeding that data back to participants in a way that suited them and matched their needs. This is in alignment with established research methodologies developed by Mack et al. (2022) in the field of human-centered methods.

2.6. Constraints and expression

When developing new technological tools for music-making, particularly alternate controllers, there is a decoupling of the sound production from the sound generation mechanism. This bond is found in traditional instruments where the acoustic properties of the instrument are in direct relationship with the physical construction of the instrument. This decoupling needs careful reconstruction to ensure that users can interact with the tool, the output is engaging, and it is suitable and understood by the user. As Bott (2010) identifies, distinguishing between access needs and learning needs is key to determining musical possibilities with an individual. These can often be interrelated, but making a distinction can start to cut through what might otherwise seem to be impenetrable complexities (ibid) when designing tools to suit individual users. Instruments and tools can be constrained by design to suit the needs of the user. The unique qualities of technology can provide opportunities to configure systems to users and constraints can be put in place to scaffold this. Sounds can be designed that are motivating to the user and mechanisms to play these can be constructed to suit the particular interaction paradigm of the user—for example, small gestures can be translated into vastly amplified sound or equally large gestures can be mapped to nuanced control of sound.

The issue of expression and constraints in the creation of DMIs has been explored in broader discourse such as the study by Jordà (2004) who discusses balancing relationships of "challenge, frustration, and boredom" (ibid, p. 60) with stimulation and placation. Balances must be struck between the learning curve, how users can interact, and what the user can ask the instrument to do. Technology use can range from using a single button to trigger one set of notes to instruments that have a one-to-one mapping of gesture to sonic output (such as the Theremin). There can be benefits, limitations, and creative playing opportunities within constrained interactions (Gurevich et al., 2010). This can be balanced against the time needed and learning curves faced, as less constrained instruments can require hours of practice to attain virtuosity (Glinsky, 2005). This can be a barrier that can alienate some users, however, heavily constrained instruments may quickly lead to mastery and boredom. The tools in this toolkit explore the lines between what Jack et al. (2020) quoted from Wessel and Wright (2002), the range of interaction through "low entry fee to no ceiling on virtuosity" (Jack et al., 2020, p. 184). Considering these elements can aid in pointing to "what might be considered essential needs for different types of musicians" (Jordà, 2004. p. 60). It is recognized that whilst constraining musical output may inspire creativity and remove frustrations by initially opening accessibility—as seen in the study of Gurevich et al. (2010)—there may be demotivation in the long run if the constraints hamper the players' ability to play the way they wish.

2.7. Third-wave human–computer interaction

The toolkit is situated in third-wave human–computer interaction (HCI) in which the focus on task-based efficiency of information transfer and technology-centric operation of computer systems by humans is interchanged for a more socially situated and embodied view of the user in a "turn to practice" (Tzankova and Filimowicz, 2018). This practice paradigm looks at HCI through a lens of real-life unfolding, processes situated in time and space, and technology as one aspect of the situation (Kuutti and Bannon, 2014). Third-wave HCI engages users through "in the wild" contexts (Bødker, 2006), with an emphasis on human meaning-making, situated knowledge, and grappling with the full complexity of the system (Harrison et al., 2007). This third wave "considers the 'messy' context of socially situated and embodied action which introduces humanistic and social science considerations into design research" (Tzankova and Filimowicz, 2018, p. 3). The practice paradigm features *in-situ*, extended activities involving people and artifacts within their daily practices, within their organizational routines, and with more developmental and phenomenological orientations being used (Kuutti and Bannon, 2014). The research described in this study saw both the development and practical use of the tools embedded within the context of use to design *for and with* central users and their facilitators through an action research methodology. At the forefront were activities that held meaning for the participants and these meaning-making activities were conducted to inform the development of the toolkit whilst incorporating their feedback through real-world usage scenarios.

2.8. Research through design

In carrying out this research with a user-centric position and the aim of creating new designs with accessibility in mind, it is relevant to acknowledge the relation of this study to other fields in which similar questions have been asked. Research through Design (RtD) is a field that aims to understand users' perspectives, behaviors, and feedback. This is then used to inform the design process and ensure that the final design solution meets users' needs, preferences, and expectations. The toolkit shares the sentiment of being "an attempt to make the right thing: a product that transforms the world from its current state to a preferred state" (Zimmerman et al., 2007) which is underlined within the methodology of RtD. This also aligns with the tenets of action research (finding improved solutions in current situations). The prototypes used design as a means of inquiry—as commonly seen in RtD. Design practice was applied to situations with significant theoretical and practical value to the users.

2.9. Technology prototypes as probes

In this project, prototypes were used as "technology probes" (Hutchinson et al., 2003). Technology probes can be defined as simple, adaptable, and flexible with three main goals:

- A social science goal of "collecting information about the use and the users of the technology in a real-world setting" (Hutchinson et al., 2003, p. 18).
- An engineering goal of field testing the technology (ibid).
- A design goal of inspiring users and designers to think of new kinds of technology (ibid).

Technology probes can be presented in various fidelities from low-fi physical hardware prototypes to rudimentary examples of the system in use, or as fully functioning prototypes. These become more refined as the research progresses and are informed by interactions with stakeholders. These types of technology probes have also been called "products" (Odom et al., 2016) in RtD, and they can be used to support investigations into distinct kinds of experiences, encounters, and relationships between humans and interactive technology (ibid). Other researchers have designed ADMIs using products (Jack et al., 2020) and outlined them as having four key features. These features are that products should be *inquiry-driven* and designed to "provoke users in their environment to engage them in an enquiry" (ibid p. 2); have a *complete finish* with a focus not on what the design could become but what the artifact is; be able to *fit convincingly* into the everyday scenario they are designed to be used within; and be able to be *used independently* of the researcher (ibid). Probes and products are useful when considering the MAMI Tech Toolkit as the three technology probe goals describe what the toolkit aimed to elicit throughout the research journey, whilst the four key features of products embody how it was ultimately designed for the final iteration.

3. Methods

3.1. Research sites and stakeholder involvement

The toolkit was developed over 5 years with four research sites (three special educational needs schools and 1 day center for adults with disabilities) as part of an engineering doctorate undertaken by the author. As part of this doctorate, the main school (school A in Table 1) was the industrial sponsor of the research. An industrial mentor (the creative technologist at school A) was assigned to mentor the researcher in the development of the project. A team of stakeholders (Table 1) whose practice is directly related to the research was also used to assist with the research. Outlined are the roles of the stakeholders, the research site that they worked within, and the activities that they were involved in as part of the research.

Eleven practitioners (industrial mentor, class teacher/head of music, assistant head teacher, digital media and sensory support technician, music therapists, digital music technician, director of music, musician, music technologist, and community musician) and nine children and young people (CYP) participated in the research as stakeholders. A limitation of this research is that the specific needs and ages of all the CYP involved were not gathered due to a combination of ethical, logistical, and time constraints.

The fieldwork involved consulting, observing, and engaging with practitioners and CYP in their naturalistic environments. Some practitioners already used music technology within their practice. This allowed us to conduct case studies to inform the creation of technical specifications. Some practitioners did not use technology, in these cases, insight was gained into the role technology might be able to play in enabling active music-making. These cases were also helpful for considering how to integrate technology with more traditional approaches, and how traditional acoustic and digital musical instruments could be combined. Direct requests from stakeholders were gathered

TABLE 1 Stakeholders, sites, and involvement.

Representational stakeholders	Code	Research site	Research activity	Action research cycle
Industrial mentor	IM	School A	Meetings/group meeting/sessions	1, 2, 3
Class teacher/head of music	CT	School A	Group meetings/sessions	1, 2
Assistant head teacher	AHT	School A	Group meetings	1, 2
Digital media and sensory support technician	DMSST	School A	Group meetings	1
Music therapist	MTA	School A	Group meetings	1, 2, 3
Digital music technician	DMTB	School B	Meetings	2, 3
Director of music	DoMB	School B	Meetings	3
Musician	MC	School C	Meetings/session observations	3
Music technologist	MTC	School C	Meetings/one-to-one sessions	3
Music therapist	MTDC	Day center	Meetings/session observations	3
Community musician	CMDC	Day center	Meetings/session observations	3
Seven children and young people	n/a	School A	Group sessions using commercial technology	1
Child one	CO	School A	Using bespoke technology within a group session	3
Child two	CT	School A	Using bespoke technology within a group sessions	3

including specific desires for features and functionality, alongside explorations into issues of form, function, and context of use. From this, a set of design considerations were developed (Ward et al., 2017). Prototypes of the tools were presented to stakeholders throughout—most regularly to the industrial mentor (a person appointed by the industrial sponsor school as a lead stakeholder). These prototype tools were used as probes to iteratively provide physical manifestations of the stakeholder's requirements and were used to both gain their feedback and engage them to think of the next design steps to refine designs further or add features or functionality.

3.2. Action research

Action research (AR) was the methodology used to conduct the research. AR seeks to bring together action and reflection, theory, and practice, in a participatory process concerned with developing practical solutions to pressing concerns for individuals and communities (Reason and Bradbury, 2008). The focus of AR is to create a democratic atmosphere whereby people work together to address key problems in their community to develop knowledge embedded in practical activities for human flourishing (ibid). In the case of this research, the pursuit was providing access to music-making through the mobilization of the expertise of the stakeholders. This informed the final designed outcome and situated the tools within their context of use. The use of AR allowed practitioners to be added into the research as stakeholders as their expertise was needed and the flow of the design process to follow organically the needs of those at the center of the research.

This "orientation to inquiry" (Reason and Bradbury, 2008, p.1) involved working closely in collaboration with stakeholders through phases of *planning, action, and reflection*. These three phases form the cycles of AR with the outcome of each cycle informing the activities of the next cycle. The first cycle aimed to conduct a situation analysis. In the *planning* phase, the stakeholders work with the researcher to put a plan of action together. In the *action* phase, the research activity is conducted, stakeholders are interviewed, and the data are summarized and analyzed. In the *reflection* stage, "researchers and practitioners reflect on, and articulate lessons learned and identify opportunities for improvement for subsequent research cycles" (Deluca et al., 2008, p. 54). Reflection on both techniques and methods used occurs to allow contributions to knowledge to begin to form. The first cycle of AR then feeds into the next cycle's planning stage. These basic steps are then germane to further cycles. Although this research is separated into distinct cycles, there was substantial overlap between cycles with an element of feedback and feedforward that informed the iterative development of the tools—as reflected in the red arrows (Figure 1).

3.3. Data collection

The research approach used throughout the data collection is an example of "starting where you are" (Lofland and Lofland, 1994; Robson, 2002, p. 49), in that it began through personal interest and with several connections to stakeholders already in place. The research endeavored to *meet people where they were*, following stakeholders' leads where possible. The use of this method connects with the underpinning research methodology of action research—a methodology focused on creating solutions to problems that are pertinent to the lives of stakeholders within the research, using cycles of action and reflection to develop practical knowing (Reason and Bradbury, 2008)—see Section 3.2 for more details. This meant both physically going to their sites of practice and interacting and elucidating information through naturalistic methods such as observation of practice, and relaxed and open discussion about their practice. The aims and research agenda also remained open

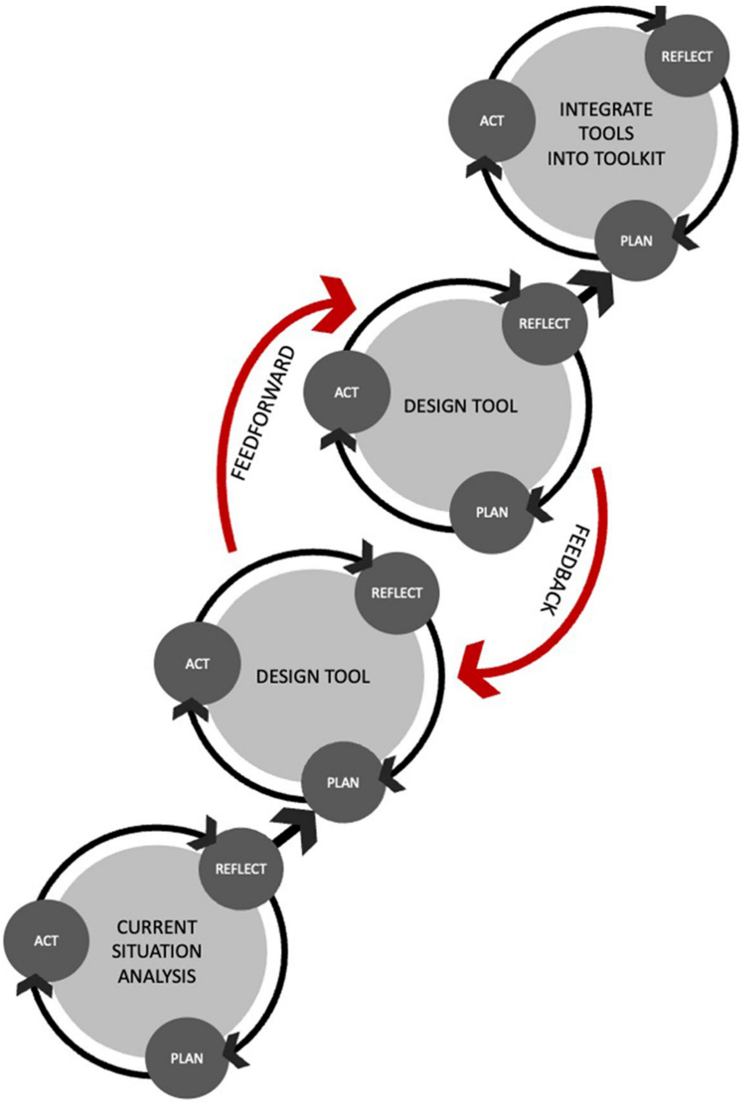

FIGURE 1
Four completed cycles of action research (adapted from Piggot-Irvine, 2006).

to follow the needs of the stakeholders and to be able to focus on what they considered important.

The data collection methods used for this research were an amalgamation of both Creswell (2014) and Yins (2018) categories of evidence sources. Yins (2018, p. 114) sources used within this research were documentation (emails, agendas, key points synthesized, and observational notes from stakeholder interactions), interviews (face-to-face one-to-one in person, telephone conversations, emails, video calls, and focus groups), direct observations (in the sessions using technology, observation of practitioners using music therapy in their sessions, and observing CYP in music therapy), and physical artifacts (technology probes) (Hutchinson et al., 2003). Creswell's (2014) table of qualitative data collection types further illustrates sources of data collection including self-reflective documentation (self-memos on tech developments, on research and analysis, and recording of events), technical documentation (journaling technical development using e-notebook and GitHub), and audio-visual material (videos in sessions, photographs, and audio recordings).

In total, six sessions of CYP using off-the-shelf music technology were conducted throughout the four cycles of action research. These were group sessions with seven CYPs in two separate age groups—primary aged 5–11 years and secondary aged 11–16 years. Also conducted were three focus group meetings with a consistent group of practitioner stakeholders; 29 meetings with the industrial mentor; around eight separate meetings with individual practitioners; three observations of practitioners using music therapy in a live session; and 19 music therapy sessions incorporating elements of the toolkit with two consistent CYP.

All sessions were planned with the relevant stakeholders. Observational field notes were taken by the author and then analyzed. These key points were then provided as feedback to stakeholders covering issues and successes and developing a plan for the next iteration. For observing practitioners running their sessions, field notes were taken by the author and then analyzed to provide data on how to develop technology, how technology was being used, or how it could be integrated. Interviews/meetings were conducted face-to-face on a one-to-one basis, over the telephone, via email, and with focus groups. Focus groups were used initially to allow the opening of communicative spaces (Bevan, 2013) for the interdisciplinary stakeholders, and to facilitate open discussion around what had been done previously in the setting, what was being done, and what could be done next. The meetings were recorded and transcribed before analysis.

A literature review was also conducted and updated throughout the research to identify design considerations/requirements for the toolkit. Searches of Google Scholar, Google, and The Bournemouth University Library Catalog were used with the following keywords: music technology for music therapy, new interfaces for musical expression, music technology and special education needs, music technology SEN, and music technology complex needs. The Nordoff Robbins Evidence Bank (2014) (specifically account no. 16) and Research and Resources for Music Therapy 2016 (Cripps et al., 2016) were also consulted. The selection of articles expanded as the literature was reviewed. Articles were scanned for significance pertaining to the use of technology (both novel or off-the-shelf) with users with complex needs for active music-making or sonic exploration, or that they featured details of such technologies in use, or that they explored issues around and/or reviewed music technology in use. Some gray literature was also consulted (O'Malley and Stanton Fraser, 2004; Department for Education, 2011; Farrimond et al., 2011; Ofsted, 2012) as this provided a different perspective on technology usage in practice.

3.4. Data analysis

Thematic analysis was used to analyse the data gathered throughout the research using a process as described by Saunders et al. (2015, p. 580). This was done solely by the author. Data were coded manually on an activity-by-activity basis, with codes being grouped into themes. This was done after each activity to inform the technical development of the toolkit and then these themes were further analyzed more in-depth for the doctoral thesis. Both inductive and deductive methods were used during coding using a data-driven approach. Codes were assigned by the author based on describing the unit of data, and themes were then grouped by describing the overall subject of the individual codes to allow the data to be presented to stakeholders as a report of key findings.

3.5. Action research cycles

Four cycles of action research were conducted within the project span (see Figure 1).

3.5.1. Cycle one

The first cycle aimed to understand the current situation by focusing on working with stakeholders to explore the current use of technology and the issues surrounding this use. Existing technology for active music-making was explored through the use of the Orphion app (Trump, 2016) on an iPad connected to an amplifier within the six sessions with CYP. The findings from these activities were used to inform the developments within the cycle two.

3.5.2. Cycle two

The second cycle aimed to develop the first two tools in the toolkit: filterBox and the squishyDrum. These were iteratively prototyped with weekly input (nine meetings) from the industrial mentor and prototypes were presented to some of the stakeholders at two separate focus groups (2 months apart). Within this cycle, an additional school (B) came on board as stakeholders.

3.5.3. Cycle three

Cycle three aimed to develop another bespoke tool—The Noodler—in close collaboration with a music therapist at school A. In this cycle, the research moved away from creating tools that purely suited individual users into the direction of creating technology that worked for practitioners as these practitioners facilitate the use of technology within the setting. This allowed a "zooming out" to occur to see the tool as embedded within a bigger context, in recognition that music-making in the settings, and with the users involved in this research, constituted the messy real-world scenario described in the third-wave of HCI (Bødker, 2006). This necessitated the inclusion of input from more stakeholders, observing their practice closely, and integrating the tools into practice to inform the final designs. Other practitioners were consulted from a third site (School C) and a fourth research site (day center). Meetings with stakeholders from School B occurred and sessions were observed with each of the music therapists, community musicians, and musicians from all of the research sites. These activities were conducted to determine what a usual session might entail and where technology might be able to be deployed, or how it was being integrated if it was being used. Sessions were held with two children at school A spanning around 6 months of weekly sessions (in term times) using the Noodler—to aid in the development of the software and mappings for that specific tool, which could be classed as the most evolved design because of this extended contextualized use. Meetings were conducted with the industrial mentor to discuss the developing tools.

3.5.4. Cycle four

Cycle four aimed to finalize the prototypes of the tools in the toolkit alongside the software elements needed to create a stand-alone kit. This cycle featured analysis and integration of emergent concepts from preceding AR cycles into the finalized toolkit. A fourth bespoke tool (the touchBox) was developed during this cycle as a stand-alone tool that did not require a computer connection and had onboard sound production. Further to this was the development of the tools into a cohesive toolkit. Software to accompany the filterBox, squishyDrum, and Noodler was finalized, alongside an overall software application from which to launch the individual pieces of software. Resources for use were also developed—such as a quick start guide and user manual—and iPad connectivity was integrated.

10.3389/fcomp.2023.1113078

FIGURE 2
Tools in the MAMI tech toolkit. **(Top left)** filterBox, **(Top right)** squishyDrum, **(Bottom left)** Noodler, **(Bottom right)** touchBox.

4. Results

4.1. Technical design

Three of the tools feature a software component thus requiring a computer for processing and sound production, and one tool is stand-alone with an integrated microprocessor, screen, and speaker. Three of the tools in the toolkit (filterBox, squishyDrum, and the Noodler) consist of hardware and software elements as they connect wirelessly (using nRF24L01radio connected to Arduino microprocessor boards as transmitters and receivers). One tool (Touchbox) featured a Teensy 3.2 microcontroller with an audio adaptor. Max/MSP (Cycling'74, 2023) was used to develop the software for the system which was then turned into an application that could run on a Mac/PC tested at technical specification levels found on the school computers, and the ability to run directly from USB stick to ensure installing would not be a barrier. The audio featured synthesized sounds using Max/MSP via embedded virtual studio technology (VST) plug-ins or Max/MSP's synthesis engine, sample triggering (short clips of sounds or music), or musical note triggering via MIDI. The final toolkit featured an underpinning software element developed by the industrial mentor that was augmented and extended considerably to only work specifically with the tools in the toolkit and formed the bedrock that the toolkit software was built-on. A video demonstration of the toolkit in action can be found here: https://youtu.be/HIce1nJX4Us.

4.2. Mapping strategies

Mapping strategies involved taking input(s) from the user's gestures as a starting point and processing and attaching this to some sonic output. Two main mapping strategies were used to design the user gesture-to-sound couplings within the tools in the kit. Simple mapping strategies are exemplified as a one-to-one gesture to sound (such as pressing a button to trigger a sample or note) and complex mapping strategies are exemplified as many-to-one. Complex many-to-one mappings can be considered by analogy of the violin. To bow a note on the violin, the player must control both the pressure and movement of the bow on the string as well as place a finger on the fret to change the note as needed, hence many inputs are needed to achieve one output. The different tools in the toolkit used different mapping strategies to allow for the continuum between instant access with strong cause and effect—through the potential for virtuosic and heavily nuanced control.

4.3. Features and accessibility of the toolkit

This section outlines the functionality and accessibility of each element of the toolkit as well as the software developed. Some design decisions were aimed at creating a more user-friendly device—such as all tools in the toolkit using commonly found batteries (9v and AA) housed in easy-to-open compartments. All the tools feature mounting fixtures and can be affixed to clamps or stands to enable users to control them without having to hold them—and to achieve the best positioning for the user. The designs produced were considered to reflect a combination of the author's and stakeholder's decisions about effective ways to address the opportunities and challenges presented by these situations. By reflecting on the outcomes of this process, various topical, procedural, pragmatic, and conceptual insights had the potential to be identified and articulated (Gaver, 2012). The products of this

TABLE 2 filterBox features and design requirements.

filterBox features	Design requirements identified through literature search and activities with stakeholders
• Hinged lid • On-board force sensitive resistor (FSR) • 2 buttons • Light dependent resistor (LDR) • Built in 9V battery compartment • Separate USB receiver (wireless connection) • Accompanying software	• Tangible hand-held tool • Enabling interaction styles akin to those used with traditional instruments • Potential to explore sound using fine-motor control • Direct translation of gesture into sound (i.e., squeeze harder to make louder) • Ability to move up and down common scales • High-fidelity and motivating sounds

process were the tools developed as artifacts through various stages of prototyping.

4.4. filterBox

The filterBox (Figure 2 top left) is an acrylic lacquered high-gloss finish on laminated softwood with acrylic facia plates and integrated high-quality sensors. Table 2 shows the features and design requirements of the filterBox. The use of these materials and finish suggests a quality and robustness invoking a familiarity analogous to acoustic instruments. The final look of the tool aimed to give the look, smell, feel, and overall sense akin to traditional acoustic instruments.

4.4.1. Features and accessibility

The filterBox was created to be hand-held in either hand. The idea was to offer the chance to practice fine motor control, manual dexterity, and coordination with complex mappings being used. Two buttons on the hardware control the notes, which can be moved up with one button and down with the other, through the selected scale. A force-sensitive resistor (FSR) situated around the middle of the tool detects pressure and controls the volume of the notes (off when no pressure is applied), and the light-dependent resistor (LDR) detects light and controls a filter on the notes. The buttons are positioned in a way that allows for accessibility of pressing but mitigates some elements of accidental presses by only being activated along one axis—i.e., a user must push down on the center and cannot push into the side to activate a sound. This enables the user to move around in a guided access way to find the direct spot for activation with space between the two buttons so that when one is being activated the other is less likely to be accidentally pressed. The proximity of them enables one finger/body part to be used to move between the two. The lid can be opened either by using a body part to lift it or by tipping the device upside down. Provided a user could grip the device and move it in the air, then the lid could be opened to trigger the LDR, but similarly fine motor skills could be practiced in opening and closing the lid. The filter-box could be clamped around the center to facilitate pressing the buttons alone to trigger notes and moving the lid to modulate the sound. The LDR activated by the lid could also be activated by covering it with the body or other materials, enabling users to utilize themselves or objects they may already interact with to control the sound.

As the complex mapping included control data from the FSR, LDR, and the buttons all working simultaneously, there is the opportunity for different elements to be controlled by different users as a multiplayer instrument. The smooth action hinge in the style of those used on a piano key cover can be opened and closed to increase/decrease the light to the LDR altering a filter on the sound. Using a commonly used interaction (the opening of a hinge), the lid mechanism aimed to provide a resistive mechanism for the user to work against to explore the sonic properties of the filterBox. The changing filter cutoff provided a familiar connection between the energy input via movement to match the sonic output. Expectation was used when designing some mappings—such as the idea of what a sound being put in a box would sound like to design the filter as the lid was closed. The hinged lid mechanism provided interaction opportunities that were analogous to those found in both traditional instruments—such as playing a trumpet with a mute—or in more contemporary music practices—such as scratching like a DJ. The position of the lid can be used to interpret how much light is being let in both by sight and by feel. There was an element of harnessing interactions that would be encountered in the everyday of the participants—such as opening a box and peeking in.

Many-to-one complex mappings are employed as a strategy to access the tool. Instant access and initial ease of use were balanced with the chance to achieve a nuanced and sophisticated control of sound and more technical exploration over time. The sounds were constrained to selectable scales to facilitate in-tune playing with other tools in the kit or other players. Complex physical manipulation could be navigated as the potential of the tool was explored by the user—for example, pressing the button to trigger an ascending note whilst undulated force on the FSR and simultaneously opening and closing the lid heralded different results to rapidly pressing the buttons whilst holding pressure on the FSR. The mobility of the tool allowed it to be used within the existing ecosystem of the user via their lap if in a wheelchair or on a table in front of them, and to be easily passed around without wires getting in the way or becoming a hazard or a point of distraction.

4.4.2. Software

Some software functions feature in the GUI of all three computer-connected tools. These functions are as follows:

- Sound off/on button.
- Volume control.
- Ability to switch output from stereo to left or right speakers (to enable two tools to be connected to one computer and both speakers to be used as independent sound sources).
- Access to the computers' audio system (to be able to select the audio output/input device).

The software GUI (Figure 3) can be used to change the VST to find a compelling sound for the user and the scale (and

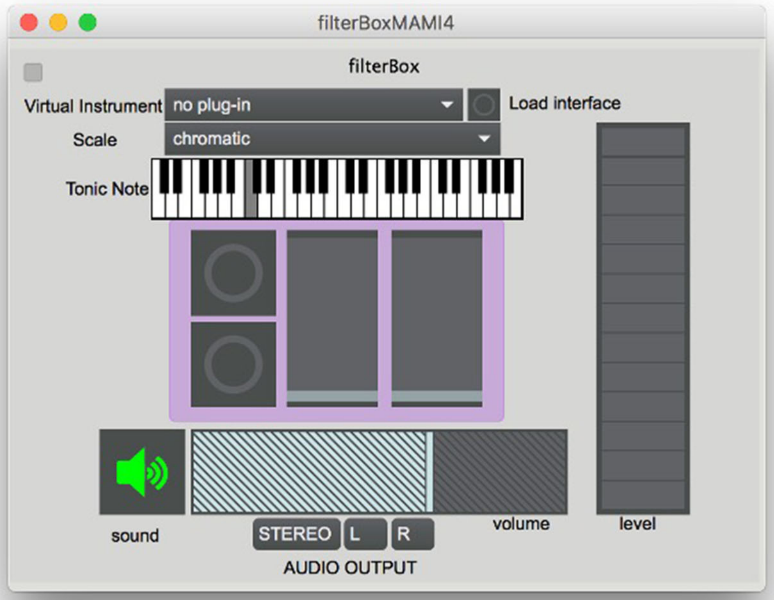

FIGURE 3
filterBox graphical user interface.

tonic note) can be selected to constrain the sounds. The filterBox software can be controlled via the Mira app on an iPad to foster accessibility by allowing settings to be adjusted to suit the user and by giving the option to interact with the iPad screen to change settings (and/or with the computer mouse)—depending on the needs of the users. The set-up of the software forms a series of constraints that the user must navigate to explore the sound gamut available (Magnusson, 2010; Wright and Dooley, 2019) providing opportunities to select a range of VST instruments and access each GUI—opening opportunities for endless augmentation of the sound to suit the user. This mechanism allows the user to delve into augmenting the sound whilst providing easily selectable presets. Several VST instruments were selectable (and settings within them accessible) to give the user access to a choice of high-fidelity sounds that were motivational to use and highly customizable. By revealing hierarchical levels of control of the settings as needed/wanted by the user, there was an attempt to enable and support the users without overwhelming them. There are visual elements on the GUI that show the real-time interaction with the buttons, force-sensitive resistor strip, and light sensor, as well as the levels of the sound and the master volume level. The GUI can be used with an interactive whiteboard allowing users to touch the screen to adjust levels.

4.5. squishyDrum

The squishyDrum (Figure 2 top right) is a round tool of 150 mm diameter which can be held in the hands or placed on the lap or surface and uses the same materials as the filterBox. Table 3 shows the features and design requirements of the squishyDrum. On the top, there is a 3-mm thick silicon skin. The materials used evoke a sense of robustness and similarity to an acoustic drum.

TABLE 3 squishyDrum features and design requirements.

squishyDrum features	Design requirements identified through literature search and activities with stakeholders
• 3 force sensitive resistors (FSR) • 2 piezo discs • Built-in 9V battery compartment • Separate USB receiver • Accompanying software	• Wireless tool that could be hand-held or placed on a lap or surface • Move away from fine motor control • Malleable surface to which bodily pressure could be applied • Ability to hit like a drum • Ability to record own samples and play back

4.5.1. Features and accessibility

The squishyDrum has a wireless connection so it can be moved around easily and placed in a position suitable to the user. The main interaction mode for the squishyDrum is applying pressure on the silicon surface under which there are three FSRs. The sensors under the surface provide accessibility—with very small amounts of pressure effectively magnifying the user's gestures into amplified sonic outputs. Two piezo discs send out a signal when tapped to allow for tapping on the drum or hitting with a stick. The squishyDrum could be leaned on to trigger sound

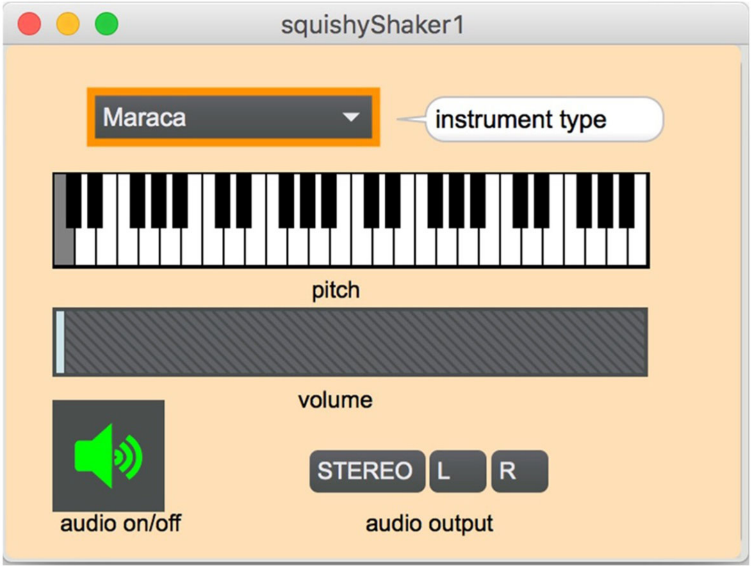

FIGURE 4
squishyShaker graphical user interface.

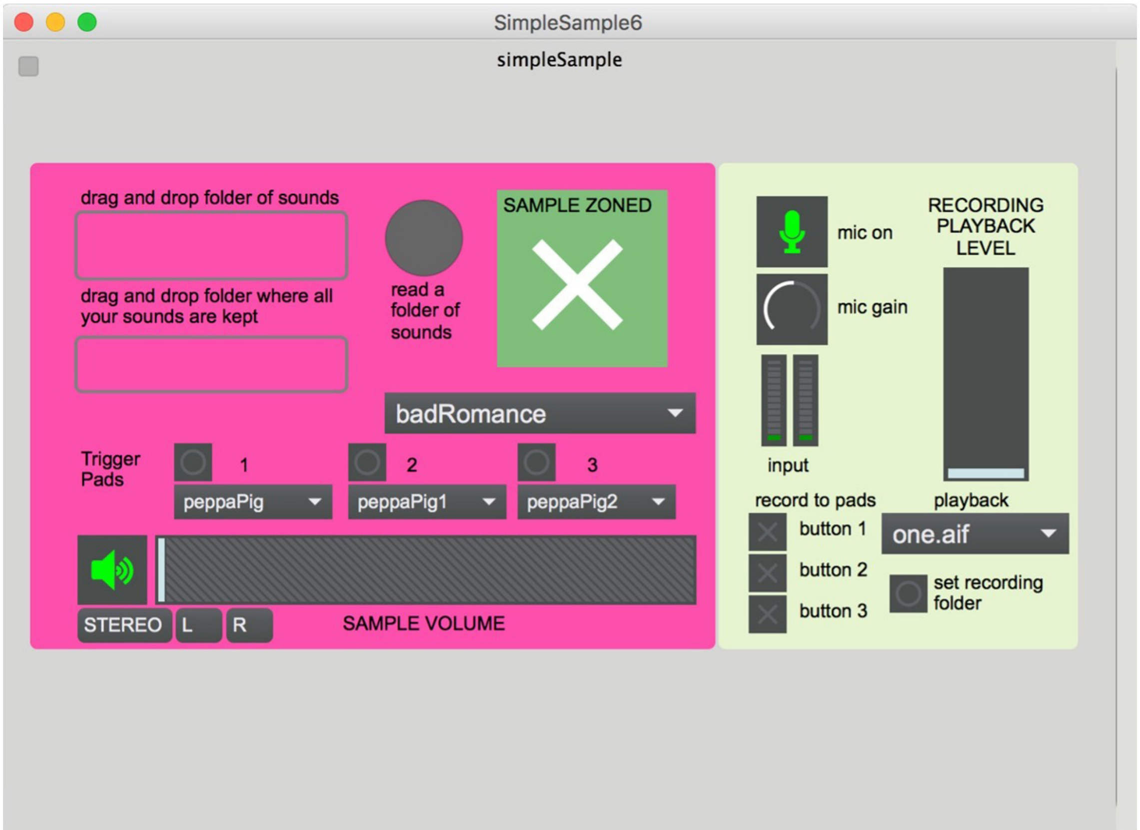

FIGURE 5
Simple sample graphical user interface.

or pushed on to any body part or object. After many requests by stakeholders, a rudimentary microphone recording capability was added to this tool. The recording was achieved by using the squishyDrum software GUI and the microphone connected to the computer with playback being triggered via the FSRs on the tool. The ability to record and add folders of sounds enables users (and those that know them best) to have ownership and tailorability of the output. Stickers were added to denote where the FSRs were located so that users would know where to put pressure.

4.5.2. Software

There are two software elements for the squishyDrum. The first, *squishyShaker* (Figure 4), allows for physical modeling synthesis to be triggered and modulated. Pressing the FSRs will determine the amplitude of the sound and the number of "objects" (coins, screws, and shells) in the digital shaker. The amount of effort exerted on the surface directly correlates to the intensity of the sound with the added element of having to press two of the force-sensitive resistors in tandem to trigger the sound. This was designed to encourage exploration of the surface to discover and coax out the sound. The user can hold the sound for as long as the surface is depressed, something that is almost exclusive of digital musical instruments (other than drone instruments). This self-sustain has the potential to enable the user to engage on a deep level by providing time to process the sound and to make the decision to stop. Also selectable is the pitch of the sound. The second software, named *Simple Sample* (Figure 5), allows for three samples to be triggered by pressing each of the three FSRs, or by pressing the on-screen button via the computer or the iPad (which can be used to remotely control the software). The ability to trigger via the iPad and the FSRs was to allow two users the opportunity to play together. One user could mirror the action of another user and vice-versa—in a call-and-response style interaction. The ability to record from a microphone was added to respond to stakeholder requests to magnify the users' voices, enabling them to hear themselves, as well as giving some ownership, involvement, and autonomy in the creation of triggered content. The gain of the recording microphone can be changed, as well as the playback volume of the recorded and in-built samples. The user can trigger both in-built samples and their stored recordings at the same time and mix the levels of these independently. Sample folders can be dragged and dropped (using either a single folder or a folder of folders—and folder hierarchy will be maintained). Folders can also be loaded by clicking to open a pop-up dialog box on the computer. A setting was added to the software GUI to allow switching between a triggered sample playing to the end or retriggering upon repressing the pressure pad—to allow for sonic layering to be achieved.

4.6. The Noodler

The Noodler (Figure 2 bottom left) is a wireless tool based around the Nintendo Wii chuck controller which has a built-in accelerometer to detect movement across the *x*, *y*, and *z* axes, an *x*/*y* axis joystick, and two buttons. The Wii chuck has an ergonomically designed form factor (albeit aimed at a user with typically developed hands), with smooth rounded edges and high production-level quality. Table 4 shows the features and design requirements of the Noodler.

TABLE 4 The Noodler features and design requirements.

The Noodler features	Design requirements identified through literature search and activities with stakeholders
• Removable Wiichuck controller • On-board accelerometer • x/y joystick • 2 buttons • Built-in 9V battery compartment • Separate USB receiver • Accompanying software	• Tangible hand-held wireless multi-modal device • Customisable ability to trigger sounds "Drawable" trigger zone templates to allow individual user mappings of gesture to sound • Ability to add user media to create motivating interaction • Provision of commonly used presets in the form of a variety of instruments/scales/sound effects to allow instant access • Use of familiar input devices (joystick and buttons)

4.6.1. Features and accessibility

For this research project, the Wii chuck was made detachable from the transmitter unit to be able to allowing users to keep their own Wii chuck or replace it should it cease functioning. The compact form factor of the Noodler meant that it was lightweight to move around and was accessible to a person with muscle weakness or weakened stamina, with the heavier part of the tool (transmitter containing batteries) on the end of a 50-cm cable which could be stowed either in a belt or on a chair nearby. The Noodler is a recognizable tool both in being a joystick and a controller for the popular computer console the Wii, which builds on commonly used interaction (that of controlling things with a joystick or pushing a button). The Noodler provides access to triggering notes and samples through its joystick movement and buttons. The joystick and buttons could be accessed with the thumb and fingers but also by holding the Noodler and pushing either input onto a surface to trigger sounds. This enabled it to be used against different body parts/against tables to activate the sonic output. The Noodler was able to be physically manipulated and moved into comfortable positions by the users to gain access to the modes of interaction available (one participant used their lip to control the joystick). The ability to change both sonic content and triggering gestures meant the tool could be tailored to the individual. By allowing the user to select from samples provided or the ability to add their own, they could appropriate the system to suit their tastes to motivate engaged use.

4.6.2. Software

There are two modes that the Noodler software (Figure 6) can facilitate: MIDI note triggering or sample triggering. When MIDI note triggering is selected, a MIDI instrument can be selected (from 128 choices) as the octave (−2 octaves to +4 octaves) or MIDI drums. When sample triggering is selected, a range of preset samples can be triggered featuring commonly requested sound effects or songs, or the user could upload their own samples. The user folders can be loaded via a drag-and-drop mechanism or pop-up dialog box. Samples triggered by the buttons can also be selected. One of the key accessibility features of The Noodler is

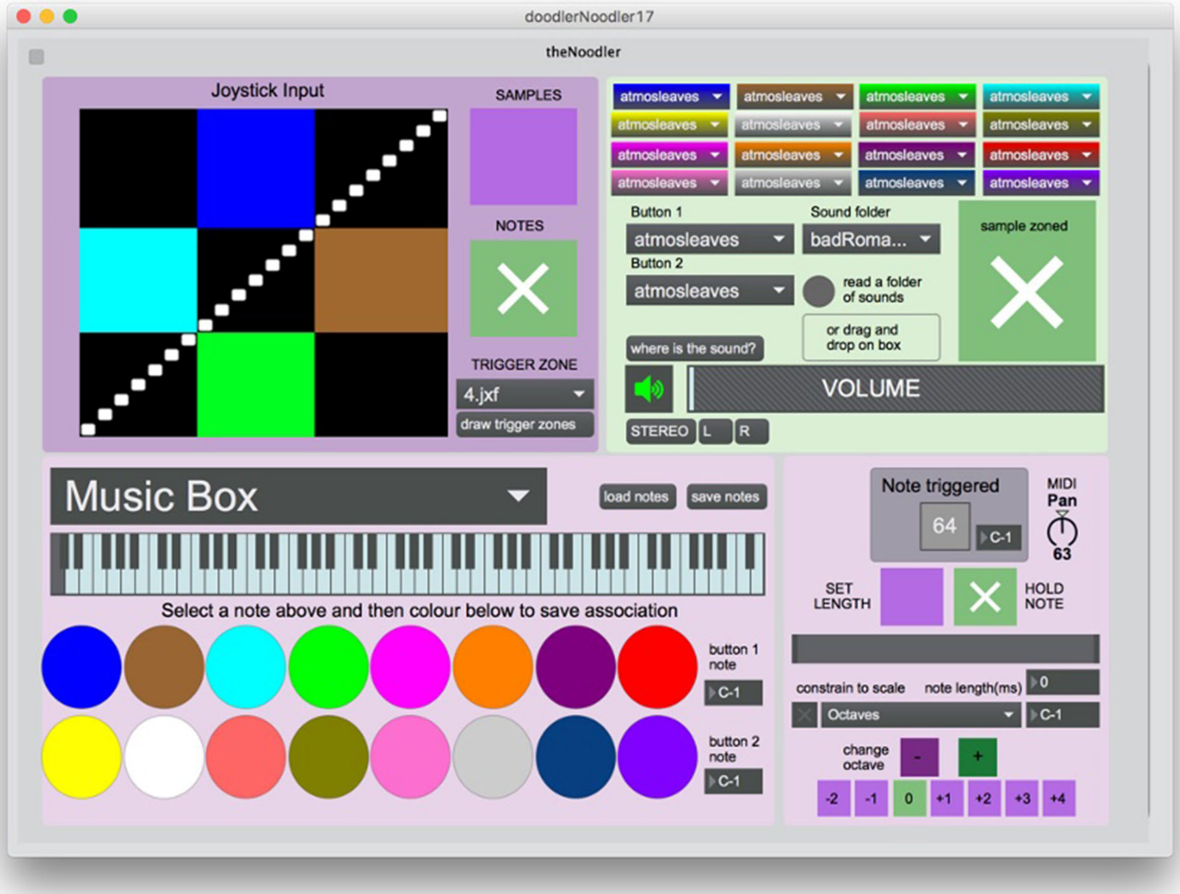

FIGURE 6
Noodler graphical user interface.

the ability to draw triggerable zones on the GUI that are activated by the motion of the Noodler joystick. The familiar mechanism of drawing (with the mouse) can be used to create these zones within the GUI. The triggerable area (shown below the title "joystick input" in Figure 6) is virtually mapped to the gestural dimension space that the joystick can move around. Different color pens (up to 16 colors) can be used to color areas of the triggerable zone—the patterns of which can be saved and reloaded. Each color can then be mapped to trigger a MIDI note or sample with areas left black as rest zones. This allows the joystick to effectively become scalable in sensitivity facilitating use for both fine and gross motor control depending on the trigger pattern drawn (Figure 7 shows four presets that come pre-loaded). These customisable trigger zone templates allow individual user mappings of gesture to sound as the user moves the joystick in a way comfortable to them, and the trigger zones can then be drawn to suit. A dot moving around the square in the GUI is provided as an on-screen visual representation of the position of the Noodler within the trigger zone—to give users some visual feedback on the effect of their actions and to establish cause and effect by ensuring explicit visual mapping between sites of interaction and sonic generation.

Several presets were included within the software—featuring varying amounts of colors and zones. All the software loads up with mappings enabled so that sound can be interacted with straight away. This tailors to a specific request from stakeholders to include presets and set-ups that open with instant sonic output upon loading the software. When MIDI note triggering is selected, there is the ability to select whether the note stays on whilst the target dot stays within a trigger zone or whether the triggered note has a set duration—which can also be modified. There is a pan dial to allow the MIDI output to be panned to the left or right.

The Noodler sacrificed complex mappings (although it has the data streams to be able to design this in the future) to become a sample/note-triggering tool—a request that came up through working with stakeholders. These simple mappings can be layered up to create a complex system such as the Simple Sample app in which the triggered sample can be cut off when moving out of the trigger zone or continue to play out depending on what the user selects within the software.

Within the software, the state of the system can be seen from a variety of GUI components with musically analogous elements (e.g., keyboard slider and faders in Figure 6)—to provide a system that made sense to the user. More time to iterate over the look of the software would have been helpful in creating an interface that better matched the user's needs. The current interface may alienate some users—by being difficult to interpret (describing the functionality in a simple manner without using jargon whilst retaining an accurate description of said functionality is a challenge) and by having inaccessible usability qualities (icons that are too small, no text-to-speech, or use of icons to assist visually impaired users or those cannot read), both of which could form barriers to some users. The addition of iPad integration via the Mira app was used to help alleviate this in some areas.

4.7. touchBox

The touchBox (Figure 2 bottom right) is a stand-alone box with eight jack sockets for detachable capacitive touch 50 mm × 50 mm copper pads (each pad with 3 mm jack tipped 1 m retractable cables), two dials, five buttons, built-in LCD, internal speaker, headphone and ¼ inch jack socket for output, and toggle switch for toggling between internal speaker and ¼ in output. Table 5 shows the features and design requirements of the touchBox.

4.7.1. Features and accessibility

This self-contained unit had the aim of being able to turn on itself and play without involving an external computer and to open accessibility to the controls of the device to users by using tangible controls. These controls provide a means to grasp against. Materially, it matches the design aesthetic of the filterBox, squishyDrum, and Noodler. The pads are of hand-held size and can be held, placed on a surface, or mounted, meaning they can be positioned to suit the user. Up to eight pads can be used at one time. The pads require a light touch on a copper conductive plate to trigger and stay activated until the touch is removed. This provides accessibility for those who can only apply very small amounts of pressure (or even just place their body part down on the plate) and gives control beyond that of triggering sound to choosing when the sound stops. This gives the users a chance to rest and take in the sound, giving sometimes vital processing time needed to truly realize cause and effect. The movability of the pads gives users some autonomy in the set-up of their own instrument—as appropriation is common in other musical instruments where each player has their own unique set-up. The touchBox is polyphonic meaning that more than one note can be played concurrently.

The main unit has buttons to control: the timbre of the sound, the octave of the notes (one button to move up and one button to move down), the scale, and the tonic note of the scale (cycling around the setting as pushed). Two dials on the main unit control, the volume and note, decay (allowing for short staccato notes or long legato notes). Each button and knob on the main unit have a different style of casing for the different controls that are offered and are also in different colors. These design decisions provide the ability for the user to develop a relationship with the tool by touch alone and for visually impaired users to be able to distinguish between controls. The retractable cords enable ease of putting away the tool, and whilst this may seem like an innocuous feature, it can be argued that these features add to the overall usage experience of the tool. The way the instrument is stored and retrieved, connected, and set-up contributes to this practical use. This begins with the decision to use the tool and ends when it is returned to storage as one of the biggest stakeholder issues was getting the technology set-up and having all the accouterments to achieve successful use.

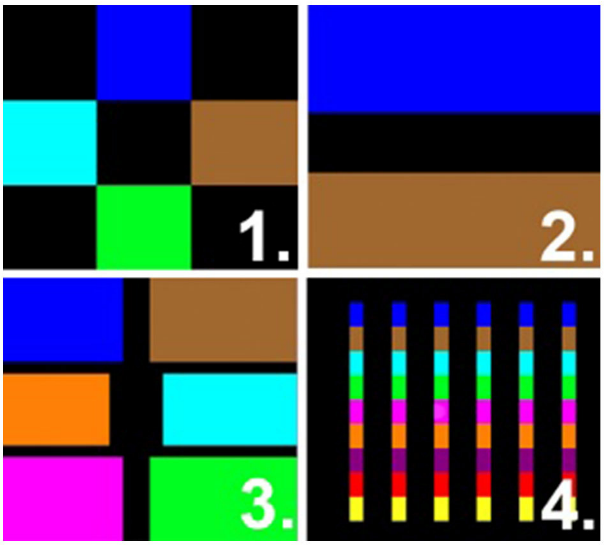

FIGURE 7
Noodler trigger pattern presets.

TABLE 5 touchBox features and design requirements.

touchBox features	Design requirements identified through literature search and activities with stakeholders
• Stand-alone box with 8 jack sockets • 8 capacitive touch copper pads (each with 3 mm jack tipped 1 m retractable cables) • 2 dials • 5 buttons • Built-in LCD display • Internal speaker, headphone and ¼ in jack socket for output, toggle switch (between internal speaker and ¼ in output) • 2xAA battery compartment	• Self-contained unit with on-board speaker • Turn on and play • Light touch to activate • Polyphonic • Headphone socket • Screen display • Operate by touch alone • triGger notes with selectable waveform, scale, tonic note, octave select, volume, note decay length

4.8. The MAMI tech toolkit features and accessibility

Table 6 shows the features and design requirements of the whole kit. The tangibility of the tools "takes advantage of embracing the richness of human senses developed through a lifetime of interaction with the physical world" (Ishii and Ullmer, 1997, p. 7) to provide rich multisensory experiences and an interface to grasp against. The final tool construction infers quality (see Figure 8). The tangibility of the tools also provides a mechanism for the users to experience their bodies. An analogous concept might be to think of

TABLE 6 MAMI tech toolkit features and design requirements.

MAMI Tech Toolkit features	Design requirements identified through literature search and activities with stakeholders
• filterBox • squishyDrum • The Noodler • touchBox • USB cables and receivers • Aluminum flight case • Software on USB stick • Instruction manual • Quick start guide • Laminated kit contents tick sheet	• Easy-to-set up • Wireless • Used by a variety of users • Move toward alleviating a fear of use • Separation of controls from interface • Ability to control settings whilst away from the computer • Cohesive kit with tools that could be used together or individually • Focus on quality of materiality that can sit alongside traditional musical instruments such as the acoustic guitar • Ability to attach to stands/clamps/arms • Presets that featured commonly used scales/notes/instruments • Following an open-source philosophy • Use easy-to-access and affordable components • Move away from screen-based interaction toward tangible user interfaces • Kit that can stay within research sites after the research is over • Kit that is sensitive to typical practice-based use in context of the research sites • Visual feedback on the system state such as indicator lights

weighted blankets that are used to provide the sensation of being embraced to alleviate anxiety or stress. The use of the weighted blanket can be seen to provide an edge and a stopping point against which a person can delineate their own edges in a proprioceptive manner. In the same way, the tools in the toolkit provide a means for the user to experience both their gestures as co-constituted with the tools, providing an opportunity to explore their own body. By extension of this mechanism, the sound can also provide an "edge" against which to interact and explore the sound/body relationship.

The ability to split the different tools in the toolkit up (each tool had its own receiver) facilitates the tools being able to be taken home by the user and practiced with for the chance to develop a relationship with the tool. The tools can be used alongside each other with a similar range of presets. Presets that are featured within the toolkits are commonly used scales/notes/instruments to allow for ease of integration with acoustic instruments and to allow for use with the existing canon of repertoire that was commonly used by the stakeholders. These can be selected and made to be cohesive with each other or to fill different ranges of frequencies, timbres, feelings, and movements, much like an orchestra or an ensemble would be constructed to work together filling the sonic space.

The three tools that connect to the computer can be controlled either via a computer or via the iPad. The Mira app allowed mirroring of the GUI from the Max/MSP software on the computer to the iPad, which in turn means the iPad can be used as a controller for the settings. The Mira app (£9.99 at the time of writing) manages the connection between the iPad and Max/MSP software allowing settings to be controlled away from the computer, thus removing the need to touch the computer during the interaction. This can be useful for some users who might find the computer a distraction and removes the need for the users to physically sit at a computer which could become a barrier to interaction. An iPad is also a much more familiar and enticing control unit with direct access to enable quick changes of the set-up or modification of the controls—some elements of the sound can be triggered by the iPad also to allow the iPad user and the tool user the potential to interact with each other.

The state of the system can be understood via visual elements that represent its states—for example, LEDs were used in the receivers to indicate that the units were active. The toolkits are contained in a metal flight case with all the components needed (minus the computer). The choice of a sturdy metal box is both analogous to transporting important artifacts and provides a practical storage solution for tools that are robust. The aim was to consider how the toolkit would fit into practice and have an overall sense of cohesion, as well as a feeling "ownable" by being portable. The hard flight case also considers the ritual of use that runs from deciding to use, using and placing it back in storage. Moreover, in a school setting, any help to mitigate parts going missing is usually welcome (hence the addition of the laminated list of contents with pictures to help users locate and replace items).

The toolkit software was also provided on a USB stick (as well as being downloadable via GitHub) to ease distribution and use in practice. Whilst these details may not involve the direct use of the tools in active music-making, they mediate the use of the tool. By providing tools that holistically consider their whole context of use, tools may integrate more easily into the context within which they are used. They have an authenticity that is considerate of the practice that they are part of. In this way, the tools enmesh with the practice within which they sit. The toolkit was created with an open-source philosophy and as such the internal diagrams of components, bill of materials, construction/code, and Max/MSP patches (the code created in Max/MSP) were made freely available on the GitHub link (see the Data availability statement section below) with a focus on using easy-to-access and affordable components throughout.

5. Discussion

5.1. Material qualities and cause and effect

Both the construction of the tool in terms of the aesthetical look and feel and the type of sonic content used (high-end synthesis or high-quality samples) received positive feedback from the stakeholders. The fidelity of the sounds offered using synthesis or VST instruments and the expressive potential that is built into their programming were more successful in engaging the stakeholders than when the tools triggered standard MIDI-based instruments/sounds. Settings and options were given in a particular order to scaffold the practitioners in their set-up and use of the tools. This could have been further explored in terms of hierarchical systems of access to settings as these could have been tailored more specifically to users depending on their confidence in using technology. By delivering the ability to change settings in a phased way, overwhelm can potentially be minimized.

FIGURE 8
MAMI tech toolkit complete in flight case.

Technology can be used to leverage interactions already associated with acoustic instruments whilst providing the ability to design interactions from the ground up to suit users. The breaking down and reconstructing of instruments in this way has a danger of confusing the user if cause and effect are not made meaningful to them and can trigger "learned helplessness" (Koegel and Mentis, 1985), when the user loses motivation and no longer believes that they can achieve the outcome or have any control if an interaction does not match their expectation or is not fully made accessible to them. The problems of this decoupling can be mitigated by clear signposting of how a system works including demonstrating the features and functionality to allow users to become comfortable; obeying commonly used interaction mechanisms (pressing something harder makes it louder); and creating robust systems with few technical issues—or at least mechanisms in place to rapidly find solutions or workarounds to problems to help users get things working again. The issue of this successful reconstitution of sound to gesture is one that has been discussed within the general realm of creating new interfaces for musical expression (Calegario, 2018). Questions arose within the research project giving appropriate feedback to ensure the user knew what they were doing and how that changed the sound, and ensuring cause and effect was meaningfully achieved. A successful method was to allow users to get to know their sound before joining group sessions—with the sound-producing mechanism (speaker or amplifier) positioned as closely as possible to the user. If possible, even placing the amplifier in contact with the user's chair to provide haptic feedback as well as auditory.

In the school setting, other people present such as carers and teachers/teaching assistants were useful in supporting individuals. They could facilitate the use of the tool, forge connections between the tool and the sound, and reduce ambiguity between the gesture and the sound. This ambiguity can be further widened by actions such as constantly changing how the tool works, changing the scale or instrument, or changing sounds. Sticking with a particular set-up could be beneficial to gaining mastery over an instrument but equally some users want to hear all the options they have before settling on a particular favorite. Again, a balance must be struck.

5.2. Mappings and design constraints

The mappings used throughout this research were constrained in terms of constraining notes to particular scales as a tool to aid in achieving inclusivity by giving the users the chance to play without feeling worried that they were going to play something wrong (except for the Noodler in which users could select any note they wanted to play). The use of these mechanisms provides suitable, appropriate, and—even more key—acceptable scaffolding for user interactions that strike that balance of difficulty and simplicity. The two ends of the continuum must be balanced on a case-by-case basis with one end leading to overwhelm and the other to boredom.

The appropriate use of technology should aim to support access or a cognitive need of the user and fit their desires. This is a bonus of ADMIs in that they can be made to scale and tailor to individual users—enabling (by the change of a dial) a range of notes to be triggered by moving millimeters or meters.

There are many goals that the tools in the toolkit could support from wellbeing to social inclusion to occupational therapy or more obviously musical expression. It may be that virtuosity is seen as the epitome of musical goals but what is suggested here is that we consider virtuosity as a relationship between typical and extraordinary *for a particular individual*. Pressing a switch to trigger a sound might demonstrate a level of virtuosity for some. If the goal is to play something in time (whether to a particular beat per minute, to a desired rhythm, at a desired point by the user, or to create a desired effect such as layering sounds, for example), then to be able to push a button in time (for reasons as suggested above) is considered a successful movement toward this goal. If the goal is to decide when to respond and move toward an intentional response, then the individual pressing the switch whenever the user wants would be a successful movement toward this goal. If the desire is to allow fine control over the pitch of a note, then pressing a button to trigger a preset sound does not facilitate that type of outcome. This should be considered when setting users up with technology to support them.

The toolkit can facilitate both the triggering of single events and continuous control of sound—thereby providing a level of expression that matches the users' needs. Within the toolkit, there was a balance between providing a flexible, understandable, easy-to-use, and customisable system without it becoming overwhelming. Initially, the idea of the toolkit was to be completely modular—in that different sensors could be selected and then attached to different musical outputs—in a plug-and-play manner. However, it was quickly evident that this kind of development would be beyond the scope of the research resources. There was a dichotomy between bespoke tailoring to one user or modular flexibility that may be "good enough" for many users—as such the final application was tailored to be easy-to-use and featured functionality for use within a range of typical scenarios—and use cases—that stakeholders requested or were observed during practice. The stakeholders desired that tools be compatible both with one another and current technologies and work alongside traditional instruments. This would be a future goal—even the ability to use generic assistive technology style switches with the toolkit would be beneficial.

5.3. The tools in use

The Noodler was the only tool that had extensive testing and was used successfully in several group sessions alongside traditional instruments, and with settings that matched the repertoire being covered within these "in the wild" sessions. These songs were always discussed in advance of sessions with the music therapist, but the Noodler settings could be changed throughout the session to match the song that was being used at that time. The Noodler was attached to its own amplifier and positioned near the user—who became more accustomed to its use over time. A particularly successful example of this was when a well-known song (one of the participants' favorites) was cut into short (3–5 s samples) that could be triggered by the Noodler. The user and the music therapist engaged in some call and response interplay with the music therapist playing the same section of the song on an acoustic guitar—and the participant using the Noodler buttons to respond. It was clear to those present that the value of these forms of technology was made concrete in witnessing that interplay. The participant was clearly empowered by using the technology, and the playing field between practitioner and recipient was seemingly leveled. Those present commented that it was the most engaged and laughter-filled time they had witnessed from the user ever. The music therapist also analyzed a video of a different participant using the Noodler for a performance in front of an audience and described how they thought the tool opened the user to their surroundings and encouraged real-time engagement with the music and the space.

5.4. Barriers and next steps

This research suggests that the barriers to technology use can be placed into four categories: barriers to finding appropriate technology, barriers to setting technology up, barriers to integrating technology into practice, and barriers to using technology within the session. The above categories also interlink depending on the goals and needs of the practitioners and the individuals using the technology. Each barrier could be considered to have its own skill set and different training needs to overcome and each points to potential gaps in provision and potential ways to break down these barriers by providing technology that addresses them, or supportive documentation to understand technology or examples of integrated technology in practice. It is hoped that future work with the toolkit will add to this discourse as this was limited in this research. It is the opinion of the author that what is needed in terms of the next level of technological advancements are easier-to-use tools that account for heterogeneous users; tools that focus on the context of use; development in the field of interaction with tools for users with the most profound needs (especially in areas of assessing their needs with regard to provision for music-making, and assessing interaction with tools for music-making); and more resources and examples of best practice for use of all of the above that synergistically combines with as many existing resources as possible—such as, for example, the Sounds of Intent framework (Vogiatzoglou et al., 2011).

5.5. Limitations

There were many limitations to the toolkit, and much further work could be conducted around it. The main limitations were that there could have been further exploration of the sound world that the elements of the toolkit connected to—the mappings could have been richer. The GUI of the software could have been further developed to be more accessible, and in general, the whole toolkit could benefit from extensive testing in context. The organic following of the needs of the stakeholders meant that the data were messy to pick apart as there were no driving research questions and only tentative aims with stakeholders guiding how the research should go. Sometimes data-informed discussion around contextual issues of using music technology (in terms of making sure individual needs were met by considering logistical matters such as sound levels) highlighted a technical issue (showed something that needed fixing) or informed future design (a feature to add). Data could be all three at once and would triangulate to inform the three perspectives. An example would be the piece of data that said "one child found the sound level too loud"—this meant the requirements in design needed to consider how to make the child comfortable, to create a method to control sound, and to add the ability to access that control quickly. The integrated sound (and haptic feedback) would have been great additional improvements and fall into the category of "future work." Adding LEDs to show the state of the system (for example, to show when the device was on, connected to the receiver, sending data, and even utility lights such as when the battery was low) could also have been further developed.

A limitation in the ethical consent process of the project meant that video of interactions featuring CYP playing with the MAMI Tech Toolkit could not be shared due to explicit consent not being obtained for video footage. However, the author would argue that it is of utmost importance to include the voices of participants alongside those of proxy representatives and practitioners, especially in the context of underrepresented voices.

6. Conclusion

In this study, a description of the development of the MAMI Tech Toolkit was given. This covered the cycles of action research used as part of the methodology to work with stakeholders and draw out tacit knowledge. The knowledge and who contributed was outlined, and how this was used to shape the direction and goals of the research and the design of the tools was discussed. Third-wave HCI methods were used to link the embedded exploration of people using technology "in the wild" and how this has shaped the technology developed throughout the action research methodology. The "technology prototype as probe" was outlined as the mechanism used to engage stakeholders with the design process as both a chance to see what was possible and as a point of departure into further ideas and design iterations. The hardware and software of the toolkit were described in terms of the engineering elements, features, and functionality. Each element of the toolkit was explored (and the toolkit as a whole) to analyse the accessibility and logistical design decisions that were made and links were highlighted to the stakeholder data that contributed to this. A discussion was provided which reflected on the integration of such tools into practice alongside some of the themes that emerged from the embedded iterative development of the system. Finally, some key factors of discourse that surround these types of tools such as material qualities, cause and effect, mapping, design constraints, barriers, and next steps were discussed.

In summary, this study described the use of an action research methodology to develop an accessible music technology toolkit. There was an emphasis on using participatory and iterative methods to both elicit current issues and barriers with music technology and to develop novel tools to address these gaps in provision. The contributions of the study advance knowledge around active music-making using music technology, as well as in working with diverse users to create these new types of systems.

Data availability statement

The datasets analyzed for this study can be found in the BORDaR at Bournemouth University http://bordar.bournemouth.ac.uk/196/. GitHub repository for the code created can be found at: https://github.com/asha-blue/MAMI-Tech-Toolkit-Final-Edition.

Ethics statement

The studies involving humans were approved by Bournemouth University Ethics Panel. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

The funding was granted by the EPSRC (code: EP/L016540/1).

Acknowledgments

This research was carried out as part of the author's engineering doctorate candidacy within the Centre for Digital Entertainment at Bournemouth University under the supervision of Dr. Tom Davis and Dr. Ann Bevan. The author thanks Luke Woodbury for industrial mentorship throughout and the Three Ways School for industrial sponsorship.

Conflict of interest

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

Publisher's note

References

AlphaSphere (2023). *AlphaSphere*. Available online at: https://alphasphere.com/ (accessed April 6, 2023).

Arduino (2015). *Arduino Uno*. Italy: Arduino.

Bela (2023). *Bela Mini kit*. Available online at: https://shop.bela.io/ (accessed December 22, 2022).

Bevan, A. (2013). Creating communicative spaces in an action research study. *Nurs. Res.* 21, 14–17. doi: 10.7748/nr2013.11.21.2.14.e347

Bødker, S. (2006). "When second wave HCI meets third wave challenges," in *Proceedings of the 4th Nordic Conference on Human-computer Interaction: Changing Roles* (New York, NY: ACM), 1–8. doi: 10.1145/1182475.1182476

Bott, D. (2010). Towards a More Inclusive Music Curriculum - The Drake Music Curriculum Development Initiative, *Classroom Music Magazine.*

Calegario, F. (2018). *Designing Digital Musical Instruments Using Probatio: A Physical Prototyping Toolkit.* Berlin: Springer. doi: 10.1007/978-3-030-02892-3

CENMAC (2021). *Soundbeam: The Invisible Keyboard in Space*. Available online at: https://cenmac.com/2021/11/soundbeam-the-invisible-keyboard-in-space/ (accessed March 5, 2023).

Creswell, J. (2014). *Research Design*. Los Angeles, CA: SAGE Publications Ltd.

Cripps, C., Tsiris, G., and Spiro, N. (eds) (2016). *Research and Resources for Music Therapy 2016: A Compilation Developed by the Nordoff Robbins Research Team.* London: Nordoff Robbins. Available online at: https://www.nordoff-robbins.org.uk/sites/default/files/research_and_resources_for_music_therapy_2016_final.pdf (accessed November 2, 2017).

Cycling'74 (2023). *Cycling74.com*. Cycling'74. Available online at: https://cycling74.com/ (accessed June 21, 2023).

Dabin, M., Narushima, T., Beirne, S., Ritz, C., and Grady, K. (2016). 3D modelling and printing of microtonal flutes. *New Interfaces for Musical Expression*, July 11–15. (Brisbaine: NIME), 286–290. Available online at: https://www.nime.org/proceedings/2016/nime2016_paper0056.pdf (accessed June 2, 2023).

Davanzo, N., and Avanzini, F. (2020a). Hands-free accessible digital musical instruments: conceptual framework, challenges, and perspectives. *IEEE Access* 8, 163975–163995. doi: 10.1109/ACCESS.2020.3019978

Davanzo, N., and Avanzini, F. (2020b). "A dimension space for the evaluation of accessible digital musical instruments," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Birmingham: Birmingham City University). NIME, 214–220. doi: 10.5220/0009816106200628

Deluca, D., Gallivan, M. J., and Kock, N. (2008). Furthering information systems action research : a post-positivist synthesis of four dialectics. *J. Assoc. Inf. Syst.* 9, 48–72. doi: 10.17705/1jais.00150

Department for Education (2011). *The Importance of Music: A National Plan for Music Education*. gov.uk, Department for Education. Available online at: https://www.gov.uk/government/publications/the-importance-of-music-a-national-plan-for-music-education (accessed March 4, 2017).

Drake Music (2016). *The Kellycaster*. Available online at: https://www.drakemusic.org/blog/becky-morris-knight/the-kellycaster-accessible-guitar/ (accessed March 5, 2023).

Drake Music (2023). *Homepage*. Available online at: https://www.drakemusic.org/ (accessed April 6, 2023).

Farrimond, B., Gillard, D., Bott, D., and Lonie, D. (2011). *Engagement with Technology in Special Educational and Disabled Music Settings*. Youth Music.

Förster, A. (2023). Accessible digital musical instruments in special educational needs schools - design considerations based on 16 qualitative interviews with music teachers. *Int. J. Hum.-Comput. Interact.* 39, 863–873. doi: 10.1080/10447318.2022.2049141

Förster, A., Komesker, C., and Schnell, N. (2020). "SnoeSky and sonicdive - design and evaluation of two accessible digital musical instruments for a SEN school," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. Birmingham, July 21–25. New Interfaces for Musical Expression (NIME), 83–88.

Frid, E. (2018). "Accessible digital musical instruments - a survey of inclusive instruments," in In *Proceedings of the International Computer Music Conference.* International Computer Music Association (Daegu: ICMC), 53–59.

Frid, E. (2019). Accessible digital musical instruments - a review of musical interfaces in inclusive music practice. *Multimodal Technol. Interact.* 3, 57. doi: 10.3390/mti3030057

Frid, E., and Ilsar, A. (2021). "Reimagining (accessible) digital musical instruments: a survey on electronic music-making tools," in *NIME 2021* (Shanghai: PubPub). doi: 10.21428/92fbeb44.c37a2370

Gaver, W. (2012). "What should we expect from research through design?" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12: CHI Conference on Human Factors in Computing Systems* (Austin, TX: ACM), 937–946. doi: 10.1145/2207676.2208538

Glinsky, A. (2005). *Theremin: Ether Music and Espionage,* 2nd ed. Champaign, IL: University of Illinois Press.

Graham-Knight, K., and Tzanetakis, G. (2015). "Adaptive music technology: history and future perspectives," in *International Computer Music Conference*. Denton, TX, USA, 25 September—1 October (Ann Arbor, MI: Michigan Publishing). doi: 10.1145/2769493.2769583

Gurevich, M., Stapleton, P., and Marquez-Borbon, A. (2010). "Style and constraint in electronic musical instruments," in *Proceedings of the Conference on New Interfaces for Musical Expression*, Sydney, Australia, June 15–18th. New Interfaces for Musical Expression (NIME).

Harrison, J., Chamberlain, A., and McPherson, A. P. (2019). "Accessible instruments in the wild: engaging with a community of learning-disabled musicians," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19: CHI Conference on Human Factors in Computing Systems* (Glasgow: ACM), 1–6. doi: 10.1145/3290607.3313037

Harrison, S., Tatar, D., and Sengers, P. (2007). "The three paradigms of HCI," in *Alt. Chi. Session at the SIGCHI Conference on Human Factors in computing Systems* (San Jose, CA), 1–18.

Heart n Soul (2023). *Homepage*. Available online at: https://www.heartnsoul.co.uk/ (accessed April 6, 2023).

Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., et al. (2003). "Technology probes: inspiring design for and with families," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY: ACM), 17–24. doi: 10.1145/642611.642616

Ilsar, A., and Kenning, G. (2020). "Inclusive improvisation through sound and movement mapping: from DMI to ADMI," in *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility. ASSETS '20.* Virtual Event Greece (New York, NY: ACM), 1–8. doi: 10.1145/3373625.3416988

Ishii, H., and Ullmer, B. (1997). "Tangible bits: towards seamless interfaces between people, bits, and atoms," in *Proceedings of the ACM SIGCHI Conference on Human factors in Computing Systems* (New York, NY: ACM), 234–241. doi: 10.1145/258549.258715

Jack, R. H., Harrison, J., and McPherson, A. P. (2020). "Digital musical instruments as research products," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (New Interfaces for Musical Expression), 446–451. doi: 10.5281/zenodo.4813465

Jordà, S. (2004). "Digital instruments and players: part I - efficiency and apprenticeship," in *Proceedings of the 2004 Conference on New Interfaces for Musical Expression* (New Interfaces for Musical Expression), 59–63.

Kirwan, N. J., Overholt, D., and Erkut, C. (2015). Bean: a digital musical instrument for use in music therapy," in *The Sound and Music Computing Conference*, ed. J. Timoney (Maynooth: Maynooth University).

Knight, A. (2013). Uses of iPad® applications in music therapy. *Music Ther. Perspect.* 31, 189–196. doi: 10.1093/mtp/31.2.189

Koegel, R. L., and Mentis, M. (1985). Motivation in childhood autism: can they or won't they? *J. Child Psychol. Psychiatry* 26, 185–191. doi: 10.1111/j.1469-7610.1985.tb02259.x

Krout, R. (2014). "Engaging iPad applications with young people with autistic spectrum disorders," in *Music Technology in Therapeutic and Health Settings*, ed. W. Magee (London: Jessica Kingsley Publishers), 181–198.

Krout, R. E. (2015). *Evaluating Electronic Music Technology Resources for Music Therapy*. Dallas, TX: Barcelona Publishers.

Kuutti, K., and Bannon, L. J. (2014). "The turn to practice in HCI: towards a research agenda," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY: ACM), 3543–3552. doi: 10.1145/2556288.2557111

Larsen, J. V., Overholt, D., and Moeslund, T. B. (2016). "The prospects of musical instruments for people with physical disabilities," in *New Instruments for Musical Expression* (Brisbane), 327–331.

Leeuw, H. (2021). "Virtuoso mapping for the electrumpet, a hyperinstrument strategy," in *NIME 2021* (Shanghai: PubPub). doi: 10.21428/92fbeb44.a8e0cceb

Lofland, J., and Lofland, L. H. (1994). *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*. Belmont: Wadsworth Publishing Company.

Lucas, A., Ortiz, M., and Schroeder, F. (2019). "Bespoke design for inclusive music: the challenges of evaluation, NIME 2019," in In *Proceedings of the International Conference on New Interfaces for Musical Expressio*. NIME. Available online at: https://www.nime.org/archives/ (accessed June 12, 2020).

Mack, K., McDonnell E., Potluri, V., Xu, M., Zabala, J., Bigham, J., et al. (2022). "Anticipate and adjust: cultivating access in human-centered methods," in *CHI Conference on Human Factors in Computing Systems. CHI '22: CHI*

Conference on Human Factors in Computing Systems (New Orleans, LA: ACM), 1–18. doi: 10.1145/3491102.3501882

Magee, W. L. (2006). Electronic technologies in clinical music therapy: a survey of practice and attitudes. *Technol. Disabil.* 18, 139–146. doi: 10.3233/TAD-2006-18306

Magnusson, T. (2010). Designing constraints: composing and performing with digital musical systems. *Computer Music J.* 34, 62–73. doi: 10.1162/COMJ_a_00026

Matthews, C. (2018). *Customising Thumbjam for Accessible Music*. Available online at: http://www.drakemusic.org/blog/charles-matthews/customising-thumbjam/ (accessed January 10, 2020.).

Misje, R. (2013). *Music Technology in Music Therapy* [Master's thesis]. Bergen: University of Bergen.

Moog Music Inc (2023). *Theremini*. Moog Music Inc. Available online at: https://www.moogmusic.com/products/etherwave-theremins/theremini (accessed March 22, 2023).

Musii Ltd (2023). *Musii*. Available online at: https://musii.co.uk/ (accessed April 6, 2023).

Nordoff Robbins Evidence Bank (2014). *Music Therapy and Music and Health References and Resources*. NordoffRobbins.com. Available online at: https://www.nordoff-robbins.org.uk/wp-content/uploads/2019/06/research_and_resources_for_music_therapy_2016_final.pdf (accessed January 9, 2016).

Odom, W., Wakkary, R., Lim, Y. K., Desjardins, A., Hengeveld, B., Banks, R., et al. (2016). "From research prototype to research product," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (New York, NY: ACM), 2549–2561. doi: 10.1145/2858036.2858447

Ofsted (2012). *Music in Schools: Wider Still, and Wider*. Gov.uk: Ofsted. Available online at: https://www.gov.uk/government/publications/music-in-schools (accessed March 4, 2017).

Oliver, M. (1996). *Understanding Disability: From Theory to Practice*. Basingstoke: Palgrave Macmillan. doi: 10.1007/978-1-349-24269-6

O'Malley, C., and Stanton Fraser, D. (2004). *Literature Review in Learning with Tangible Technologies*. Available online at: http://opus.bath.ac.uk/9506/ (accessed August 12, 2016).

OpenOrchestras (2023). *Instruments*. Available online at: https://www.openorchestras.org/instruments/ (accessed March 5, 2023).

OpenUp Music (2023). *Home*. Available online at: https://www.openupmusic.org/ (accessed April 6, 2023).

Piggot-Irvine, E. (2006). Sustaining excellence in experienced principals? Critique of a professional learning community approach. *Int. Electron. J. Leadersh. Learn.* 10 doi: 10.1080/13674580601024432

Raspberrypi.org (2020). *Raspberry Pi 4*. Available online at: https://www.raspberrypi.org/products/raspberry-pi-4-model-b/ (accessed May 4, 2021).

Reason, P., and Bradbury, H. (eds) (2008). *Handbook of Action Research: Participative Inquiry and Practice*. London: Sage. doi: 10.4135/9781848607934

Robson, C. (2002). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*, 2nd ed. Chichester: John Wiley and Sons Ltd.

Saunders, M. N., Lewis, P., and Thornhill, A. (2015). *Research Methods for Business Students, 6th Edn*. Harlow: Pearson Education Limited.

Skoogmusic (2023). *Skoogmusic*. Available online at: https://skoogmusic.com/ (accessed April 6, 2023).

Skuse, A. H., and Knotts, S. (2020). "Creating an online ensemble for home based disabled musicians: disabled access and universal design-why disabled people must be at the heart of developing technology," in *Proceedings of the International Conference on New Interfaces for Musical Expression. Birmingham, July* 21-25 (New Interfaces for Musical Expression), 137–143.

Sonosaurus (2009). *ThumbJam*. Available online at: https://thumbjam.com/ (accessed September 10, 2020).

Soundbeam (2023). *Soundbeam*. Available online at: https://www.soundbeam.co.uk/ (accessed April 6, 2023).

Stoffregen, P. (2014). *Audio: Teensy Audio Library*. Available online at: https://github.com/PaulStoffregen/cores (accessed July 13, 2017).

Streeter, E. (2007). Reactions and responses from the music therapy community to the growth of computers and technology - some preliminary thoughts. *Voices* 7. doi: 10.15845/voices.v7i1.467

Sullivan, J., and Wanderley, M. M. (2019). "Surveying digital musical instrument use across diverse communities of practice," in *Computer Music Multidisciplinary Research (CMMR)* (Berlin; Heidelberg: Springer-Verlag), 745.

Swingler, T. (1998). 'That Was Me' Applications of the Soundbeam MIDI Controller as a Key to Creative Communication, Learning, Independence and Joy," in *Technology for Persons with Disabilities* (Northridge, CA: California State University Northridge).

Trump, B. (2016). *Orphion [Mobile application software]*. iOS. Bastus Trump.

Tzankova, V., and Filimowicz, M. (2018). *New Directions in Third Wave Human-Computer Interaction: Volume 2-Methodologies*. Berlin: Springer. doi: 10.1007/978-3-319-73374-6

Vogiatzoglou, A., Ockelford, A., Welch, G., and Himonides, E. (2011). Sounds of intent: interactive software to assess the musical development of children and young people with complex needs. *Music Med.* 3, 89–195. doi: 10.1177/1943862111403628

Waiswisz, M. (1985). "The hands, a set of remote MIDI-controllers," in *The International Computer Music Conference*, Burnaby. 19-22 August (Ann Arbor, MI: Michigan Publishing).

Wanderley, M. M. (2001). "Gestural control of music," in *The International Workshop Human Supervision and Control in Engineering and Music, Journal of New Music Research*, ed J. Gunnar.

Ward, A., Davis, T., and Bevan, A. (2019). Music technology and alternate controllers for clients with complex needs. *Music Ther. Perspect.* 37, 151–168. doi: 10.1093/mtp/miz006

Ward, A., Woodbury, L., and Davis, T. (2017). "Design considerations for instruments for users with complex needs in SEN settings," in *Proceedings of 2017 International Conference on New Interfaces for Musical Expression* (Copenhagen).

Wessel, D., and Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Comput. Music J.* 26, 11–14. doi: 10.1162/014892602320582945

Wright, J., and Dooley, J. (2019). "On the inclusivity of constraint?: creative appropriation in instruments for neurodiverse children and young people," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, eds R. Schramm, and M. Johann (Porto Alegre: Institute of Informatics), 3–6.

Yins, R. (2018). *Case Study Research and Applications, 6th Edn*. Los Angeles, CA: Sage.

Ymous, A., Spiel, K., Keyes, O., Williams, R. M., Good, J., Hornecker, E., et al. (2020). "'I am just terrified of my future' epistemic violence in disability related technology research," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI: ACM), 1–16. doi: 10.1145/3334480.3381828

Zimmerman, J., Forlizzi, J., and Evenson, S. (2007). "Research through design as a method for interaction design research in HCI," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (San Jose, CA: ACM), 493–502. doi: 10.1145/1240624.1240704

frontiers | Frontiers in Computer Science

TYPE Review
PUBLISHED 05 October 2023
DOI 10.3389/fcomp.2023.1156693

OPEN ACCESS

EDITED BY
Arianna Astolfi,
Polytechnic University of Turin, Italy

REVIEWED BY
Marcia Epstein,
University of Calgary, Canada
Veronica Amodeo,
Università di Firenze, Italy
Kjetil Falkenberg,
Royal Institute of Technology, Sweden
Martin Ljungdahl Eriksson,
University West Trollhättan, Sweden, in collaboration with reviewer [KF]

*CORRESPONDENCE
Sara Lenzi
✉ s.lenzi@tudelft.nl;
✉ sara.lenzi@deusto.es

RECEIVED 01 February 2023
ACCEPTED 18 September 2023
PUBLISHED 05 October 2023

CITATION
Lenzi S, Spagnol S and Özcan E (2023) Improving the quality of the acoustic environment in neonatal intensive care units: a review of scientific literature and technological solutions. *Front. Comput. Sci.* 5:1156693. doi: 10.3389/fcomp.2023.1156693

Improving the quality of the acoustic environment in neonatal intensive care units: a review of scientific literature and technological solutions

Sara Lenzi[1]*, Simone Spagnol[2] and Elif Özcan[1,3]

[1]Critical Alarms Lab, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands, [2]Department of Architecture and Arts, Iuav University of Venice, Venice, Italy, [3]Department of Adult Intensive Care, Erasmus Medical Center, Rotterdam, Netherlands

There is an increased awareness of how the quality of the acoustic environment impacts the lives of human beings. Several studies have shown that sound pollution has adverse effects on many populations, from infants to adults, in different environments and workplaces. Hospitals are susceptible environments that require special attention since sound can aggravate patients' health issues and negatively impact the performance of healthcare professionals. This paper focuses on Neonatal Intensive Care Units (NICU) as an especially sensitive case representing a hostile acoustic environment in which healthcare professionals have little awareness of how unwanted sounds impact the perceived quality of the soundscape. We performed a semi-systematic review of scientific literature on sound assessment studies in NICU from 2001. A thematic analysis was performed to identify emerging themes that informed the analysis of 27 technological solutions for the assessment of sound quality in indoor and outdoor environments. Solutions were categorized by functions and evaluation methods and grouped according to the characteristics of the design components, i.e., acquisition, computation, and communication strategies. Results highlight a lack of solutions to assess the qualitative characteristics of indoor environments such as NICU and forecast the footprint that different sound sources have on the indoor soundscape. Such solutions are urgently needed to empower healthcare professionals, and especially nurses, to actively modify and prevent the negative impact of unwanted sounds on NICU and critical care soundscape.

KEYWORDS

neonatal intensive care unit (NICU), acoustic environment evaluation, indoor sound quality, soundscape perception, Machine Listening, indoor soundscape assessment, indoor soundscape modelling

1. Introduction

We, as a society, have developed a certain tolerance toward unwanted sounds. Yet, several studies have shown that sound pollution-defined by sound level thresholds beyond which the exposure to sound can negatively impact health (Kang et al., 2016)-has adverse effects on many populations, from infants to adults (Gupta et al., 2018; Teixeira et al., 2021), in different environments and workplaces. Within the broader spectrum of sound studies, soundscape research considers environmental sound "as resource rather than a waste" (COST TUD Action, 2013), with various actions and interventions put in place to actively improve it (Hellström et al., 2014; Moshona et al., 2022; Henze et al., 2023), instead of "silencing" it (Thibaud and Amphoux, 2013). Rather than quantitively measuring environmental sound levels, soundscape studies approach the assessment of the acoustic environment by identifying qualitative descriptors (e.g., pleasantness, perceived annoyance, appropriateness, quietness, and so on) and investigating their potential correlation with the "footprint" of different sound events (i.e., human- or machine-generated, and natural sounds) to characterize the acoustic environment as listeners perceive it (Aletta et al., 2016).

In the past decade, this approach has informed the development of novel algorithms for the automatic assessment of the acoustic environment. These algorithms are typically modeled on the outdoor soundscape and build on the growing availability of low-cost sensors for the continuous monitoring of sound levels (De Coensel and Botteldooren, 2014; De Coensel et al., 2015). This line of research aims to develop intelligent systems (Wei and Van Renterghem, 2014; Socoró et al., 2017; Quinn et al., 2022) and novel indices (Brocolini et al., 2012; Graziuso et al., 2022) to characterize, interpret, and forecast the acoustic environment beyond traditional sound level measurements and closer to how sound is perceived by humans. A recent review of 24 studies on predictive models of urban soundscapes (Lionello et al., 2020) indicates that the combination of *quantitative* (i.e., acoustic and psychoacoustic metrics) and *qualitative* analysis (i.e., subjective and perceptual information on the perceived affective quality of a soundscape) greatly improves the performance of algorithms to model and predict the outdoor soundscape, compared to algorithms that only use acoustic and psychoacoustic indicators. Additionally, more reliable results seem to be achieved when descriptors of the affective quality of the soundscape are combined with the *categorization* of sound sources and information on the appropriateness to a specific context.

Within professional socio-technological environments (i.e., functional settings with a specific mission that relies on time-sensitive actions and teamwork) such as the hospital, it becomes critical to identify and define the role of humans as both recipients and producers of sound events who impact the quality of the soundscape. Such identification will eventually lead to increased awareness among healthcare professionals on the impact of sound on the functionality of their shared acoustic space, and improved guidelines for the design of more "actionable" (Özcan et al., 2022b) healthcare spaces. Previous research on auditory affordances supports the claim that people understand and relate to soundscapes through their potential to induce and guide action (Rosenblum et al., 1996; Nielbo et al., 2013). How the potential for action relates to annoyance of sound events has also been recently investigated (Misdariis et al., 2019) with the goal of defining computational models for the evaluation of the quality of the urban soundscape.

In the hospital context, medical alarms, sounds from medical equipment, and the continuous human activity within units cause the typical soundscape to be perceived as poor (Bliefnick et al., 2019). However, existing studies on this topic mainly refer to a progressive and harmful increase in sound levels in hospitals (see Busch-Vishniac et al., 2005; Busch-Vishniac and Ryherd, 2019 for a review of the past 40 years). While research has shown how different sound sources (e.g., medical alarms) can create a stressful environment for medical staff (Johnson et al., 2017; Varisco et al., 2021) and how sounds from staff conversations and activities negatively impact the patients' soundscape quality (MacKenzie and Galbrun, 2007; Konkani et al., 2014; Lenzi et al., 2023), to the knowledge of the authors no comprehensive study has been conducted to assess, model, and predict the quality of the hospital soundscape beyond sound level measurements. In this paper we take a role to understand the status quo of the current technological solutions and the technological trend for near future applications for monitoring the sound of professional socio-technological environments. Our intention is to find opportunities to envision a targeted solution for hospitals at large and more specifically neonatal intensive care units in focus.

NICU soundscapes: problem statement

Of all shared spaces in society, neonatal intensive care units (NICU) are especially vulnerable environments in which patients, their families, and healthcare professionals are particularly subject to the harmful consequences of excessive sounds (Özcan et al., 2019; de Lima Andrade et al., 2021). The NICU is designed for premature neonates who are not necessarily ill upon their arrival. Nonetheless, they need special care to grow and survive, and their likelihood of getting ill inside the NICU is high since their bodies are not fully developed. Moreover, neonates are likely to experience physiological limitations, central nervous system limitations, and dependency on intensive care, which makes them more vulnerable to the whole NICU sound environment (Blackburn, 1998). Concerns regarding the impact of acoustic stimuli in the NICU were first addressed during the 1970s (Lawson et al., 1977) through direct observation of the environment. Later in the 1990s, an increased awareness of how environmental stimuli affect neonates' clinical conditions, and their neurodevelopment was recorded (Philbin et al., 2000). This led to several studies by which sound level measurements were carried out in the NICU with professional equipment such as sound level meters (Thomas, 1989; Thomas and Uran, 2007). Sound measurement studies continued throughout the years, providing recommendations for reducing the high sound level issue through room redesign (Chen et al., 2009), use of earmuffs (Duran et al., 2012), or the implementation of educational programs (Elander and Hellström, 1995; Calikusu Incekar and Balci, 2017).

Whereas, the benefits of training nurses and other hospital staff to decrease noise levels by incorporating behavioral changes is clear as an important first step (Carvalhais et al., 2015), previous research from the authors (Özcan et al., 2022a; Spagnol et al., 2022) shows that NICU occupants are often unaware of the contributors to the noisy sound environment and feel they have no control to change the sound quality of their environment. Thus, there is an urgent need to create a shared awareness about the contributors to the decreased sound quality in NICU to be able to take collective action. A recent study by the authors (Spagnol et al., 2022), that we further expand in Section Beyond sound measurements: a review of technological solutions, gaps, and opportunities of this paper, shows that current technological solutions for the assessment of indoor acoustic environments focus on collecting and measuring basic acoustic metrics such as sound pressure level (SPL, expressed in decibels). However, SPL measurements are difficult to interpret by non-experts (Alsina-Pagès et al., 2021). Additionally, they are not an appropriate indicator for describing the affective quality of sounds as perceived by humans (Aletta et al., 2016). Lastly, SPL measurements do not allow users to identify the cause of unwanted sounds (i.e., noxious sounds that can lead to harmful consequences). Sound identification is crucial to increase awareness on the NICU soundscape toward its active improvement, and it is a highly context-dependent cognitive function (Özcan and van Egmond, 2007, 2009; Axelsson, 2015). Finally, protocols for the analysis of the affective qualities of the soundscape, such as the Swedish Soundscape Protocol (Axelsson et al., 2010) were developed in the context of outdoor environments, and the application of such protocols to indoor environments is currently under development (Torresin et al., 2020). Therefore, we see a clear opportunity to explore and design intelligent solutions for the algorithmic modeling of indoor soundscapes and make it technologically viable.

Across the globe, critical care departments ranging from neonatal to adult care feel the urgency to improve their sound quality for better patient experiences and working conditions. Hospitals run exhaustive studies with extensive resources and human effort to be able to characterize their existing soundscape and plan interventions accordingly in the lack of available standardized solutions (Özcan et al., 2022b). Therefore, a sustainable solution that can automatically assess and characterize hospital soundscapes is needed and will be timely in the era of data-centric approaches employed in healthcare.

In this paper, we review the state of the art of technological solutions for the assessment of the NICU acoustic environment to highlight current technological gaps and identify opportunities for the design of novel solutions to improve the NICU soundscape. We first present a semi-systematic review of 77 publications on the topic of environmental sound in the NICU (Section Monitoring the sound quality of NICU: a semi-systematic review). We then define an evaluation framework and discuss the results of a review of 28 current technological solutions for the assessment of sound quality in NICU and, more broadly, in indoor and outdoor spaces (Section Beyond sound measurements: a review of technological solutions, gaps, and opportunities). The review supports the definition of the design requirements for a solution able to increase nurses' awareness of the impact of sounds on the NICU soundscape. The characteristics of this novel technological solution are discussed in Section Conclusions.

2. Monitoring the sound quality of NICU: a semi-systematic review

This section reports a semi-systematic literature review carried out to have an overview of the sound monitoring studies that had been conducted inside the NICU. Contrary to systematic reviews, which identify and analyze all the available empirical evidence to quantitatively answer specific research questions or hypotheses, a semi-systematic review has a broad research question, examines research areas and follows their evolution over time, and synthesizes the main themes from the literature using meta-narratives instead of quantitative methods (Snyder, 2019). The resulting themes allow collecting insights and limitations from the literature, which will guide the review of technological solutions in Section Beyond sound measurements: a review of technological solutions, gaps, and opportunities and will be later translated into opportunities for research and design in Section Conclusions.

2.1. Methodology

Our research started by identifying relevant studies on environmental sound in the NICU. First, to retrieve relevant titles, we queried the popular academic literature search engine PubMed with the following search string, *[NICU OR (Neonat* AND "Intensive Care")] AND (Noise OR "Sound Level*")*. The query returned 77 articles. Then, we excluded (1) articles published more than 20 years ago, i.e., before 2001; (2) articles written in another language than English; (3) non-journal publications; (4) entries without a full text available; (5) duplicate entries. After this filtering phase, 59 articles were left.

Upon carefully reading all the 59 articles, we generated a table of different factors that could hint at potential research themes, such as targeted listener (e.g., neonate, nurse, family); methodologies used for assessment (e.g., measurements, questionnaires, structured interviews); devices used for sound monitoring. Finally, we further excluded those articles that did not actually report the results of environmental sound recordings inside a NICU. The final sample included 41 articles. Thematic analysis, i.e., a qualitative data analysis method, was used to code, analyze, and report patterns in the form of themes (Braun and Clarke, 2012). In particular, the coding phase consisted of highlighting sections of text and coming up with shorthand labels, or *codes*, to describe their content. For this task, the ATLAS.ti 9 software was used. Once codes were written for all articles, higher-level categories were formed from patterns in the codes. Lastly, categories were clustered into four main themes. When interpreting and explaining themes, insights and limitations emerged. Figure 1 schematically reports the above-described process.

2.2. Results

Table 1 reports an overview of the four themes that emerged from the thematic analysis of the literature focusing on recorded sound/noise levels in NICUs. Each theme covers several categories, which are listed below it. We now present an overview of the themes along with the corresponding insights and limitations.

Theme 1: Collecting and processing sound focuses on methods used for sound measurement and recording, along with their outcomes.

1. Measurement time spans are generally scattered, and studies hardly follow the same protocol [for instance, a 24-h period every week for a total of 44 weeks (Brandon et al., 2007), 168 consecutive hours (Aita et al., 2021), or eight separate 1-h recordings (Krueger et al., 2007)]. Currently, all we can find are studies that are episodic rather than continuous. Most of

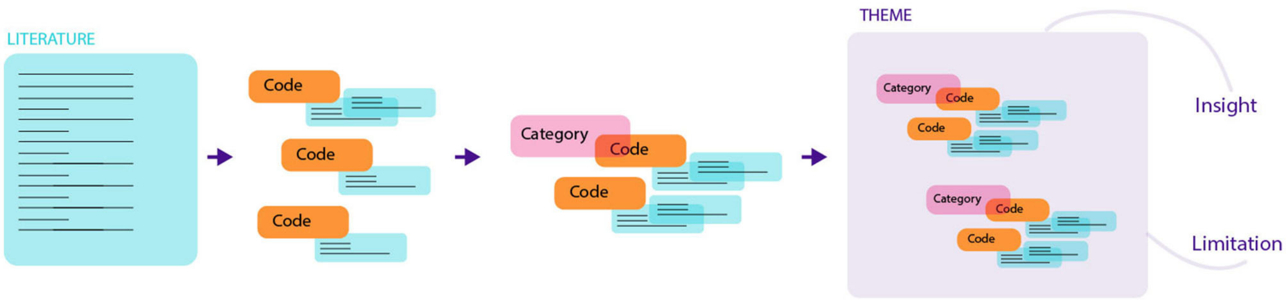

FIGURE 1
Thematic analysis as a semi-systematic review methodology.

TABLE 1 The four main themes from the thematic analysis and their underlying categories.

Theme 1: collecting and processing sound	Theme 2: assessing the NICU environment
Length of the study	Standards and organizations
Measurement intervals	NICU occupancy
Outcome variables	Monitoring area
Monitoring devices	Device placement
Device calibration	Comparing different environments
Theme 3: interactions with sound	**Theme 4: beyond sound measurements**
Sound sources in the NICU	Subjective factors and outcomes
Environmental variables	Qualitative assessments
Noise control	Behavioral programs
Alarm management	Sound education
	Structural changes

them focus on measuring sound levels to report how harmful the auditory environment can be but do not focus on implementing potentially long-lasting sound monitoring solutions.

2. Sound level meters are the most used devices for measurement, followed by sound dosimeters (Liu, 2010; Ramm et al., 2017; Smith et al., 2018). Only a few studies use available sound level monitoring solutions for healthcare (Milette, 2010; Casey et al., 2020) or smaller devices such as probe microphones (Surenthiran et al., 2003) to measure neonates' exposure to sound. Sound level meters generally provide extremely accurate yet objective measurements of the auditory environment. More intuitive interpretations are needed to give individuals a concrete means of evaluating the sound environment as they experience it.
3. The outcome variables are almost exclusively measurements expressed in decibels (dB), most often A-weighted (dBA). Only a few studies also conduct spectral analysis, therefore analyzing sound power at different frequencies (Surenthiran et al., 2003; Livera et al., 2008; Lahav, 2015). However, excluding a time-frequency analysis from a sound recording limits a complete perspective of sound events occurring in the NICU. Privacy issues might be the main reason why sound recordings are not stored and analyzed, which limits their possible use as training data for sound event detection approaches.

Theme 2: Assessing the NICU environment is about the environmental factors and experimental configurations that studies aim to assess.

4. The goal of most of the considered studies is to report sound levels exceeding the recommended thresholds. Baseline levels are established by national and international organizations such as the American Academy of Pediatrics (AAP), the US Environmental Protection Agency (EPA), or the World Health Organization (WHO) (Williams et al., 2007; Darcy et al., 2008). The general problem with these baseline levels is that they are seen as too low and therefore hard to reach within an environment that several different individuals visit or work in.
5. A substantial number of studies focus on not only reporting sound levels in a specific unit but comparing different environments: for instance, NICUs of different levels of care (Levy et al., 2003), open bay units vs. single-family rooms (Liu, 2012; Szymczak and Shellhaas, 2014), or before vs. after a structural change in the unit (Krueger et al., 2007; Aita et al., 2021). These comparisons aim to report and give evidence on the most suitable environment for the wellbeing of neonates, parents, and nurses. Unfortunately, the NICU characteristics in which every study is conducted are unique. Among the characteristics that change we can count patient census, number of beds, number of nurses working during shifts, presence of parents, to name but a few. All these factors can potentially contribute to an increase in average and/or peak sound levels. It is even more challenging to compare outcomes from different studies since they do not share the same settings.
6. In the analyzed literature, most researchers explain where measurement devices are positioned. If the goal is to measure environmental sound, devices are often positioned at the center of the room (Livera et al., 2008; Lahav, 2015). Conversely, they are placed close to neonates' heads when the goal is to measure either subjective exposure or care activities nearby the incubator area (Surenthiran et al., 2003; Liu, 2010). A few studies give a more extensive mapping by placing measurement devices in several different locations within the unit (Krueger et al., 2005; Wang et al., 2014). As single measurement devices are used,

measurements in different locations are not time synchronized. The main reason for this experimental choice could be the lack of resources and/or budget. The use of a set of independent devices would give a more complete picture of sound levels and events within the unit.

Theme 3: Interactions with sound is about auditory events in the NICU and how the staff (mainly nurses) deal with them.

7. There is consensus on the most relevant sound sources in the NICU. The most cited categories are related to equipment, i.e., alarms, incubators, mechanical ventilation systems (Lasky and Williams, 2009; Liu, 2012; Restin et al., 2021), and speech (Lahav, 2015; Hernández-Salazar et al., 2020). Alarm levels can neither be set below a certain threshold, nor turned off, meaning that alarm-induced sound level issues can only be addressed through rules and regulations and with the collaboration of stakeholders involved in the manufacturing and supply chain. Furthermore, it is very difficult to avoid voice communication, especially in such a human-centered environment. The goal toward reducing sound levels must go in accordance with the care activities carried out in the unit.
8. Nurses are naturally considered as the main source for human-induced sound nuisance because of their constant presence and continuous activities within the NICU. Although nurses commit to keeping a quieter NICU environment for the wellbeing of neonates, a commonly seen issue is that they are unaware of how loud the sounds they produce can be and how susceptible their environment is (Darcy et al., 2008; Ahamed et al., 2018). The lack of (real-time) feedback in the NICU, such as a visual representation of the impact of human activity on the overall sound level, might de-prioritize individual sound awareness.

Theme 4: Beyond sound measurements includes assessments outside the domain of sound, as well as strategies aimed at reducing sound levels.

9. Subjective measurements are necessary to assess individual sound exposure. Literature generally presents two different approaches, depending on the targeted population. In the case of neonates, it is correlated to alterations in heart rate, blood pressure, and oxygen saturation (Williams et al., 2009; Smith et al., 2018). Conversely, questionnaires and interviews are the classical methods used to evaluate staff tolerance and awareness toward sound (Darcy et al., 2008; Trickey et al., 2012). However, studies focusing on the repercussions of the sound environment on parents are scarce, if not absent.
10. Alongside structural changes, behavioral change strategies and the implementation of educational programs are recurrent patterns in literature (Milette, 2010; Wang et al., 2014; Ahamed et al., 2018), although some authors already point out that they are not effective in the long term (Liu, 2010; Carvalhais et al., 2015). It is indeed uncertain to which extent behavioral strategies can be sustained long-term without periodic reinforcement. Therefore, it can be possible that the found effects are only temporary.

The results of the literature review reveal the already existing practice for sound recording and analysis in NICUs, albeit brief and for research purposes. However, there seems to be a need to understand the effect of NICU sound environment on nurses and patients (and on families, to date scarcely investigated) by psychological and physiological measurements and the need to use objective measurements for long-term behavior change through knowledge. This acknowledges that sound is an issue in NICUs worldwide and a threat to the wellbeing of its occupants. Moreover, NICU soundscapes are also found to be susceptible to human-environment interactions indicating that NICU occupants do contribute to sound levels. Yet, the nurses especially seek to understand individual sound sources and monitor the behavior of sound events over time to be able to take action to reduce sound pollution. For design purposes, these outcomes support the need for an intelligent system that enhances nurses' understanding of environmental sounds by continuously monitoring, analyzing, and explaining the acoustic environment in terms of sound sources and perceptual characteristics. In other words, we need to move beyond the often-sporadic measurement of noise levels with a physical descriptor that does not consider human perception and is difficult to make sense of for the non-expert. In the next Section, we closely look at 27 existing solutions to assess their capacity to measure, characterize and interpret (and possibly, forecast) the acoustic environment for long-term sound awareness and noise management in the NICU.

3. Beyond sound measurements: a review of technological solutions, gaps, and opportunities

To complement the literature review, we conducted a technology search to identify those products that are available on the market and could potentially serve as a solution to reducing sound pollution in the NICU. The search was initially based on sound monitoring solutions for hospitals. Given the extremely limited availability of such solutions (see Section Conclusions), our search criteria were extended. Our approach was then to start the search broad and include solutions that would cover multiple contexts (indoors and outdoors), then funnel the solutions into indoors, healthcare, and NICU, respectively. We included both commercial solutions that are marketed products and services for the assessment of the acoustic quality as well as academic response to the sound monitoring needs that would have produced concept solutions (i.e., demonstrators with prototypes). The latter is important as it showcases the trend for future applications and indicates where the technology might be best applied. Previous research from two of the co-authors on six case studies (Spagnol et al., 2022) categorized solutions by product complexity, customization options, active or passive feedback, and interpretative and predictive power. The latter two attributes define the threshold between solutions that only quantify the physical quality of sound events in terms of standard indices such as dB levels and solutions that use qualitative data to provide meaningful information that empowers the user to improve the acoustic environment. The study showed how some of the existing sound monitoring solutions can also generate some sort of reports that describe the acoustic environment at a deeper level (i.e., through the behavior of noise levels over time) but are unable to characterize perceptual properties and interpret it in terms of sound sources and even less so, to predict its dynamic nature by a well-defined model under different circumstances or over time. In this review, we increased the number of cases, and we deepened the analysis taking into account the entire design life cycle of a technological solution and its three main components: the type of data collected from the environment (Acquisition), how data are processed to gain insights into the phenomenon under study (Computation), and how information is displayed (Communication) for the final user to make sense of the sound phenomenon.

3.1. Methodology

The 27 solutions we considered in this review were mainly retrieved through a manual search on the most popular web search engines with the keywords "sound measurement tools," "acoustic measurements systems," "noise evaluation tools" "noise measurement tools." We also combined all these searches with the keywords "healthcare," "wellbeing," "NICU," "intensive care," and "critical care." To this search, we added solutions of which we had direct knowledge and solutions retrieved through the previous literature review. We excluded off-the-shelf sound level meters such as those manufactured and commercialized by Castle or Amprobe as the technological value of these tools relies fundamentally on the quality of the integrated microphones rather than in the design of the solution as we define it in this review (Acquisition, Computation, Communication). Additionally, professional tools such as these require the presence of a trained operator both to capture the sound levels and to interpret the results, of difficult interpretation for lay people. As mentioned, we are interested in solutions that can increase nurses' awareness of the contribution of sound to an unhealthy experience in the NICU and therefore in solutions that do not need the presence of audio experts. We also excluded both freemium and premium mobile applications marketed with the same purpose. These are simplified versions of hardware sound meters that use lower-level audio equipment (i.e., the smartphone's own microphones). As such, they provide non-expert users with less reliable basic information such as an average SPL at a given time.

Because the focus of the review is on technological solutions for measuring the quality of acoustic environments, we first listed current methods that evaluate environmental sounds. These methods include (i) acoustic and psychoacoustic indices for the quantitative measurement of sound, (ii) indicators of perceived affective qualities of soundscapes (Aletta et al., 2016), and (iii) applications of machine learning (ML) to categorize sound events and forecast their behavior in a given context. For each of the 27 solutions we analyzed available public documentation (websites, scientific publications, demos, and videos) to list specific indices, indicators, and ML techniques used by manufacturers. Through an iterative process, we grouped methods by the function they fulfill. Figure 2 summarizes the assessment methods identified across the 27 solutions and the four Functions they define: Measure, Characterize, Interpret, and Forecast. As shown in Figure 2, each Function builds upon the other and integrates the previous evaluation method on a scale of complexity that goes from standard acoustic measurements to the characterization of soundscape in terms of its perceptually relevant dimensions (e.g., pleasantness, eventfulness, annoyance, monotony, homogeneity), then to the classification of sound events to be able to interpret the acoustic environment in terms of the footprint of sound events, and finally, the use of Artificial Intelligence (AI) to model soundscapes and forecast the dynamic behavior of acoustic environments. This distinction of functions will allow us to distribute the existing technological solutions (i.e., cases) in an organized way.

The correlation of the four Functions with the three data-driven design components (Acquisition, Computation, and Communication) defines a matrix (see Table 2) that guides the analysis of the cases. Based on the matrix, we describe each component for each selected case in an iterative process to identify common features around which the cases were grouped.

Solutions that allow users to Measure the acoustic environment only extract physical features from sound (i.e., SPL) in real time. These measurements are typically used by private and public enterprises to comply with existing regulations on noise levels (such as the European Parliament., 2000/14/EC) both outdoors (e.g., in public urban spaces) and indoors (e.g., in offices, school, hospitals). Particularly interesting to this analysis are those solutions marketed for healthcare environments and specifically for NICU and/or ICU (McLennan Sound Monitoring, Sound Intelligence). In these solutions, data is collected manually by trained users with specialized equipment or automatically by sensors permanently installed at the customer's premises (Noisemote, Noisescout). Data are displayed in real-time (as in hand-held acoustic cameras like Sorama, which produces and displays real-time heatmaps of the indoor acoustic environment, see Figure 3) and sent to a centralized repository where analytics and reports are generated and regularly sent to customers (gfai tech, Norsonic and Sorama). Among the most common methods to represent and communicate findings to the end users are heatmaps (in the form of so-called noise maps), spectrograms and other basic diagrams representing the evolution of the physical descriptors of sound (e.g., frequency and amplitude) over time (see Figure 3 for an example of how data are displayed for the user). All the solutions in the Measure category are commercialized and have a high Technology Readiness Level (TRL) e.g., between 7 and 9.

Solutions that not only measure the physical characteristics of sound but aim to Characterize the soundscape as perceived by humans in context (ISO 12913-1., 2014) tend to offer both real-time measurements and historical data of dB levels for monitoring and trend analysis. Sometimes acoustic data are coupled with other information such as air quality or temperature (Quietyme, SonicU) or are scored to represent the acceptable ranges (e.g., sleep score by Quietyme). These products are commercialized to couple compliance with regulations with data intelligence that can support decision-making on noise mitigation policies both in private and public enterprises, indoors and outdoors (e.g., hospitals, construction sites). Other solutions that belong to this group try to engage users in the data acquisition process to increase awareness on the impact of noise. Typically, a mobile app is provided to collect data such as audio recordings and surveys on qualitative and perceptual characteristics of sound events (e.g., peaceful, uneventful, chaotic, pleasant) (MosART, Harmonica Project). Data

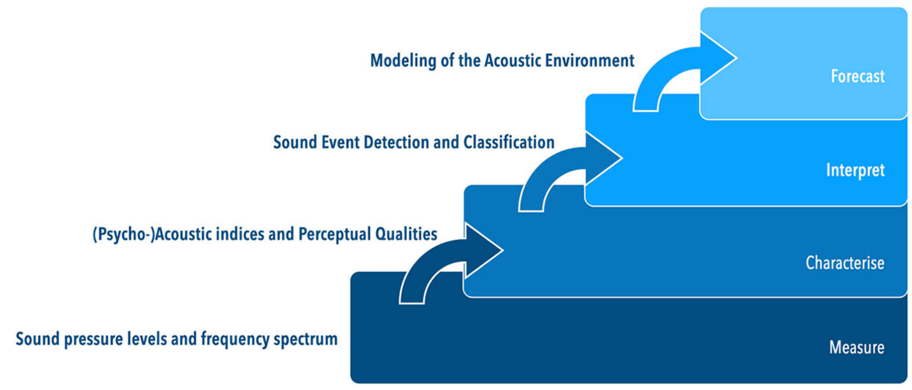

FIGURE 2
The four functions of technological solutions for sound quality evaluation, defined by the sound evaluation methods used. Each function builds and integrates the previous one.

TABLE 2 Evaluation matrix for the 27 technological solutions.

Data-driven design components	Functions of the design solution			
	Measure	Characterize	Interpret	Forecast
Acquisition (data collection)	Real-time data	Real-time and time-series data	Real-time and big data	Big data from multiple sources (e.g., GIS data)
Computation (data processing)	Basic signal processing	Basic signal processing Perceptual evaluations	Advanced signal processing Machine learning	Machine learning
Communication (data display)	Sound level displays Spectrograms Plots, diagrams Heat maps	Sound level displays Warning signals Data analytics (trendlines, noise maps) Surveys, questionnaires Web sound maps	Dashboards Web sound maps	Dashboards GIS maps

is later uploaded to a central repository for analysis and, in some cases, public communication (for instance, via crowdsourced soundscape maps on the web as in HushCity, Radicchi et al., 2021). Solutions in this category are for the vast majority available on the market both for purchase and as a free download (in the case of mobile applications), with an average TRL of 7. A small percentage is still in an R&D phase (TRL 3 or 4). These solutions mainly refer to academic or publicly funded research to mitigate the impact of sound pollution in contexts such as public health (Mietlicki et al., 2014; Misdariis et al., 2019). As we discuss in Section Conclusions, we believe that the inclusion of solutions that are currently only modeled and applied to the outdoor context can provide relevant information to support the design of novel technological solutions for the assessment of the quality of the indoor soundscapes such as the NICU.

We define solutions that can Interpret the acoustic environment as products or systems that leverage big data collected through networked sensors over a longer period. These solutions use AI methods such as ML or the so-called Machine Listening (i.e., the processing of sounds through a computer in a way that mimics human auditory cognition) in combination with signal processing techniques to interpret the sound environment in terms of detection of sound events, and classification of sound sources so that the "footprint" (i.e., the impact on the quality of the acoustic environment) can be assessed. As we will further detail in the following Section, these solutions are characterized by a lower TRL (between 2 and 5). Some solutions, notably those that address the specificity of the NICU context, are currently concept solutions (Özcan et al., 2022c; Spagnol et al., 2023). These solutions pay special attention to how information is displayed to the user (see Figure 3) to support sense-making of data for long-term behavior change, and their proposed interface has been designed considering the nurses' needs and expectations as collected through qualitative research (interviews, observations). Lastly, some solutions focus on detecting and, in some cases predicting the behavior of unwanted sound events in outdoor environments as part of academic endeavors, sometimes in collaboration with industrial partners (Salamon et al., 2016; Sevillano et al., 2016; Misdariis et al., 2019). No commercially available solutions fit in this category. Authors recognize that current measurements are still "insufficiently understood by the general public and authorities" and use sound descriptors that are "complicated to explain and relatively far-removed" from human perception (Mietlicki et al., 2014). These solutions are interesting to the present study as they focus on the algorithmic modeling of an acoustic environment to inform the end-user on the auditory footprint of different sound sources, with the goal of supporting better decision-making to mitigate the negative effects of sound

FIGURE 3
Examples of interfaces for each function from the database of solutions under study.

pollution. In Section Conclusions, we comment on how similar approaches could be applied to the NICU context in order to increase nurses' awareness with actionable knowledge on the role of different sound sources.

Solutions that can Forecast the acoustic environments in the longer term can be understood as like weather forecasts, i.e., they should provide information that we can easily consult to make daily and long-term decisions that are informed by patterns found in big and historical data used for generalization and prediction (i.e., modeling). When applied to the assessment of acoustic environments, this area of investigation is still in its infancy and despite a clear increase in the research effort (Bianco et al., 2019) no viable solutions are yet available on a large scale, let alone commercial applications. In the context of outdoor soundscapes, experimental solutions apply Deep Learning and other ML techniques on multiple data streams at the same time (e.g., geospatial data, sound recordings, traffic data, weather data related to seasonal conditions, etc.) to ultimately augment the human listening capabilities in terms of sound events recognition, classification and prediction, and perceptual interpretation of the acoustic environment (eventfulness and pleasantness) (Mitchell et al., 2021). In the authors' intention (Sharan and Moir, 2016; Pedersen et al., 2018), this augmented knowledge will allow for better management and planning to take action, both at the institutional and individual level. In the NICU context, a solution that can predict the quality of the soundscape while also assessing the impact of different sound categories (e.g., speech by nursing staff, machinery, medical equipment) could support behavioral change among nurses but also, in the long-term, guide the management in better planning for the NICU activity and inform design decisions for device manufacturers and architects.

Once defined the matrix to classify the solutions, we proceeded to analyze the selected cases in terms of Function (Measure, Characterize, Interpret and Forecast), Scope (Outdoor/Indoor) and application to the Healthcare and NICU context. In the following section, we present our findings for each category and discuss specific representative cases.

3.2. Analysis and discussion of results

Figure 4 offers an overview of the selected solutions by Function. Of the 27 cases analyzed, 12 are used to measure the acoustic environment, eight aim to characterize it, seven provide for a more sophisticated interpretation, and one only case represents the current research effort in applying various degrees of ML techniques to augment human listening capabilities to look into the future of the acoustic environment. The reader can find the complete list of cases with the metadata used in this review study and references to external resources as Supplementary material.

In the Measure group, Noise Aware represents a baseline solution that provides customers with a real-time alarm whenever it detects a breach of the established noise threshold. It comes with a noise level detector installed on mostly real-estate properties for one's "peace of mind" and "protection of profits" (Noise Aware website., 2022). NIOSH app is a representative of a broad category of hardware and software devices for the real-time measurement and display of noise levels in decibels (dB). We consider it a special case as it is released by a public authority, the United States National Institute for Occupational Safety and Health "to help workers make informed decisions about their noise environment and promote better hearing health and prevention

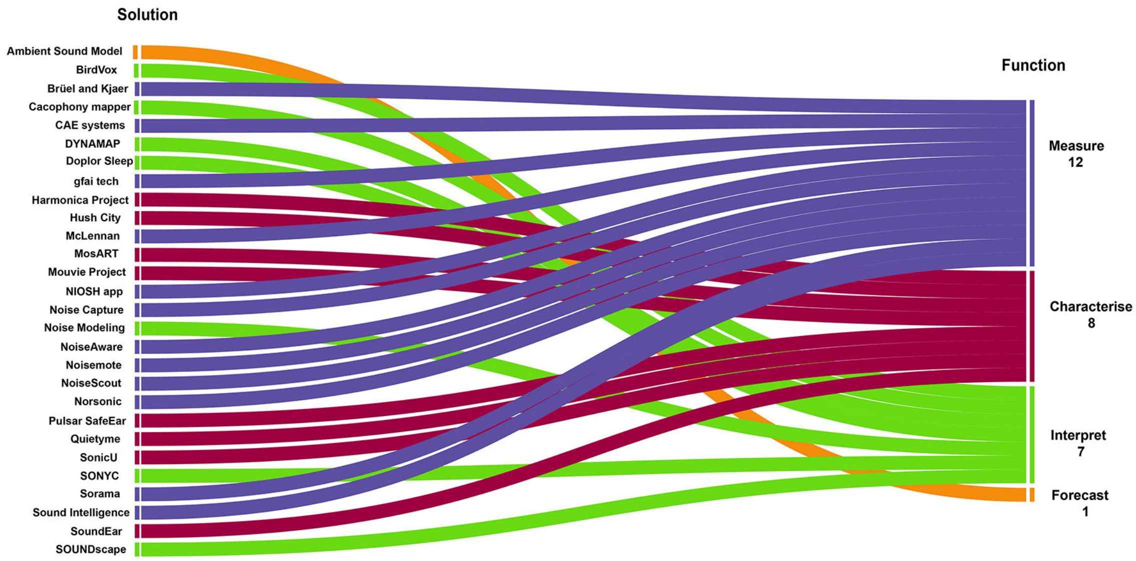

FIGURE 4
List of solutions by function (measure, describe, interpret, and forecast).

efforts" (NIOSH website., 2022). Like NIOSH, the Noise Capture app stands out as a crowdsourced project that encourages citizens to capture the dB level in the urban space with their mobile phones and upload the measurements to a shared web map, in the context of the European Union effort on sensitizing the population on the negative effects of urban noise on health. The solutions of Noisemote, Noisescout and McLennan Sound Monitoring leverage permanently installed networks of noise sensors, mainly in an outdoor context, to continuously measure sound levels. Next to real-time alerts, they also provide access to a dashboard for continuous monitoring of dB average and peak levels along the network. Finally, solutions provided by Sorama, Brüel and Kjaer and CAE systems are at the forefront of this market as they leverage holographic technology (an increasingly popular technique to estimate sound wave propagation for better source identification and localization) to provide a more accurate estimation of sound events' SPL and source localization through handheld and hand-moved acoustic cameras.

Solutions that Characterize the acoustic environment, move beyond noise measurement to inform users on the psychoacoustical and perceptually relevant dimensions of sounds. Off-the-shelf tools such as Quietyme, SonicU, Pulsar Safe Ear, and Sound Ear are particularly relevant since, as we will discuss below, they target the healthcare sector and have been applied to the context of NICU. Both Quietyme and SonicU provide integrated solutions that rely on permanent installations of networked sensors and store data for analysis, forensic investigations of incidents and correlation with other information sources over time. Reports with historical data on average noise levels and peak noise events over a certain time are regularly sent to customers to support a deeper awareness of the acoustic environment and decision-making processes. Both products integrate different data sources, such as temperature and air quality. In an effort to get closer to the human perception of sound and characterize noise level in terms of the effects it has on people, Pulsar Safe Ear and Sound Ear provide their customers with an artifact: a physical display that is permanently installed on the premises. In the case of Sound Ear, the display is ear-shaped to attract the attention of users, both listeners and producers of sound events, on the perceptual effects of noise levels. The display uses a simple visual cue to communicate information in real-time: similarly, to a traffic light, the color changes from green ("good" noise level) to red ("critical," "harmful" noise level). The exact dB measurement, represented by a number, is also displayed.

Cases such as MosART, Hush City and the Harmonica and Mouvie projects deserve special attention as their aim is to provide users with meaningful information on the acoustic environment as mediated by human perception. MosART (Mobile Soundscape Appraisal and Recording Technology) is a smartphone application prototyped in the context of the MosART+ intervention (Kosters et al., 2022), a research and commercial endeavor that aims "to increase auditory awareness in healthcare professionals, to research the experience of music festivals by residents and visitors, and to study possible effects on sound annoyance of different constructions techniques" (Soundappraisal website., 2022). The MosART app prompts users to sample their acoustic environment collecting short recordings then labeled according to perceptual qualitative characteristics such as calm, boring, chaotic, and lively. It is also possible to label sound sources that the user identifies while recording. It is particularly relevant to this review that the main goal of the project is to increase awareness of nursing staff on the impact of the acoustic environment on healthcare professionals and, conversely, the impact of our own actions as sound producers. The increased awareness acts as a facilitator to support more informed institutional decision-making and individual action-taking to contrast the negative impact of noise in the healthcare space (Kosters et al., 2022). Hush City (Radicchi, 2021) also proposes a smartphone app that citizens can use to record, geotag, and share on an open web map quiet places in their cities. It is a

 10.3389/fcomp.2023.1156693

solution framed within the citizen science movement that wants to empower people in collectively assessing the quietness of the urban soundscape "with the potential of orientating plans and policies for healthier living" (Hushcity website., 2022). While it shares some characteristics with the Noise Capture app described above (it provides the user with real-time dB measurements), it also prompts users to make a subjective judgement on the perceived quietness of the place thus interpreting a physical dimension such as SPL through the lenses of human listening in context.

The Mouvie (Mobility and Quality of Life in Urban Areas) project ran from 2014 to 2019 to develop new metrics for the assessment of urban noise generated from vehicles traffic (Misdariis et al., 2019). Researchers explicitly aimed to move from a "normative" approach to the evaluation of the urban acoustic environment, based on "objective measurements, sound level thresholds and operational solutions" to a "sensitive" approach based on subjective metrics that consider the psychoacoustical, cognitive, and social dimensions of sound events (Misdariis et al., 2019, p. 2). In an experimental validation, several psychoacoustic metrics along with automatic sound event classification are used in combination with listeners annotations to characterize the urban acoustic space (notably, traffic sounds) in terms of its annoyance level in relation to potential for action. The combined metric forms the basis for the definition of a computational method to automatically evaluate the annoyance level of traffic noise and its impact on human activity in urban space. Finally, the Harmonica Index developed in the context of the Harmonica Project (Mietlicki et al., 2014) defines a novel approach to noise measurements that interprets and displays the physical features of sound (such as average and peak SPL) in a way that is closer to what people perceive. The Index combines two existing metrics, the background noise level, and the peak noise level (that refers to salient sound events such as aircrafts, rail traffic, trucks passing by). The quantitative information extracted from these two metrics is presented to users on a simple 0 to 10 scale, thus favoring the interpretation by non-experts both at citizen and institutional level. This representation, explicitly inspired by current consumer solutions for air quality monitoring, provides listeners with information on the "real feeling" of sound (analogous to the real feeling of temperature provided by weather forecasts), rather than analytical, accurate information on SPL. Thanks to an engaging data visualization display, the Harmonica Index wants to increase awareness on the role of different sound sources and their impact at different times and in different locations in the urban context.

In all solutions, Characterization seems to have been achieved by making the perceptual quality of the sounds explicit for the non-expert user by combining quantitative data on sound pressure levels with information on how it affects listeners i.e., how annoying, pleasant, quiet, harmful, or peaceful the soundscape can be. Thus, an overall quality assessment is the result of these sound evaluation methods.

Solutions that aim to Interpret the acoustic environment replicating the cognitive process of the human ear can be clustered around two main areas. DYNAMAP (Sevillano et al., 2016), Noise Modeling (Le Bescond et al., 2021), BirdVox (Salamon et al., 2016), and SONYC (Dove et al., 2022) rely on low-cost sensor networks already deployed in cities to collect sound data later used to train ML algorithms for sound events classification, with the ultimate goal of designing reliable models of the urban and natural soundscapes. Except for BirdVox, which is targeted to the identification of bird calls for scientific and ecological purposes, all other solutions address, once again, the issue of urban noise and the harmful impact it can have on public health by providing advanced real-time monitoring systems and noise mapping. All solutions are academic research projects at various degrees of validation, in general characterized by a lower TRL (between 4 and 5). They are all designed for the outdoor acoustic environment, which highlights a lack of research in sound events detection and source classification in indoor spaces. DYNAMAP's goal is to develop low-cost solutions for the real-time update of noise maps. Its algorithm builds a "sound layer" on general purpose Geographical Information System (GIS) platforms (used to manage geographical data for research, land management, and urban planning). According to the authors, this strategy reduces the need of public institutions to rely on expensive dedicated acoustic software and hardware (Sevillano et al., 2016) to produce mandatory municipal noise maps, while at the same time it provides a more accurate description of the impact of urban noise (mainly, traffic sounds). Similarly, SONYC (Sounds of New York City) leverages existing low-cost sensors within the city of New York to understand the urban soundscape in terms of noise pollution. Machine Listening algorithms were developed (Salamon and Bello, 2017) to predict noise levels and identify patterns and outliers in the propagation of urban noise from specific sound sources. As an additional layer, SONYC also promotes citizen participation for the collection of data. Samples of urban noise recordings can be collected by individuals and added to the database. BirdVox applies ML and automatic sound recognition to the natural soundscape with the goal of cataloging free-flying birds calls. The algorithm will be deployed in natural environments in conjunction with audio sensor networks for the monitoring and interpretation of birds' migrations paths to support and promote actions for the protection of avian species. Two prototypal solutions, Doplor Sleep and SOUNDscapes are of relevance to this study as they are designed for healthcare and, in one case, NICU. Doplor Sleep addresses the issue of sound-induced sleep disturbance in hospitals. Through a smartphone app, the system captures sound events to visualize, through a friendly and attractive interface that targets non-expert users, critical information such as sound levels but also classification of sound sources (alarm, speech, incidental sounds, or snore) for increased awareness on sleep disturbances (Özcan et al., 2022c). Doplor Sleep also has a nurse interface and displays the analysis of the sound events occurring at night. Both the patient and nurse solutions also use characterization method for displaying the acoustical quality of the nighttime sound environment. SOUNDscapes is a digital platform that detects, localizes and classifies sound events occurring at the NICU. Data is then displayed to inform nurses on real-time sound levels, trends (e.g., during day or nighttime), type of sound sources and their localization (Spagnol et al., 2023), to provide healthcare professionals and, over time, hospital management, with specific knowledge to address the issue of noise in NICU. Like Doplor Sleep, SOUNDscapes uses a metaphorical and visual description to provide the listener with an overall evaluation of the quality of the acoustic environment.

The Interpretation of the acoustic environment is achieved by combining data and information from all previous stages. The Measurement of quantitative metrics i.e., sound levels, and the Characterization of perceptual characteristics such as annoyance and pleasure are combined with information on the footprint (i.e., the impact on the overall soundscape's quality) that different sound events have. The goal of these solutions is to present the user with a comprehensive understanding of the acoustic environment so that they can take action to improve it, both by individually changing their behavior (e.g., producing less "noise") and working toward better management of the space and work activity. Thus, users have access to measurement data, its perceptual characterization, and the footprint of the different sound categories.

We see **Forecast** solutions as guided by the ambition of providing human users with augmented capabilities to make-sense and predict the behavior of acoustic environments at a previously unimaginable scale. In this sense, these solutions will be "empowered hearing systems" that, while they face "the same challenge that biological hearing systems have evolved to solve-to make sense of sound and thereby infer the state of the world" (Sharan and Moir, 2016), aspire to overcome the limits of human listening and cognition. Ultimately, these solutions could be used to "answer specific biological, ecological, and management" (Bianco et al., 2019) questions. In the NICU context, these questions might include: How does sound influence the sleeping pattern of premature babies? How do decisions on personnel shifts (e.g., when nurses change shift, when food is delivered, when cleaning service is administered) impact the sound quality of a NICU unit? How does the design of the rooms impact the perception of different sound sources? The capability to answer these questions would provide both individuals and institutions with usable knowledge to take action toward a healthier sound environment in critical care, both in the short and in the long term.

Along the course of this study, we identified several research efforts that move toward this direction exploring applications of ML techniques (for a recent review see Bianco et al., 2019) and the development of new acoustic, psychoacoustic and bioacoustics computational models (Brocolini et al., 2012; Sueur et al., 2014; Alsina-Pagès et al., 2021). Although none of these studies focus on indoor environments, we include them in this review as they constitute a possible scenario for a solution that can forecast the behavior of the NICU environment, as described above. In particular, here we consider a case that-albeit sharing some of the characteristics of embryonic research and low TRL (3)-has been empirically validated as an integrated solution, rather than an isolated set of novel indexes. The study by Pedersen et al. (2018, 2021) leverages ML to predict ambient sound levels across the United States. Sound samples from more than 600 locations were correlated with GIS measurements from more than 100 sites in the U.S. to train a model to predict the impact on the acoustic environment of changes in land management, such as the introduction or the removal of an airport or a high-traffic road. This project is of particular interest to this review as the authors acknowledge the importance of accurate soundscape modeling for "public health studies and urban development," potential "commercial applications for real estate and urban development" and "implications for social justice" (Pedersen et al., 2018, p. 2). In the context of NICU soundscape, a similar algorithm that can forecast the impact of different auditory footprints and display it for nurses would support both short-term increased awareness and longer-term improved decision-making toward a healthier acoustic environment in critical care.

As part of the next step in our analysis, we categorized the solutions as a function of their Technology Readiness Levels (TRLs). The definition of the TRL is based on the official description adopted by the European Union and can be found here https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf. The TRL analysis by Function (Figure 5) shows that solutions in the Measure and Describe function are characterized by a higher TRL (between seven and nine). As we move toward Interpret and Forecast, the TRL level decreases. Specifically, the TRL level of solutions that Interpret the acoustic environment starts at two, with a maximum of five points, while the TRL level of the solution representative of the Forecast function is three.

As shown in Figure 6, of the 27 cases considered in this review 12 are solutions for the evaluation of the outdoor soundscape while only eight are presented as indoor solutions. Eight are described as applicable to both outdoor and indoor contexts. This might be due to several factors including critical technical aspects (e.g., the relative lack of variety of sound sources and the higher complexity of the acoustic space) that are made more complex by the difficulty to access and collect the needed data in indoor spaces due to privacy concerns. For instance, in the context of NICU, the limitations imposed by privacy rules in an extremely sensitive environment make the development of algorithms that can identify and categories different sound sources difficult to achieve. Additionally, social aspects such as the citizens' perception and public discourse around noise facilitate the advancement of research in the context of urban environment while leaving the indoor soundscape relatively under-investigated.

Our findings highlight a clear gap in the research, development, and commercialization of solutions that address the needs of indoor environments. However, functional socio-technological environments such as control rooms, workspaces, healthcare environments, where sound is conducive to social, environmental, and instrumental interactions (Özcan et al., 2022a) greatly impact the everyday life of people and therefore, innovation in this area is urgently needed. Critical care environments, and especially the NICU, represent a unique case of a particularly self-contained functional environment with peculiar design characteristics that greatly differentiate it from other indoor environments, notably the strict control of interchanges with the external world. In fact, only a handful of solutions among those we consider in this study are applied to healthcare environments (Figure 7). The majority belongs to the Measure and Characterize functions and are marketed by manufacturers as off-the-shelf solutions to monitor noise levels. Some of them (SonicU, Quietyme) offer integrated environmental monitoring of critical environments that couple the measurement of sound levels with temperature and air quality. Others, like McLennan Sound Monitoring and SoundIntel provide healthcare personnel with real-time alerts of unusual sound events that might represent a threat to patients' safety. However, the definition of what constitutes a "threat" is based solely on the

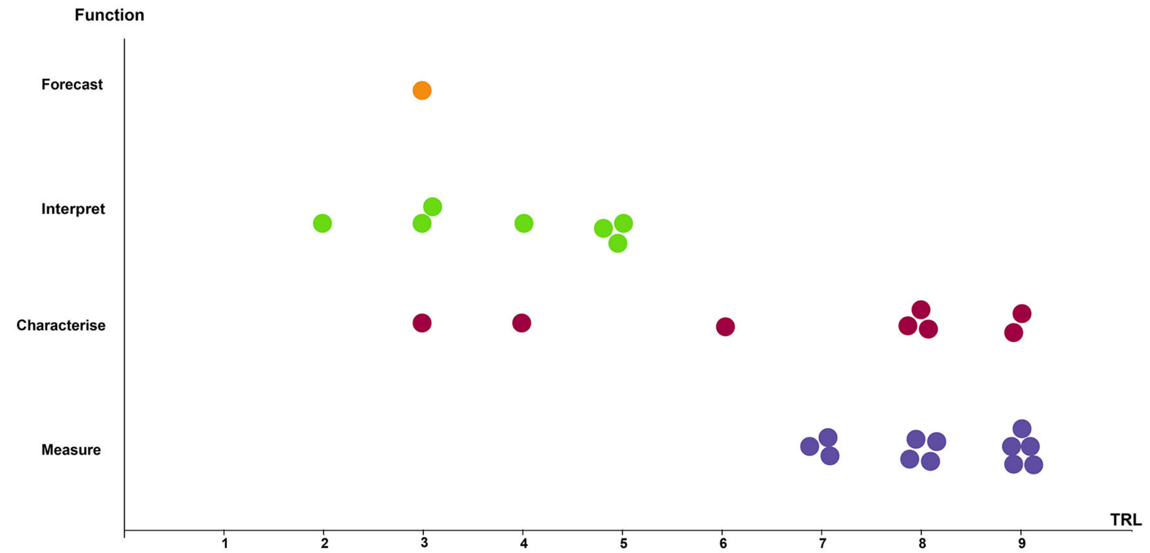

FIGURE 5
Technology readiness levels by function.

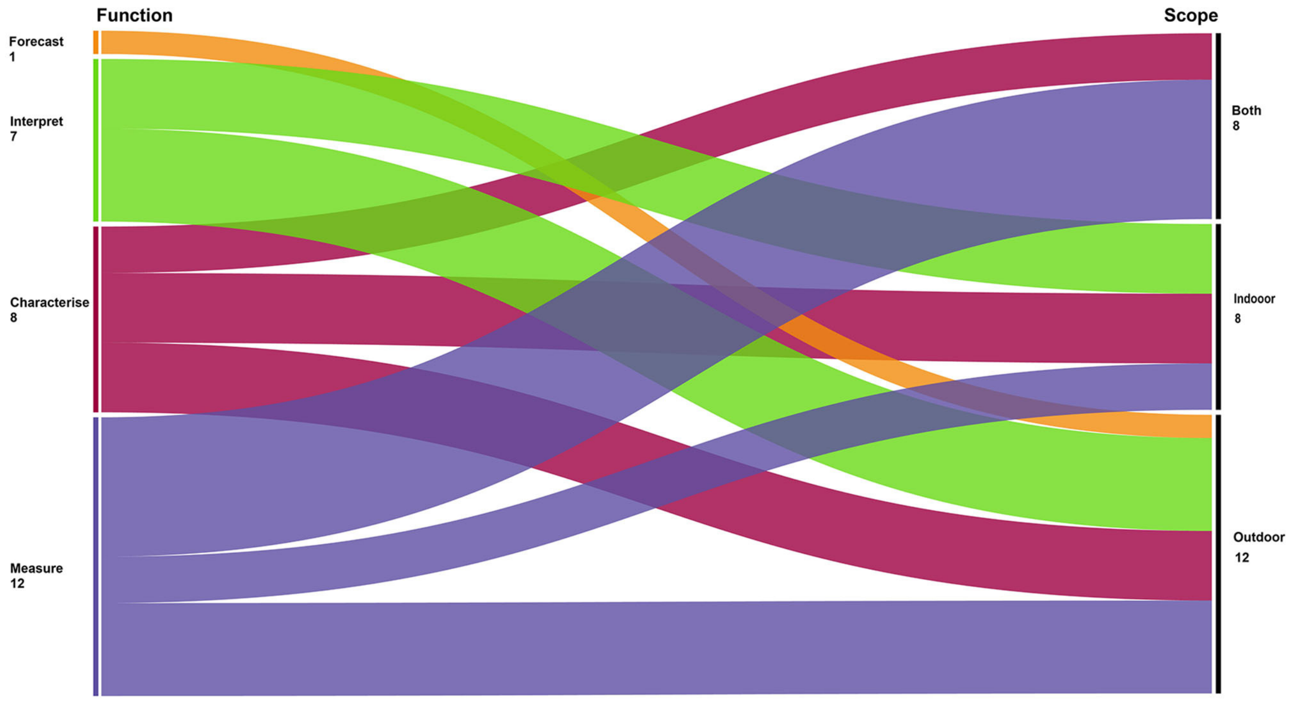

FIGURE 6
Correlation between function and use outdoor or indoor of the solutions.

detection of sound levels (e.g., the sudden peak in environmental sound level caused by a body falling) that does not provide any insight into the cause of the event. Among the solutions considered in this study, only MosART+ (described above, Kosters et al., 2022) engages nursing staff in the qualitative categorization of sound events in the hospital environment by means of a dedicated mobile application. While this system collects data on the perceived qualities of different sound sources, it does not automatically inform users on the footprint of hospital sounds so that they can take action to improve it.

Of the five solutions that explicitly target NICU (sometimes along with other critical and intensive care environments such as PICU and ICU), the only interpretative solution (in green in Figure 6), SOUNDscapes, is currently in the state of concept design

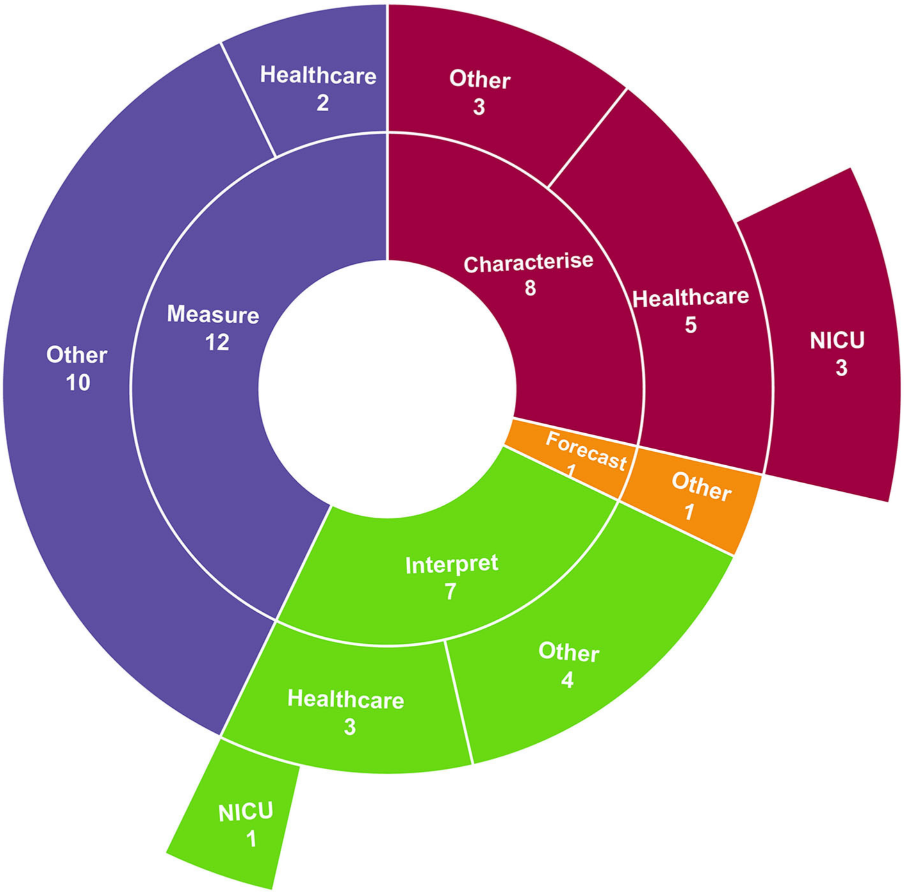

FIGURE 7
Application to the healthcare and NICU context of analyzed solutions, by function.

(i.e., it has not been prototyped yet). Only three solutions-Sound Ear, Pulsar Safe Ear, and SonicU-are commercialized to be applied in NICU. This finding highlights the smaller market represented by NICU as compared to healthcare in general (NICU solutions are <50% of all solutions for healthcare). NICUs, as reminded on various occasions in this article, are extremely delicate acoustic environments where patients are particularly vulnerable to sounds and where extra care is required when it comes to the introduction of new technology and the collection of soundscape data. At the same time, patients of NICU are unique in that they are exposed to the negative effect of unwanted sounds, but they rarely have a role in producing sounds that can negatively impact the acoustic environment (as premature babies are in general extremely quiet and cannot express verbally). For the very same reasons, we see NICU as an opportunity for the design of novel solutions that want to move beyond a descriptive approach to support a more informed sense-making of the sound quality. These novel solutions should, as Doplor Sleep and SOUNDscapes, assess the impact that sound events have on the quality of the NICU soundscape, and provide nurses with actionable information so that they can actively contribute to increase it. Solutions that prove to be efficient in the NICU could be scalable to other critical care contexts such as PICU and adult ICU.

4. Conclusions

Our semi-systematic review of more than 70 scientific publications on the acoustic environment of the NICU shows that there is an urgent need for solutions that can provide nurses with comprehensive, holistic information on how different categories of sounds (e.g., speech, medical alarms, machinery) impact the quality of the soundscape. This information should include quantitative measurements of sound levels but also increase awareness of how qualitative characteristics of sound-which are subjective and highly context-dependent (Axelsson, 2015)-are perceived (e.g., annoying, pleasant, chaotic, calm).

The review of technological solutions highlights that existing products marketed for critical care tend to monitor the acoustic environment by measuring the overall sound level episodically and in real time. While this measurement strategy complies with privacy regulations that limit the collection and identification of personal data (such as data contained in speech), it greatly limits the possibilities for nurses (i.e., the guardians of NICU patients and their wellbeing) of understanding how different sound sources impact the perceived quality of critical care soundscape and consequently, the possibility to take immediate action to improve it. Current research that leverages low-cost and networked acoustic sensors to collect long-term data and algorithmically model the soundscape characteristics only focuses on outdoor spaces, where such networks exist as part of a public effort in mitigating noise pollution. Additionally, existing protocols to assess the affective quality of soundscapes were designed for the outdoor context and their applicability for indoors is still debated (Torresin et al., 2020). Consequently, we see a clear opportunity for the design and development of new solutions that focus on the collection of rich soundscape data in the NICU environment to allow for *quantitative* (i.e., based on physical properties of sound), *qualitative* (i.e., based on the perceived characteristics), and *categorical* (i.e., based on the sources of different sound events) analysis. Such rich data should be used to extract the appropriate soundscape descriptors (Aletta et al., 2016) to model the NICU soundscape and develop novel algorithms to characterize, interpret, and predict its behavior over time in a holistic perspective. Particular attention should be given to the design of how nurses interact with the soundscape data, i.e., how data are translated into usable information that becomes actionable knowledge (Masud et al., 2010). A technological solution that integrates these algorithms with a user-centered display of information on the auditory footprint and quality of sound events in NICU, would provide nurses with actionable information to actively improve the soundscape. However, critical ethical and privacy concerns are faced to collect the appropriate soundscape data from the indoor environment, especially in the context of critical care. These concerns are recognized by providers of commercial solutions for audio data collection for AI (Javahid, 2023) as well as by researchers (Nautsch et al., 2019). The exploration of automatic data cancellation (for instance regarding speech) and technological solutions such as on edge data collection where acquisition and processing happen on the device is necessary although beyond the scope of this article.

4.1. Design implications

We commenced this review study not only to assess the state-of-the-art of the measurement of the acoustic quality of NICU, but also to understand what characteristics future solutions should present. Awareness that current acoustic measurement systems are neither facilitating radical change in the quality of the auditory experience in NICU, nor preventing health and cognitive risks connected with the exposure to dangerous noise levels is widespread. Therefore, with our accumulated knowledge and insights, we reflected on the critical elements of a system that aims to reduce the negative impact of sound in NICU in the long-term while raising awareness to the root causes and magnitude of the impact generated. Ultimately, a combination of longer-term data collection, the development of new soundscape indices and ML techniques for the quantitative and qualitative evaluation of the sound environment, and a human-centered design of informative displays seems a promising combination of design factors for solutions that can help critical care professionals to drastically improve the quality of the NICU acoustic environment.

Figure 8 summarizes our findings and consequently highlights a vision for a future solution that integrates several main functions and provides a holistic analysis of the NICU soundscape with the aim to better inform nurses with respect to its perceived qualities and the footprint of specific sound categories. The Figure highlights the past and established efforts with sound measurement methods but also a future vision for a system that can model soundscapes and forecast the perceptual impact of sound events. However, the intermediate level of solutions (captured in blue background in the Figure) provides the sweet spot for state-of-the-art data-centric design for soundscape improvement as the current knowledge and technologies would be suitable to characterize the perceptual qualities of the sound and interpret the acoustic environment in terms of its sound sources. Yet, the solutions we have found that fit the NICU and would incorporate these functions only represent conceptual designs with low TRLs. The NICU context urgently needs an industrial and scientific effort to support their sonic needs.

A sound quality assessment solution for NICU should answer the following questions:

- What is the relationship between the physical and the perceived properties of sound, and the footprint of different sound categories in the NICU soundscape?
- What are the appropriate indicators to algorithmically model the relationship between the physical (quantitative) and the perceived (qualitative) footprint of different sound categories of the NICU soundscape?
- How does a technological solution for nurses holistically represent the quantitative, qualitative and categorical properties of the NICU soundscape in a context-relevant, human-centered way?

Ideally, a sound quality evaluation system fit for NICUs could have the following functions. First, a general scan of the acoustic environment could result in traditional and well accepted measurement and display of sound levels. This stage is also essential to collect the sound input properly and carefully to be used by more advanced functionalities (Characterize, Interpret, and Forecast). Purposely developed sound quality indices would further complement the sound measurements and build toward a holistic evaluation of the NICU soundscape for nurses to make quick judgements and take immediate action to improve it. Such an evaluation would imply the development of new soundscape descriptors and indicators based on affective qualities of sound (e.g., pleasantness, annoyance, eventfulness) and modeled on the indoor NICU environment. The automatic assessment of sound quality would represent a great advancement in the study of indoor soundscapes and could help listeners further train their listening skills and be aware of the perceptual impact of sound. The Interpret function will provide nurses with insights into the root causes of the sound quality by making sound sources (e.g., speech, alarms, support devices) and their footprint over time explainable. Finally, the Forecast function will provide a holistic overview of the NICU soundscape based on big, trained data for any given time or situation and will allow for predictions and early diagnosis for possible threats to the quality of the acoustic environment. Not only nurses but most likely unit managers and hospital technology scouts will make use of this function to make well-informed choices for structural change (e.g., purchasing decisions,

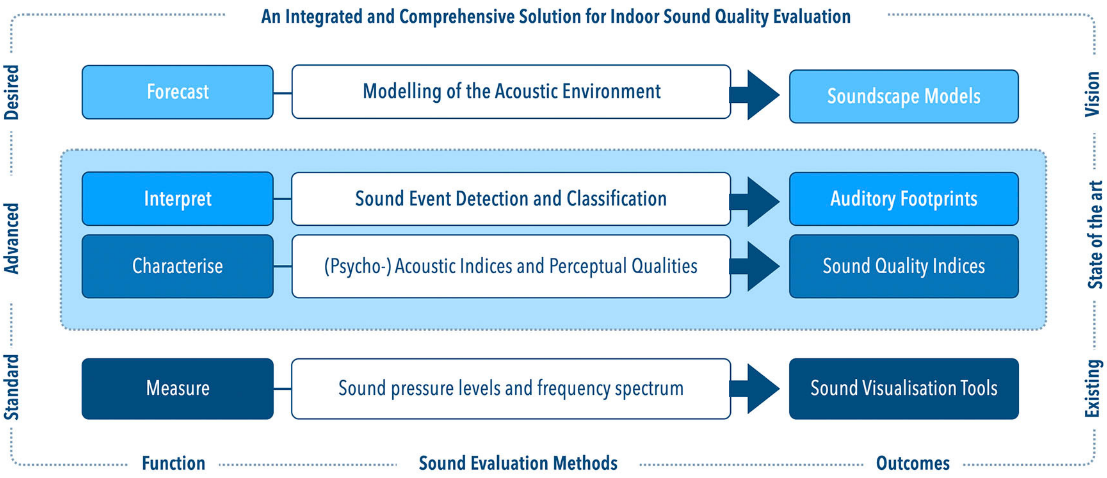

FIGURE 8
The four functions of technological solutions for sound quality evaluation and the outcome they generate for the user.

workflow analysis that can cause unwanted noise). Accordingly, all these functions will augment the listener's perceptual, cognitive, and affective skills enhancing their sense making of the NICU soundscape.

The future *holistic* solution envisaged in Figure 8 will build on a quantitative, qualitative, and categorical approach to sound analysis with the goal to computationally model "the relationship between the physical and the perceived properties of the acoustic environment" (Aletta et al., 2016, p. 68) in the context of the NICU indoor soundscape. The *quantitative* analysis would be based on audio signal processing and psychoacoustics and inform the conventional metrics which represent the prior art sound measurements. The *qualitative* analysis would be based on perceived affective quality and inform the modeling of the descriptive qualities of the NICU soundscape. The *categorical* analysis would be based on the classification of sound sources and their context relevance and inform the computational modeling of classification of sound events and their appropriateness. The combination of these three approaches would provide nurses with a comprehensive understanding of the NICU soundscape thus facilitating its active improvement both in the short term–through behavior changes, and long term–by supporting informed decisions on the organization of NICU activities and design of its infrastructure.

Author contributions

SL conceived the review study of technological solutions and wrote the Section 3. SS and EÖ conceived the semi-systematic literature review and wrote the Section 2. SL, SS, and EÖ equally contributed to Sections 1 and 4, Figures, and Tables. All authors approved the submitted version.

Funding

This study was conducted with fundings from Delft University of Technology, Faculty of Industrial Design Engineering, Department of Human-Centered Design and Erasmus Medical Center Rotterdam.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomp.2023.1156693/full#supplementary-material

SUPPLEMENTARY MATERIAL 1
List of the technological solutions analyzed, with metadata.

References

Ahamed, M. F., Campbell, D., Horan, S., and Rosen, O. (2018). Noise reduction in the neonatal intensive care unit: a quality improvement initiative. *Am. J. Med. Qual.* 33, 177–184. doi: 10.1177/1062860617711563

Aita, M., Robins, S., Charbonneau, L., Doray-Demers, P., and Feeley, N. (2021). Comparing light and noise levels before and after a NICU change of design. *J. Perinatol.* 41, 2235–2243. doi: 10.1038/s41372-021-01007-8

Aletta, F., Kang, J., and Axelsson, Ö. (2016). Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landsc. Urban Plan.* 149, 65–74. doi: 10.1016/j.landurbplan.2016.02.001

Alsina-Pagès, R. M., Benocci, R., Brambilla, G., and Zambon, G. (2021). Methods for noise event detection and assessment of the sonic environment by the harmonica index. *Appl. Sci.* 11:8031. doi: 10.3390/app11178031

Axelsson, Ö. (2015). "How to measure soundscape quality," in *Proceedings of the Euronoise* 2015 *conference*, 1477–1481.

Axelsson, Ö., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–2846. doi: 10.1121/1.3493436

Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., et al. (2019). Machine learning in acoustics: theory and applications. *J. Acoust. Soc. Am.* 146, 3590–3628. doi: 10.1121/1.5133944

Blackburn, S. (1998). Environmental impact of the NICU on developmental outcomes. *J. Pediatr. Nurs.* 13, 279–289. doi: 10.1016/S0882-5963(98)80013-4

Bliefnick, J. M., Ryherd, E. E., and Jackson, R. (2019). Evaluating hospital soundscapes to improve patient experience. *J. Acoust. Soc. Am.*, 145, 1117. doi: 10.1121/1.5090493

Brandon, D. H., Ryan, D. J., and Barnes, A. H. (2007). Effect of environmental changes on noise in the NICU. *Neonatal Netw.* 26, 213–218. doi: 10.1891/0730-0832.26.4.213

Braun, V., and Clarke, V. (2012). "Thematic analysis," in *APA Handbook of Research Methods in Psychology, Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological,* eds H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, and K. J. Sher (American Psychological Association), 57–71.

Brocolini, L., Lavandier, C., Marquis-Favre, C., Quoy, M., and Lavandier, M. (2012). *Prediction and Explanation of Sound Quality Indicators by Multiple Linear Regressions and Artificial Neural Networks.* Nantes: Acoustics, 2121–2126.

Busch-Vishniac, I., and Ryherd, E. (2019). Hospital soundscapes: characterization, impacts, and interventions. *Acoust. Today*, 15, 11–18. doi: 10.1121/AT.2019.15.3.11

Busch-Vishniac, I. J., West, J. E., Barnhill, C., Hunter, T., Orellana, D., and Chivukula, R. (2005). Noise levels in Johns Hopkins Hospital. *J. Acoust. Soc. Am.*, 118, 3629–3645. doi: 10.1121/1.2118327

Calikusu Incekar, M., and Balci, S. (2017). The effect of training on noise reduction in neonatal intensive care units. *J. Spec. Pediatr. Nurs.* 22:e12181. doi: 10.1111/jspn.12181

Carvalhais, C., Santos, J., da Silva, M. V., and Xavier, A. (2015). Is there sufficient training of health care staff on noise reduction in neonatal intensive care units? a pilot study from Neonoise project. *J. Toxicol. Environ. Health*, 78, 897–903. doi: 10.1080/15287394.2015.1051204

Casey, L., Fucile, S., Flavin, M., and Dow, K. (2020). A two-pronged approach to reduce noise levels in the neonatal intensive care unit. *Early Hum. Dev.* 146, 105073. doi: 10.1016/j.earlhumdev.2020.105073

Chen, H. L., Chen, C. H., Wu, C. C., Huang, H. J., Wang, T. M., and Hsu, C. C. (2009). The influence of neonatal intensive care unit design on sound level. *Pediatr. Neonatol.*, 50, 270–274. doi: 10.1016/S.1875-9572(09)60076-0

COST TUD Action (2013). *TD-804 March 22, Publications. (J. Kang, K. Chourmouziadou, K. Sakantamis, B. Wang, and Y. Hao, Eds.) from COST TUD Action TD-0804 Soundscape of European Cities and Landscapes.* Available online at: http://soundscape-cost.org/index.php?option=com_contentandview=articleandid=62:final-e-book-publishedandcatid=32:publicationsandItemid=9 (accessed April 15, 2016).

Darcy, A. E., Hancock, L. E., and Ware, E. J. (2008). A descriptive study of noise in the neonatal intensive care unit. ambient levels and perceptions of contributing factors. *Adv. Neonatal Care*, 8, 165–175. doi: 10.1097/01.ANC.0000324341.24841.6e

De Coensel, B., and Botteldooren, D. (2014). "Smart sound monitoring for sound event detection and characterisation," in *Proceedings of the INTERNOISE* (Melbourne).

De Coensel, B., Sun, K., Wei, W., Van Renterghem, T., Sineau, M., Ribeiro, C., et al. (2015). "Dynamic noise mapping based on fixed and mobile sound measurements," in *Proceedings of the Euronoise* (Maastrich), 2339–2344.

de Lima Andrade, E., da Cunha E Silva, D. C., de Lima, E. A., de Oliveira, R. A., Zannin, P., and Martins, A. (2021). Environmental noise in hospitals: a systematic review. *Environ. Sci. Pollut. Res. Int.*, 28, 19629–19642. doi: 10.1007/s11356-021-13211-2

Dove, G., Mydlarz, C., Bello, J. P., and Oded, N. (2022). Sounds of New York city. *Interactions* 29, 32–35. doi: 10.1145/3527726

Duran, R., Ciftdemir, N. A., Ozbek, U. V., Berberoglu, U., Duranku,ş, F., Süt, N., et al. (2012). The effects of noise reduction by earmuffs on the physiologic and behavioral responses in very low birth weight preterm infants. *Int. J. Pediatr. Otorhinolaryngol.* 76, 1490–1493. doi: 10.1016/j.ijporl.2012.07.001

Elander, G., and Hellström, G. (1995). Reduction of noise levels in intensive care units for infants: evaluation of an intervention program. *Heart Lung*, 24, 376–379. doi: 10.1016/S0147-9563(05)80058-5

European Parliament. (2000). *Directive 2000/14/EC of the European Parliament and of the Council of 8 May 2000 on the Approximation of the Laws of the Member States Relating to the Noise Emission in the Environment by Equipment for Use Outdoors.* Available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32000L0014

Graziuso, G., Francavilla, A. B., Mancini, S., and Guarnaccia, C. (2022). "Application of the harmonica Index for noise assessment in different spatial contexts," in *Journal of Physics: Conference Series, Volume 2162, 5th. International Conference on Applied Physics, Simulation and Computing (APSAC 2021)* (Salerno). doi: 10.1088/1742-6596/2162/1/012006

Gupta, A., Gupta, A., Jain, K., and Gupta, S. (2018). Noise pollution and impact on children health. *Indian J. Pediatr.*, 85, 300–306. doi: 10.1007/s12098-017-2579-7

Hellström, B., Nilsson, M. E., Axelsson, Ö., and Lundén, P. (2014). Acoustic design artifacts and methods for urban soundscapes: a case study on the qualitative dimensions of sounds. *J. Archit. Plann. Res.* 31, 57–71. Available online at: http://www.jstor.org/stable/43031024

Henze, H., Moshona, C., and Fiebig, A. (2023). "Katalogisierung von Lärmschutzmaßnahmen mit Gestaltungsaspekten," in *Fortschritte der Akustik – DAGA 2023, 49. Jahrestagung für Akustik, Deutsche Gesellschaft für Akustik e.V.* (DEGA).

Hernández-Salazar, A. D., Gallegos-Martínez, J., and Reyes-Hernández, J. (2020). Level and noise sources in the neonatal intensive care unit of a reference hospital. *Invest. Educac. Enferm.* 38:e13. doi: 10.17533/udea.iee.v38n3e13

Hushcity website. (2022). Available online at: https://map.opensourcesoundscapes.org (accessed January 10, 2023).

ISO 12913-1. (2014). *Acoustics—Soundscape—Part 1: Definition and Conceptual Framework.* Geneva: International Organization for Standardization.

Javahid, S. (2023). *Audio Data Collection for AI: Challenges and Best Practices in 2023.* Available online at: https://research.aimultiple.com/audio-data-collection/ (accessed January 27, 2023).

Johnson, K. R., Hagadorn, J. I., and Sink, D. W. (2017). Alarm safety and alarm fatigue. *Clin. Perinatol.* 44, 713–728. doi: 10.1016/j.clp.2017.05.005

Kang, J., Aletta, F., Gjestland, T. T., Brown, L. A., Botteldooren, D., Schulte-Fortkamp, B., et al. (2016). Ten questions on the soundscapes of the built environment. *Build. Environ.* 108, 284–294. doi: 10.1016/j.buildenv.2016.08.011

Konkani, A., Oakley, B., and Penprase, B. (2014). Reducing hospital ICU noise: a behavior-based approach. *J. Healthc. Eng.* 5, 229–246. doi: 10.1260/2040-2295.5.2.229

Kosters, J., Janus, S. I. M., Van Den Bosch, K. A., Zuidema, S., Luijendijk, H. J., and Andringa, T. C. (2022). Soundscape optimization in nursing homes through raising awareness in nursing staff with MoSART. *Front. Psychol.* 13:871647. doi: 10.3389/fpsyg.2022.871647

Krueger, C., and Schue, S., and Parker, L. (2007). Neonatal intensive care unit sound levels before and after structural reconstruction. *MCN Am. J. Matern. Child Nurs.* 32, 358–362. doi: 10.1097/01.NMC.0000298131.55032.76

Krueger, C., Wall, S., Parker, L., and Nealis, R. (2005). Elevated sound levels within a busy NICU. *Neonatal Netw.* 24, 33–37. doi: 10.1891/0730-0832.24.6.33

Lahav, A. (2015). Questionable sound exposure outside of the womb: frequency analysis of environmental noise in the neonatal intensive care unit. *Acta. Paediatr.* 104, e14–e19. doi: 10.1111/apa.12816

Lasky, R. E., and Williams, A. L. (2009). Noise and light exposures for extremely low birth weight newborns during their stay in the neonatal intensive care unit. *Pediatrics*, 123, 540–546. doi: 10.1542/peds.2007-3418

Lawson, K., Daum, C., and Turkewitz, G. (1977). Environmental characteristics of a neonatal intensive-care unit. *Child Dev.* 48, 1633–1639. doi: 10.2307/1128527

Le Bescond, V., Can, A., Aumond, P., and Gastineau, P. (2021). Open-source modeling chain for the dynamic assessment of road traffic noise exposure. *Transp. Res. D. Transp. Environ.* 94:102793. doi: 10.1016/j.trd.2021.102793

Lenzi, S., Lindborg, P. M., Han, N., Spagnol, S., Kamphuis, D., and Özcan, E. (2023). "Disturbed sleep: estimating night-time sound annoyance at a hospital ward," in *Proceedings of Forum Acusticum 2023* (Torino).

Levy, G. D., Woolston, D. J., and Browne, J. V. (2003). Mean noise amounts in level II vs level III neonatal intensive care units. *Neonatal Netw.* 22, 33–38. doi: 10.1891/0730-0832.22.2.33

Lionello, M., Aletta, F., and Kang, J. (2020). A systematic review of prediction models for the experience of urban soundscapes. *Appl. Acoust. Volume* 170:107479. doi: 10.1016/j.apacoust.2020.107479

Liu, W. F. (2010). The impact of a noise reduction quality improvement project upon sound levels in the open-unit-design neonatal intensive care unit. *J. Perinatol.* 30, 489–496. doi: 10.1038/jp.2009.188

Liu, W. F. (2012). Comparing sound measurements in the single-family room with open-unit design neonatal intensive care unit: the impact of equipment noise. *J. Perinatol.* 32, 368–373. doi: 10.1038/jp.2011.103

Livera, M. D., Priya, B., Ramesh, A., Suman Rao, P. N., Srilakshmi, V., Nagapoornima, M., et al. (2008). Spectral analysis of noise in the neonatal intensive care unit. *Indian J. Pediatr.*, 75, 217–222. doi: 10.1007/s12098-008-0048-z

MacKenzie, D., and Galbrun, L. (2007). Noise levels and noise sources in acute care hospital wards. B*uild. Serv. Eng. Res. Technol.* 28, 117–131. doi: 10.1177/0143624406074468

Masud, L., Valsecchi, F., Ciuccarelli, P., Ricci, D., and Caviglia, G. (2010). From data to knowledge: visualizations as transformation processes within the data-information-knowledge continuum. *In Proc. Int. Conf. Inf. Vis.* 2010, 445–449. doi: 10.1109/IV.2010.68

Mietlicki, C., Mietlicki, F., Ribeiro, C., Gaudibert, P., and Vincent, B. (2014). "The HARMONICA project, new tools to assess environmental noise and better inform the public," in *Proceedings of the Forum Acusticum 2014* (Kraków).

Milette, I. (2010). Decreasing noise level in our NICU: the impact of a noise awareness educational program. *Adv. Neonatal Care*, 10, 343–351. doi: 10.1097/ANC.0b013e3181fc8108

Misdariis, N., Marignier, L., Dianoux, C., and Leiba, R. (2019). "Experimental validation of an urban traffic noise annoyance model," in *ICSV - International Congress on Sound and Vibration* (Montreal).

Mitchell, A., Oberman, T., Aletta, F., Kachlicka, M., Lionello, M., Erfanian, M., et al. (2021). (2021). Investigating urban soundscapes of the COVID-19 lockdown: a predictive soundscape modeling approach. *J. Acoust. Soc. Am.* 150, 4474. doi: 10.1121/10.0008928

Moshona, C., Aletta, F., Henze, H., Chen, X., Mitchell, A., Oberman, T., et al. (2022). "What is a soundscape intervention? exploring definitions and identification criteria and a platform to gather real-world examples," in *Internoise, 51st International Congress and Exposition on Noise Control Engineering.*

Nautsch, A., Jiménez, A., Treiber, A., Kolberg, J., Jasserand, C., Kindt, E., et al. (2019). Preserving privacy in speaker and speech characterisation. *Comput. Speech Lang.* 58, 441–480. doi: 10.1016/j.csl.2019.06.001

Nielbo, F. L., Steele, D., and Guastavino, C. (2013). Investigating soundscape affordances through activity appropriateness. *Proc. Mtgs. Acoust.* 19, 040059. doi: 10.1121/1.4800502

NIOSH website. (2022). Available online at: https://www.cdc.gov/niosh/topics/noise/app.html (accessed January 10, 2023).

Noise Aware website. (2022). Available online at: https://noiseaware.com/ (accessed January q0, 2023).

Özcan, E., Broekmeulen, C. L. H., Luck, Z. A., van Velzen, M., Stappers, P. J., and Edworthy, J. R. (2022b). Acoustic biotopes, listeners and sound-induced action: a case study of operating rooms. *Int. J. Environ. Res. Public Health.* 19:16674. doi: 10.3390/ijerph192416674

Özcan, E., Kerimoglu, E., Müler, B. J., and Spagnol, S. (2022a). *Exploring the Soundscape of the NICU: A Semi-Systematic Review With a Human-Centric Approach,* (Technical Report).

Özcan, E., Liu, Y., Vroon, J., Kamphuis, D., and Spagnol, S. (2022c). "Doplor sleep: monitoring hospital soundscapes for better sleep hygiene," in *Proceedings of the 6th International Conference on Medical and Health Informatics (ICMHI '22). Association for Computing Machinery* (New York, NY), 69–77. doi: 10.1145/3545729.3545745

Özcan, E., Rietdijk, W. J. R., and Gommers, D. (2019). Shaping critical care through sound-driven innovation: introduction, outline, and research agenda. *Intensive Care Med.* 46, 542–543. doi: 10.1007/s00134-019-05832-6

Özcan, E., and van Egmond, R. (2007). Memory for product sounds: the effect of sound and label type. *Acta Psychol.* 126, 196–215. doi: 10.1016/j.actpsy.2006.11.008

Özcan, E., and van Egmond, R. (2009). The effect of visual context on the identification of ambiguous environmental sounds. *Acta Psychol.* 131, 110–119. doi: 10.1016/j.actpsy.2009.03.007

Pedersen, K., Gee, K. L., Transtrum, M. K., and Butler, B. A. (2018). Machine learning-based prediction of outdoor ambient sound levels: ensemble averaging and feature reduction. *J. Acoust. Soc. Am.* 144, 1791–1791. doi: 10.1121/1.5067908

Pedersen, K., Transtrum, M. K., Gee, K. L., Lympany, S. V., James, M. M., and Salton, A. R. (2021). Validating two geospatial models of continental-scale environmental sound levels. *JASA Express Lett.* 1, 122401. doi: 10.1121/10.0007368

Philbin, M. K., Lickliter, R., and Graven, S. N. (2000). Sensory experience and the developing organism: a history of ideas and view to the future. *J. Perinatol.* 20, S2–S5. doi: 10.1038/sj.jp.7200434

Quinn, C. A., Burns, P., Gill, G., Baligar, S., Snyder, R. L., Salas, L., et al. (2022). Soundscape classification with convolutional neural networks reveals temporal and geographic patterns in ecoacoustic data. *Ecol. Indic.* 138:108831. doi: 10.1016/j.ecolind.2022.108831

Radicchi, A. (2021). "Citizen science mobile apps for soundscape research and public spaces studies: lessons from the Hush City project," in *Geographic Citizen Science Design: No one left behind*, eds A. Skarlatidou, and M. Haklay (UCL Press), 130–148. doi: 10.2307/j.ctv15d8174.14

Radicchi, A., Cevikayak Yelmi, P., Chung, A., Jordan, P., Stewart, S., Tsaligopoulos, A., et al. (2021). Sound and the healthy city. *Cities Health,* 5, 1–2. doi: 10.1080/23748834.2020.1821980

Ramm, K., Mannix, T., Parry, Y., and Gaffney, M. P. (2017). A comparison of sound levels in open plan versus pods in a neonatal intensive care unit. *Health Environ. Res. Design J.* 10, 30–39. doi: 10.1177/1937586716668636

Restin, T., Gaspar, M., Bassler, D., Kurtcuoglu, V., Scholkmann, F., and Haslbeck, F. B. (2021). Newborn incubators do not protect from high noise levels in the neonatal intensive care unit and are relevant noise sources by themselves. *Children*, 8:704. doi: 10.3390/children8080704

Rosenblum, L. D., Wuestefeld, A. P., and Anderson, K. L. (1996). Auditory reachability: an affordance approach to the perception of sound source distance. *Ecol. Psychol.* 8, 1–24. doi: 10.1207/s15326969eco0801_1

Salamon, J., and Bello, J. P. (2017). Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Process. Lett.* 24, 279–283. doi: 10.1109/LSP.2017.2657381

Salamon, J., Bello, J. P., Farnsworth, A., Robbins, M., Keen, S., et al. (2016). Towards the automatic classification of avian flight calls for bioacoustic monitoring. *PLoS ONE* 11:e0166866. doi: 10.1371/journal.pone.0166866

Sevillano, X., Socoró, J. C., Alías, F., Bellucci, P., Peruzzi, L., Radaelli, S., et al. (2016). DYNAMAP–development of low cost sensors networks for real time noise mapping. *Noise Mapp.* 3, 172–189. doi: 10.1515/noise-2016-0013

Sharan, R. V., and Moir, T. J. (2016). An overview of applications and advancements in automatic sound recognition. *Neurocomput.* 200, 22–34. doi: 10.1016/j.neucom.2016.03.020

Smith, S. W., Ortmann, A. J., and Clark, W. W. (2018). Noise in the neonatal intensive care unit: a new approach to examining acoustic events. *Noise Health*, 20, 121–130. doi: 10.4103/nah.NAH_53_17

Snyder, H. (2019). Literature review as a research methodology: an overview and guidelines. *J. Bus. Res.* 104, 333–339. doi: 10.1016/j.jbusres.2019.07.039

Socoró, J. C., Alías, F., and Alsina-Pagès, R. M. (2017). An anomalous noise events detector for dynamic road traffic noise mapping in real-life urban and suburban environments. *Sensors*. 17:2323. doi: 10.3390/s17102323

Soundappraisal website. (2022). Available online at: https://soundappraisal.eu/mosart.html (accessed January 10, 2023).

Spagnol, S., Viñas Vila, N., Akdag Salah, A., Goos, T. G., and Özcan, E. (2023). "SOUNDscapes: a dashboard for promoting a healthy sound environment in the Neonatal Intensive Care Unit," in *Proceedings of Forum Acusticum 2023* (Torino).

Spagnol, S., Viñas Vila, N., Akdag Salah, A., Goos, T. G., Reiss, I., and Özcan, E. (2022). "Towards a quieter neonatal intensive care unit: current approaches and design opportunities," in *Proc. 2022 Int. Conf. of the Design Research Society* (DRS 2022) (Bilbao). doi: 10.21606/drs.2022.258

Sueur, J., Farina, A., Gasc, A., Pieretti, N., and Pavoine, S. (2014). Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica united Acust.* 100, 772–781. doi: 10.3813/AAA.918757

Surenthiran, S. S., Wilbraham, K., May, J., Chant, T., Emmerson, A. J., and Newton, V. E. (2003). Noise levels within the ear and post-nasal space in neonates in intensive care. *Arch. Dis. Child. Fetal Neonatal Ed.* 88, F315–F318. doi: 10.1136/fn.88.4.F315

Szymczak, S. E., and Shellhaas, R. A. (2014). Impact of NICU design on environmental noise. *J. Neonatal Nurs.* 20, 77–81. doi: 10.1016/j.jnn.2013.07.003

Teixeira, L. R., Pega, F., Dzhambov, A. M., Bortkiewicz, A., da Silva, D., de Andrade, C., et al. (2021). The effect of occupational exposure to noise on ischaemic heart disease, stroke and hypertension: a systematic review and meta-analysis from the WHO/ILO Joint Estimates of the work-related burden of disease and injury. *Environ. Int.* 154:106387. doi: 10.1016/j.envint.2021.106387

Thibaud, J. P., and Amphoux, P. (2013). Silencing the city?. *SoundEffects Interdiscip. J. Sound Sound Exp*. 3, 62–70. doi: 10.7146/se.v3i3.18441

Thomas, K. A. (1989). How the NICU environment sounds to a preterm infant. *MCN Am. J. Matern. Child Nurs.* 14, 249–251. doi: 10.1097/00005721-198907000-00007

Thomas, K. A., and Uran, A. (2007). How the NICU environment sounds to a preterm infant: update. *MCN Am. J. Matern. Child Nurs.* 32, 250–253. doi: 10.1097/01.NMC.0000281966.23034.e9

Torresin, S., Albatici, R., Aletta, F., Babich, F., Oberman, T., Siboni, S., et al. (2020). Indoor soundscape assessment: a principal components model of acoustic perception in residential building. *Build. Environ.* 182:107152. doi: 10.1016/j.buildenv.2020.107152

Trickey, A. W., Arnold, C. C., Parmar, A., and Lasky, R. E. (2012). Sound levels, staff perceptions, and patient outcomes during renovation near the neonatal intensive care unit. *Health Environ. Res. Design J.* 5, 76–87. doi: 10.1177/193758671200500407

Varisco, G., van de Mortel, H., Cabrera-Quiros, L., Atallah, L., Hueske-Kraus, D., Long, X., et al. (2021). Optimisation of clinical workflow and monitor settings safely reduces alarms in the NICU. *Acta Paediatr*. 110, 1141–1150. doi: 10.1111/apa.15615

Wang, D., Aubertin, C., Barrowman, N., Moreau, K., Dunn, S., and Harrold, J. (2014). Examining the effects of a targeted noise reduction program in a neonatal intensive care unit. *Arch. Dis. Child. Fetal Neonatal Ed.* 99, F203–F208. doi: 10.1136/archdischild-2013-304928

Wei, W., and Van Renterghem, T. (2014). "Monitoring sound exposure by real time measurement and dynamic noise map," in *Forum Acusticum 2014 Proceedings* (Krakow EAA).

Williams, A. L., Sanderson, M., Lai, D., Selwyn, B. J., and Lasky, R. E. (2009). Intensive care noise and mean arterial blood pressure in extremely low-birth-weight neonates. *Am. J. Perinatol.* 26, 323–329. doi: 10.1055/s-0028-1104741

Williams, A. L., van Drongelen, W., and Lasky, R. E. (2007). Noise in contemporary neonatal intensive care. *J. Acoust. Soc. Am.* 121, 2681–2690. doi: 10.1121/1.2717500

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 24 October 2023
DOI 10.3389/fcomp.2023.1158476

Check for updates

OPEN ACCESS

EDITED BY
Emma Margareta Frid,
Royal Institute of Technology, Sweden

REVIEWED BY
Alex Michael Lucas,
Queen's University Belfast, United Kingdom
Fabio Morreale,
The University of Auckland, New Zealand

*CORRESPONDENCE
Erivan Gonçalves Duarte
✉ erivan.duarte@mail.mcgill.ca

RECEIVED 03 February 2023
ACCEPTED 02 October 2023
PUBLISHED 24 October 2023

CITATION
Duarte EG, Cossette I and Wanderley MM (2023) Analysis of Accessible Digital Musical Instruments through the lens of disability models: a case study with instruments targeting d/Deaf people. *Front. Comput. Sci.* 5:1158476. doi: 10.3389/fcomp.2023.1158476

Analysis of Accessible Digital Musical Instruments through the lens of disability models: a case study with instruments targeting d/Deaf people

Erivan Gonçalves Duarte[1,2,3]*, Isabelle Cossette[1,2] and Marcelo M. Wanderley[1,2,3]

[1]Department of Music Research, McGill University, Montreal, QC, Canada, [2]CIRMMT, Montreal, QC, Canada, [3]IDMIL, Montreal, QC, Canada

Music educators and researchers have grown increasingly aware of the need for traditional musical practices to promote inclusive music for disabled people. Inclusive music participation has been addressed by Accessible Digital Musical Instruments (ADMIs), which welcome different ways of playing and perceiving music, with considerable impact on music-making for disabled people. ADMIs offer exciting possibilities for instrument design to consider and incorporate individual constraints (e.g., missing arm, low vision, hearing loss, etc.) more than traditional acoustic instruments, whose generally fixed design allows little room for disabled musicians inclusivity. Relatively few works discuss ADMIs in the context of disability studies, and no work has investigated the impact of different disability models in the process of designing inclusive music technology. This paper proposes criteria to classify ADMIs according to the medical, social, and cultural models of disability, then applies these criteria to evaluate eleven ADMIs targeting d/Deaf people. This analysis allows us to reflect on the design of ADMIs from different perspectives of disability, giving insights for future projects and deepening our understanding of medical, social, and cultural aspects of accessible music technology.

KEYWORDS

Accessible Digital Musical Instruments, disability studies, music and disability, participatory design, Deaf culture, deafness

1. Introduction

In the last few decades, the disability rights movement has sought to affirm basic human rights, promoting equal access and inclusion for people experiencing different disabilities (Howe et al., 2015). Such activism has motivated landmark legislation that has raised challenges related to the inclusion of disabled people in music-making and music classroom activities (Hammel and Hourigan, 2011). Participation in music activities allows for emotional growth and sociability and the development of fine motor skills and cognitive capabilities (Hammel and Hourigan, 2011). These benefits highlight the importance of promoting broad access to music, which is particularly important for disabled people, whose physical and social barriers frequently exclude them from full participation in society (Adams et al., 2015).

The Accessible Digital Musical Instrument (ADMI)[1] field opens up new opportunities for inclusion by accommodating bodily differences and valuing multiple ways of perceiving music, such as through tactile and visual means (Davanzo and Avanzini, 2020). Scholars have observed that the development and design of ADMIs are often based on physiological "needs" at the expense of social and cultural ones (Frid, 2019a; Lucas et al., 2020). Authors have argued that studies on ADMIs have not dealt with the possibility of thinking about disability more deeply (Adams et al., 2015; Frid, 2019a; Lucas et al., 2020), highlighting that most studies in the field have focused mainly on approaches that can be identified with the medical model of disability (Frid, 2019a).

This paper provides original research by proposing criteria to classify ADMIs according to the medical, social, and cultural models of disability and apply these criteria to evaluate eleven ADMIs targeting d/Deaf[2] people. This analysis allows us to reflect on the design of ADMIs from different perspectives on disability, giving insights for future projects and deepening our understanding of medical, social, and cultural aspects of accessible music technology. This has not yet been explored in the existing literature. The paper addresses the following main needs: (1) the research in accessible music technology should connect with the different disability approaches found in the literature (Haag, 2017; Lucas, 2023), (2) researchers in the field should develop a literacy of disability models and present discussions about the impact of the model in their study (Haag, 2017; Lucas, 2023), and (3) studies in the ADMI field should consider approaches other than the medical model (Frid, 2019a; Lucas, 2023).

We address the three needs mentioned in the previous paragraph by proposing an analysis of ADMIs according to the three main disability models examined in the literature and by discussing the impact the different disability models may play on the ADMI design process. To enable this analysis and discussion, we proposed criteria to analyze ADMI development. These criteria are inspired by aspects of the medical, social, and cultural models of disability and can be used at any point in the development process, including before and during design. They can also guide future ADMI projects. First, our study addresses need 1 by providing an introduction and discussion of different disability models in the context of music and technology. Second, our criteria can be used to analyse, classify, and develop ADMIs according to the medical, social, and cultural models of disability, leading researchers to reflect on the importance of developing a literacy of disability models (need 2). Third, we propose a discussion about the implications of our criteria to the design of new ADMIs, highlighting the importance of considering design aspects beyond the medical model (need 3) and opening different perspectives on disabled musicians participation, inclusion, and evaluation. To our knowledge, no previous work proposes reflecting on ADMIs according to the medical, social, and cultural models nor considers how these models might affect the design of new ADMIs.

Another original point of this research is that we discuss ADMIs in the context of the cultural model of disability, presenting disability as a source of cultural diversity that can impact the development of new ADMIs. This avenue has not yet been explored in the existing literature. This discussion opens doors for reflections on music and technology from different cultural perspectives and provides researchers with a way of analyzing ADMIs that takes into account individual and collective practices and the experiences of disabled musicians.

The criteria proposed in this study were used to analyse eleven papers focused on ADMIs, all of which describe instruments that are used or potentially used by d/Deaf individuals. A search of the literature revealed that few studies are specifically designed for d/Deaf people (Frid, 2019a), suggesting that this research direction needs to be further explored. After introducing the medical, social, and cultural models of disability (Section 2.1) and discussing their implications for music and technology (Section 2.2), the paper proposes criteria to analyse ADMIs through the lens of these three models (Section 2.3). We then analyse eleven selected ADMIs according to those criteria (Section 3) and then discuss potential new insights for ADMI design in the context of the medical, social, and cultural models of disability (Section 4). The insights from using these criteria include: analyses of inclusive technology under different perspectives on disability, reflections on participatory methods, and reflections on the value of disability as a source of sociocultural diversity that can inspire and guide the development of new ADMIs.

2. Materials and methods

2.1. The medical, social, and cultural models of disability

The Western conception of disability has changed over the past 300 years. Straus (2011) explains that some conditions once legitimated by medical diagnoses no longer exist (e.g., hysteria, neurasthenia, fugue, and nostalgia) while others that never existed before (e.g., attention deficit disorder, autism, anorexia, and obesity) are now part of the discourse on disabilities. By the 18th century, disability was commonly conceived as a sign of divine disfavor, a punishment from God, and thus the external mark of an internal moral failing (Straus, 2011). In the 19th century, disability was generally characterized as a sign of divine inspiration and conceived as something to be left behind after

1 In the current paper ADMI is defined as "accessible musical control interface used in electronic music, inclusive music practice and music therapy settings" as proposed by Frid (2019a, p. 3). There are many terms describing research focusing on making digital music instruments accessible, for example, "adaptive digital musical instruments" and "assistive music technology" (Frid, 2019a). However the author explains that the word "assistance" implies an external source that provides aid to a person in need, whereas "adaptive" implies a constant state of refinement and adjustment to the musician (Frid, 2019a).

2 The term "Deaf" (capitalized) refers to people who value their Deafness and identify culturally as Deaf, usually communicating in Sign Language, and utilizing visual and tactile cultural behaviors (Jones, 2015; Holmes, 2017). Whereas deafness with a lowercase d denotes an audiological condition and is used here to identify people who have a hearing impairment (Jones, 2015; Holmes, 2017). We use the term d/Deaf when referring to both groups. Although conceptions of d/Deafness are more complex than these definitions imply, it is beyond the scope of this article to articulate them further.

spiritual transcendence (Straus, 2011). Since around 1800, disability has been largely considered as an individual pathology or defect that could be corrected by medical procedures (Straus, 2011). From the 1960s and 1970s, in the context of civil rights, disability began to be thought of as a sociocultural construct, not a fixed medical condition (Garland-Thomson, 1997).

In the next subsections, the three main models of disability (medical, social, and cultural) are presented and discussed in the context of music and technology as a basis for proposing criteria for ADMI analysis. Llewellyn and Hogan (2000) explain that models of disability are not synonymous with theory as their usage does not involve data collection and are not based upon data collection or methodology; however, they may serve as a tool to help us solve problems in research. They may have some usage as generators of hypotheses and enable us to represent information in a way that may aid understanding and generate explanations, providing us with different ways of examining the world of the disabled person (Llewellyn and Hogan, 2000).

Llewellyn and Hogan (2000) discuss the danger of overgeneralization when using disability models, arguing that researchers have to ask themselves if the advantages of a given model justify usage. In the present paper, we use the medical and social models as they have dominated the discussions in disabilities over the last 50 years (Toro et al., 2020), and they have been discussed in the ADMI literature (Frid, 2019a; Davanzo and Avanzini, 2020; Lucas et al., 2021). In addition, we choose to also include the cultural model, intended to provide us with a way of analyzing ADMIs by taking into account individual and collective practices and experiences of disabled musicians. However, other models or ways to conceptualize what "disability" means can be found in the literature: the Transactional Model (Llewellyn and Hogan, 2000) that is based on the basic premise that, beyond non-sport environments, disability is caused and sustained by problematic social relationships; the Ecological-Enactive Model (Toro et al., 2020), intended to disentangle the concepts of disability and pathology, locating the difference between pathological and normal forms of embodiment in a person's capacity to adapt to changes in the environment; and the Social-Ecological Models (Shogren et al., 2018) that challenge the assumption that disability resides only within the person and provides a conceptual framework to operationalize the understanding of disability.

2.1.1. Medical model

The medical model is based on the disease model used in medicine, considering all disability as the result of some clinically observable physiological impairment in bodily structure or function that needs appropriate "treatment" (Llewellyn and Hogan, 2000; Toro et al., 2020). Overall, the goal is to return the individual to a state of normalcy, allowing full participation in society (Lucas, 2023). The human being is considered as adaptable to the environment while society is fixed and unalterable (Llewellyn and Hogan, 2000). The medical model assumes a normative standard from which disability is an individual deviation that could be remedied and adjusted. Within this model view, disability is conceptualized as an individual body problem, and the solution is provided by medically oriented institutions, experts, and professionals (Beaudry, 2016). These experts and professionals have the task to treat the individual for the disabling condition, an idea that is often internalized once a person receives training, acquires expertise, and works in an environment dominated by the medical model (Oliver, 1983).

While the medical model recognizes a myriad of environmental factors linked to individual functional limitations, disability is understood in terms of the body of the disabled person as described objectively and scientifically, failing to recognize the lived embodiment of disabled persons (Toro et al., 2020). Waldschmidt (2018, p. 69) explains that from the medical approach "everything begins from the impairment, by presupposing that there are bodily, mental or psychological conditions that cause restrictions of participation and result in disability". For example, the appropriate medical response to learning problems is to treat impairments (such as hyperactivity disorder or dyslexia) through therapeutic intervention and to support the individuals to come to terms with their deficits (Waldschmidt, 2018).

Llewellyn and Hogan (2000) explain that this model has been seen as a force only to change disabled people into some more normal beings, disregarding social factors and not allowing disabled people to claim their major role in defining their own disability. Lucas (2023) points out that, in its extreme, the medical model sees diseases as attributable to biological and somatic factors that are possible to identify in a laboratory or clinical setting, omitting psychological, social, and cultural factors. The medical model is inclined to accept what is considered physiologically, socially, and culturally "normal" and does not allow much room for discussions about changes in the social status of disabled people (Straus, 2011).

The medical model could be associated with the status quo of Western music traditions and institutions that often seem to dictate what is "normal" in music, creating ideas about the right way musicians should play, compose, use instruments, etc. (Howe, 2015). A biased perception that music has to follow "normal" standards has the potential of leading researchers in music technology to design instruments intended to promote the "right" way musicians should play, compose, use instruments, etc. (Howe, 2015). From the medical perspective, music technology researchers are the experts responsible for understanding how individual impairments prevent participation in music and proposing technological solutions to overcome the disability, allowing the person to make music "normally" (Haag, 2017).

2.1.2. Social model

The social model switches the focus from individual limitations to the way the physical and social environment imposes limitations upon certain groups of people (Oliver, 1983). This model presents disability as a form of inequality caused by societal practices of disablement or incapacity to remove obstacles faced by disabled people rather than by impairments within the individual (Beaudry, 2016; Waldschmidt, 2018). From this perspective, disability is not a fixed, medical condition; rather, it emerges from a society that chooses to accommodate some bodies and exclude others (Garland-Thomson, 1997). Under the social model, adjustment becomes an issue for society, and the reflections about the able or disabled are displaced from the individual to the design of buildings, housing environments, expectations of others, working conditions, organizations of production, etc. (Oliver, 1983).

Within a social model of disability, it is argued that disability exists insofar as it is socially constructed and imposed on people with impairments; there is a de-emphasis on the individual, putting the discussions on disabilities back into the collective responsibility of society, emphasizing that limitations experienced by disabled people are caused by factors that come from outside of the person not from their impairment (Llewellyn and Hogan, 2000; Toro et al., 2020). The social model intends to contribute to the emancipation of disabled people, posing important questions about the barriers society imposes on them and reflections on policies intended to promote inclusive arrangements (Waldschmidt, 2018).

There are two main critiques of the social model. First, it does not address individual truth, perceptions, and beliefs about disability (Llewellyn and Hogan, 2000); the focus on society leads this model to neglect the embodied lived experience of the disabled person in the world (Toro et al., 2020). Second, the social model leaves little space for differences, such as race, age, and gender, between disabled people once this model focuses on collectivity to drive political change, sustaining that disabled people will benefit by banding together (Lucas, 2023).

An important characteristic of the social model is the distinction between impairment and disability. Impairment is related to loss or diminution of sight, hearing, mental ability, etc. susceptible to individual treatment/therapy, while disability is considered to be generated by the incapacity to remove obstacles faced by disabled people and generally associated with societal practices of disablement or exclusive social environments (Llewellyn and Hogan, 2000; Waldschmidt, 2018). For example, a d/Deaf person is disabled by oral language, but not by sign language; a wheelchair user is disabled by curbs, but not by sloped curbs. Oliver (1983) explains that if the problem of housing, for example, is taken from the medical perspective, the discussion would be around terms related to getting in and out, bathing, accessing the kitchen, and so on. That way, the discussions around the social model make it possible to see the creation of disability by the way housing is unsuited to certain individuals. Thus, there is a shift in emphasis from providing personal aids and therapy to adapting environments (Oliver, 1983). Beaudry (2016) argues that exclusion is the real problem; it is caused by a social failure to make adequate inclusive arrangements, accommodating some bodies and excluding others. For example, the access barrier of built environments may be observed through the attitudes held by building designers who may have overlooked or even devalued the requirements of wheelchair users (Lucas, 2023).

Lucas (2023) explains that the broader contextual factors raised by the social model can be applied to the design of technological products or musical instruments. The history of Western music has perpetuated an idea of normalcy tied to physical constraints, constructing an idea of a normal performance body (Lucas, 2023). For example, the acoustic piano requires the players to fit the piano's dimensions, being able to sit while moving and coordinating arms and legs. The social model perceives non-inclusive musical environments (concert halls, theaters, opera houses, etc.) and tools (music instruments, interfaces, stages, etc.) as disabling factors (Howe, 2015). In this way, the design, implementation, and use of assistive techniques and technologies may be an alternative to overcome barriers in music-making (Howe, 2015). Also, a study in consonance with the social model intends to contribute to the musical emancipation of disabled people. The research under the social approach promotes the participation of disabled musicians, allowing them to contribute to the design of new technologies that have the potential of overcoming social barriers and therefore facilitate inclusion and access to music (Haag, 2017; Lucas, 2023).

2.1.3. Cultural model

The cultural model of disability challenges "normality" and investigates how normalizing practices result in disability (Waldschmidt et al., 2017). Waldschmidt et al. (2017) explains that this model considers disability neither as only an individual fate, as in the medical model, nor as merely an effect of discrimination and exclusion, as in the social model. The cultural model investigates disabilities as a category constructed within a certain cultural and historical background that defines normalities and deviations, exclusionary and inclusive practices in everyday life and different institutions (Waldschmidt, 2018). Thus, attitudes toward impairments and the relation between impairment and disability are defined according to the cultural context. For example, deafness is typically regarded as a lack of hearing by non-disabled people; however, the cultural model supports the Deaf community view of deafness as a cultural difference and a source of linguistic competence in the form of sign language (Waldschmidt, 2018).

The cultural model maintains that disability exists only, and insofar as, certain differences (bodily and embodied) can be distinguished and thought of as "relevant for health" within a given cultural and historical order of knowledge and institutional support (Waldschmidt, 2018). Following this line of thinking, research does not simply investigate, for instance, the life course experience of disabled persons, but also those of persons considered non-disabled. Such a research approach challenges stigmatized cultural identities and outworn stereotypes by asking, for example, why personal autonomy is important for modern society and what normative expectations and constraints are attached to it (Waldschmidt, 2018). Similarly to the social model, the cultural model perceives non-inclusive musical environments and tools as disabling factors. However, the cultural approach allows us to go further and also include attitudes toward music, such as musical paradigms, musical pedagogical approaches, musical technique, etc. These environments, tools and attitudes are often not directed at disabled musicians.

The cultural model allows us to have different perspectives on common ideas and practices related to music and music technology. For example, Bowman (2009) explains that the practices related to music education define inclusion based on dominant-culture notions of what is adequate or inadequate, good or bad, and successful or unsuccessful. Deaf people, for example, are disabled by hearing-centric views of music, normalized listening paradigms (Straus, 2011), and music technology resources inspired by these views and paradigms. Such conceptions of music disregard the fact that Deaf listeners, music teachers, and musicians are proposing a much more multisensory hearing experience

(Jones, 2015). Deaf musicians are creating new ways to "feel" and "see" music and use music technology within the social and cultural context of the Deaf community (Best, 2016a,b). They are creating a musical movement that covers Deaf schools (Fawkes and Ratnanather, 2009), Deaf ensembles (Swinbourne, 2016), YouTube channels (Best, 2016a), and music technology projects (Hawley, 2016). Through these musical movements, Deaf musicians introduce new ideas about hearing, composition, education, and technology. The cultural model can promote this argument by celebrating disability as a cultural difference.

2.2. Accessible music technology and disability models

Music involves the integration of a wide range of cognitive, perceptual, motor, and emotional human capabilities. As Holland et al. (2019) explain, "music is a highly embodied activity in which performers and listeners routinely engage in complex synchronized movements in time with sound" (p. 2). Disabled performers (those with fewer fingers, weaker muscles, smaller lungs, or less vision than their instruments and sociocultural context require) lead us to ask what a performer can do (with one part of the body) and what a performer cannot do (with another part of the body); accessible musical instruments can accommodate the diversity related to these performers (Howe, 2015). The complexity of music poses challenges for developments in Human Computer Interaction (HCI), which should allow the accommodation of performers' impairments in comfortable and expressive ways.

The literature concerning disability in HCI has favored approaches tied to the medical model of disability (Haag, 2017). Lucas (2023) explains that this tendency partly originates from the lack of awareness of disability models within the HCI community. In this context, research projects present disability models implicitly and "researchers rarely state the model that frames their perspective" (Lucas, 2023, p. 14). Haag (2017) points out three possible parallels between the medical model and HCI studies addressing disability: (1) a shared exploration of technology as a corrective apparatus for defective bodies, (2) the comparable role of doctors and researchers to determine the nature of disability and design, and (3) a similar underestimation of disabled people's perspectives. Most of the papers reviewed by Haag (2017) are either explicitly based on the medical model or do not present an explicit reference to a particular model of disability. Haag (2017) urges the HCI community to update its discourse with respect to the vast literature on modern conceptions of disability. The problems that emerge from this field and their potential solutions are both constrained by the underlying model of disability (Haag, 2017), thus linking these solutions to the model implicitly or explicitly used. Research can help clarify the conceptual intricacies of disability to create more effective frameworks for developing solutions that fit the needs of the end beneficiaries (Haag, 2017). As Haag (2017) argues, researchers should ideally develop a literacy of disability models and state the model to which their research adheres.

One class of HCI development is linked to Digital Musical Instruments (DMIs) (Miranda and Wanderley, 2006). In the context of inclusive music practice, these instruments are known as Accessible DMIs or ADMIs. ADMIs are typically composed of a control surface where disabled musician interactions are measured by sensors whose values are mapped to sound synthesis algorithms. These interactions are mediated by gestural controls that are decoupled from sound production (sound control dissociation) but connected according to a mapping strategy. The sound control dissociation precludes any mechanical or physical constraints on sound production or gesture controls. Therefore, ADMI designers have more freedom than builders of acoustic instruments to provide tools intended for inclusive contexts (Frid, 2019b). The dissociation also allows the use of mappings between sounds and images or haptic elements. These instruments allow us to overcome the idea of a normally performing body (Howe, 2015) which usually possesses all limbs, appendages, physical capacities, and so on. Furthermore, ADMIs dismantle the dichotomy of a "right" or "wrong" way of playing an instrument, incentivizing curiosity and exploration, and promoting a sense of empowerment (Frid and Ilsar, 2021).

2.3. Criteria for ADMI classification through the lens of the medical, social, and cultural disability models

In this section, we connect ADMI design to the medical, social, and cultural models discussed in Sections 2.1 and 2.2 proposing criteria for ADMI classification based on four design elements. Despite Haag (2017)'s evidence for the value of broader disability literacy and greater transparency about disability models, researchers still tend to adhere exclusively to the medical model and frequently neglect to disclose their chosen approach (Haag, 2017). To encourage the exploration and transparent disclosure of different disability models, we propose that researchers analyse different ADMIs through the criteria of *user participation*, *disability view*, *inclusion view*, and *impact view*. These criteria are based on the three aforementioned disability models.

As observed by Frid (2019a), ADMI projects generally do not follow a straight line from conception to project design and do not clearly explain the project development steps. However, ADMI reviews (Frid, 2018; Frid, 2019a) and evaluations (Davanzo and Avanzini, 2020; Lucas et al., 2020) shed light on four commonalities among most ADMI projects: 1) specification of the potential user, 2) frequent reflections on disability, 3) occasional consideration of access and/or inclusion, and 4) discussion of the instrument's potential impact. These four common points inspired the four criteria we propose to analyse ADMI literature: 1) participation of users in the research process (*user participation*), 2) disability references or descriptions (*disability view*), 3) mentions of inclusion or/and access (*inclusion view*), and 4) potential impact of the ADMI on music-making for musicians with or without disabilities (*impact view*). The four elements are described below in relation to each model, and then summarized in Table 1.

User participation is defined here as the degree of participation attributed to users in the ADMI research process and can be classified according to each of the three disability models. Under the medical model, ADMI research concentrates on fixing impairments. Research is typically designed and conducted exclusively by the researchers, restricting the participants' role to only evaluating the instrument according to the researcher's view, sometimes, participants with disabilities are not even involved in the research process. Participants are mostly involved in order to address impairments and discuss how ADMIs may overcome them. Under the social model, the active participation of users focuses on addressing how to overcome social barriers (music environments, music tools, etc.) that limit access to music. According to the cultural model, reflections on culture guide our understanding of disability. As in the social model, users active participation contribute to overcoming cultural barriers (musical paradigms, musical pedagogical approaches, musical technique, etc.) to accessing music.

Our investigation on participation is not intended to describe in detail the participation approaches (Harder et al., 2013) used by the researchers in the papers to be analyzed. Associating the participation approach with the disability models, we classify each approach according to the inclusion or exclusion of the disabled person in the ADMI design process. If the disabled person is part of this process, the nature of the participation is identified by pointing out medical, social, or cultural aspects or characteristics as explained in Subsection 2.1 and illustrated by the analysis of each instrument in Section 3.

The *disability view* describes how ADMI research presents disability according to the three models. Researchers following the medical model investigate which individual conditions prevent the inclusion of disabled people in music, while those following the social model explore social conditions, practices, and attitudes that prevent their inclusion. Research influenced by the cultural model strives to understand the meaning of music-making in a given culture or society, and what cultural conditions, practices, and attitudes inhibit disabled people from participating in music activities.

The *inclusion view* refers to how ADMI research conceptualizes inclusion through the three models of disability. Research aligned with the medical model inquires whether and how technology and music practices support the overcoming of "music deficits". Socially oriented research examines how technology and music practices may "empower" disabled persons and promote accessible music environments by removing societal practices of disablement in music and deviating from "normalcy" (Howe, 2015) in the design and use of musical instruments. Culturally influenced research explores how technology and music resources may incorporate conceptions or music ideas from disabled performers. The way disabled people make music may inspire the development of technology. For example, the design and development of gesture controllers can accommodate the needs of diverse bodies, and culturally diverse perceptual modes may inform systems intended to visualize or feel music by using visual and haptic devices.

The *impact view* refers to the way that ADMI research evaluates the impact of instruments according to the medical, social, and cultural models. By way of the medical model, research on the benefits of ADMIs focuses on evaluating whether the instruments are "fixing" the disabilities of people living with them. No impact of these instruments is measured for people living without disability. Socially focused research advocates the benefits of ADMIs for disabled people and encourages the design of instruments that accommodate disabilities and enable independent music-making. This branch of research thus emphasizes ADMI benefits in the context of breaking down social barriers to music access. Research following the cultural model similarly emphasizes the benefits of ADMIs for both the disabled and non-disabled by integrating cultural diversity into instrument design. Culturally diverse approaches focus on breaking cultural barriers to make musical activities more accessible, while framing the artistic output of disabled people as valuable for everyone.

2.4. Methodology

2.4.1. Data collection

In order to find works describing ADMIs targeting d/Deaf individuals, we applied the method proposed by Frid (2019a) adapted to the d/Deaf context. Accordingly, we consulted the same databases (ScoPus, Google Scholar, and Web of Science Core Collection), using the same search phrases: ("Digital Music* Instrument*" OR "New Interface* for Musical Expression" OR "music* interface*" OR "music controller") AND (accessib* OR adapt* OR assistive OR inclus* OR empower*) AND (disabilit* OR health OR need OR impairment OR therap* OR disorder*). However, in order to adapt the last group of terms to the d/Deaf

TABLE 1 Criteria for ADMI classification according to the medical, social, and cultural models of disability.

	User participation	Disability view	Inclusion view	Impact view
Medical model (M)	No participation or limited participation, intended to validate the instruments designed by the experts.	Individual conditions prevent music participation.	Technology supports disabled persons in overcoming their "musical deficits".	The research does not consider the impact on non-disabled people, focusing only on disabled individuals.
Social Model (S)	Active participation in the design and evaluation process, with a focus on music needs to overcome social barriers to music access.	Societal conditions, practices, and attitudes prevent music participation.	Technology promotes accessible music environments by removing social barriers in music.	The research considers how the instruments benefit disabled people in overcoming social barriers.
Cultural model (C)	Active participation in the design and evaluation process, with a focus on cultural elements from disabled musicians/communities.	Cultural practices and discourses prevent music participation.	Technology incorporates culturally diverse ways of hearing and music-making.	The research considers how the benefits of developing technology inspired by different cultural backgrounds can impact everyone.

TABLE 2 Workflow.

Element	Question 1	Answer	Question 2	Disability model
User Participation	Are there disabled people participating in the research?	No.		M
		Yes.	Medical-based participation.	M
			Social-based participation.	S
			Cultural-based participation.	C
Disability view	What prevents participation in musical activities?	Individual/medical conditions.		M
		Societal conditions, practices, and attitudes.		S
		Cultural practices and discourses.		C
Inclusion view	How does the ADMI promote inclusion?	Technology supports disabled persons in overcoming their "musical deficits".		M
		Technology aims to remove social barriers that prevent access to music.		S
		Technology incorporates culturally diverse ways of hearing and making music.		C
Impact view	What is the context for discussions about the impact of the ADMI?	Medical reflections focusing only on disabled individuals.		M
		The instruments social benefits impact everyone.		S
		Technology inspired by different cultural backgrounds impacts everyone.		C

context, we replaced the phrases (disabilit* OR health OR need OR impairment OR therap* OR disorder*) with the phrases (deaf* OR hearing impairment OR cochlear implant OR deaf hearing OR hearing loss). Studies that were included in our analysis had to: a) present at least one ADMI directed to d/Deaf individuals, mentioning the potential target group(s) in either the title or the abstract; b) describe the implementation of an ADMI that enabled real-time manipulation of input or control data.

Following Frid's method (Frid, 2019a), we considered conference proceedings, journal papers, PhD theses, and book chapters written in English. The review considered both academic and gray literature (meaning valuable commercial and other non peer-reviewed literature that emerges through database searching). We selected the articles most appropriate to the purpose, with the following questions in mind: Does the study present the development of an ADMI or DMI targeting d/Deaf people or potentially addressing them? Does the study mention disability? Does the study provide any description of ADMI design and evaluation?

2.4.2. Data analysis

Our analysis focuses on four distinct elements of ADMI design as shown in Table 1. We intend to analyse ADMI design; however, we have to go through the overall research process to analyse and understand how the elements related to design are linked to the different disability models presented in Section 2.1. Our workflow is presented in the next paragraphs and summarized in Table 2.

First, we identified whether the researchers plan to include disabled people in the design process, the noninclusion indicates a medical approach; when the disabled individuals are included we have to analyse the nature of the interaction between researchers and individuals regarding the medical, social, and cultural models. The nature of the interaction is classified as: medical, when the focus is on overcoming disabilities; social, when disabled individuals and researchers work together to overcome social barriers; cultural, when the disabled individuals' cultural practices are put in the foreground. That way, we have the elements to establish the project's *user participation* (Section 2.3) in the design stage. We additionally analyzed the ADMI evaluation process, classifying *user participation* in accordance with the descriptions of instrument use and identifying which model guides the collaboration process with the disabled participant (medical, social, and cultural).

Second, we scrutinized how different researchers defined the design requirements by identifying the various intended target participants and the contexts of instrument use. This preliminary definition phase also uncovered researchers assumptions about the relationship between impairment and disabilities, allowing us to classify the project in terms of the *disability view* as described in Section 2.3.

Third, we examined the design choices of different instruments to identify associations between the *inclusion view* (Section 2.3) since the design choices reflect the ways in which technology promotes access. These associations can be classified into: 1) medical model, when design choices are aimed at overcoming individual deficits; 2) social model, when design choices target musical environments and aim to empower disabled musicians; or 3) cultural model, when design choices take into account or are inspired by the cultural practices of disabled musicians.

Fourth, we reviewed the conclusion and discussion sections of each article to classify their alignment with the *impact view* (Section 2.3) and the three disability models in terms of the different instruments benefits and the targets of these benefits. The *impact view* is classified as medical when benefits are discussed in the context of disability correction. Otherwise, the *impact view* can be classified as social or cultural, depending on whether the discussion of benefits focuses on social or cultural aspects related to the analyzed ADMI.

We tried to keep the analysis of each criterion to the elements of ADMI design mentioned in the previous paragraph. However, sometimes we had to adapt our analysis when some papers presented different structures or the criteria were identified across different elements or phases. For example, the elements that made it possible to identify the *disability view* or the *inclusion view* were frequently distributed through ADMI descriptions.

We point out that the models are not mutually exclusive, which is why some categories may be associated with more than one model. We use the letters M, S, and C to identify the medical, social, and cultural models respectively. When a category is associated with more than one model, the letters are ordered according to the model dominance, which is related to the number of answers to the questions in Table 2. For example, C, S, and M indicate that the cultural model is dominant (most of the questions in Table 2 are answered/classified with the cultural model), followed by the social and medical models.

3. Results

By applying our selection criteria, as presented in Section 2.3, we found eleven instruments targeting this population. The chosen papers are briefly presented in the next paragraphs.[3]

Mórimo (Zubrycka and Cyrta, 2012) aims to provide a platform for musical expression, with an emphasis on the tactile properties of sound. The work investigates the human body-music interaction, within the performative context in particular, targeting non-deaf and deaf people. There is no mention of deaf participants in the research process. Future works are intended to include deaf and non-deaf individuals, possibly within the context of music education.

Mogat (Zhou et al., 2012) is an ADMI-like tool in the form of three musical mobile games aimed at improving musical auditory skills for deaf children post Cochlear Implantation (CI participants). The mobile design was based on feedback from deaf children to understand their "deficiency" in terms of pitch, rhythm perception, and pitch production in contrast to non-deaf peers. The authors also proposed a user study with CI participants to demonstrate the effectiveness and efficiency of Mogat.

A new DMI proposed by Burn (2016) would enhance the musical experience for deaf musicians by using visual and haptic feedback systems to represent the components of sounds produced by DMIs, delivering a multi-sensory experience similar to that provided by acoustic instruments. Burn (2016) intends to design and evaluate the system together with deaf musicians.

Music aid (Søderberg et al., 2016) explores how music creation can occur collaboratively between deaf and hearing people. The overall goal of the research was to explore first, how to design an interface that would allow a deaf person to create music, and second, how this interface would support collaboration between hearing and deaf people. The authors propose the music interpretation and interaction through the visualization of sound along with haptic feedback. The design process relied on deaf and non-deaf people. Three alternative prototypes were developed to address the needs elicited from the design phase. These prototypes were then qualitatively evaluated in a controlled environment where non-deaf and partially deaf participants explored the prototypes by creating music together, followed by in-depth interviews with the two groups.

Duarte and Tavares (2017) present an ADMI implemented as a smartphone application with two options simulating piano or drums. The instruments are intended to allow the musician to have haptic and visual feedback from music in the context of music education. The instruments were evaluated with Deaf and hearing participants.

Wearable Musical Haptic Sleeves (Trivedi et al., 2019) are wearable haptic devices that deaf people can use to experience music. The devices are combined with a visualization system and can be controlled by an interface control. Deaf people are not included in the design and evaluation process.

Toc-Tum (Chaves et al., 2021) is an ADMI-like tool in the form of a virtual reality-based educational game accessible for deaf people. According to the authors, the system was "validated" in two different ways: (a) with the target audience to evaluate the interest in the game and (b) with people who have musical knowledge or who have had contact with the deaf, to assess the game's impact on musical instruction for deaf learners.

Smartphone Drum (Iijima et al., 2021) is a smartphone app that presents a vibrotactile sensation similar to that of a drum when the musician makes a drumming motion in the air with their smartphone as a drumstick. d/Deaf people are not included in the design process, but they are included in the evaluation process.

Cavdir and Wang (2022) present three wearable musical instruments (Bodyharp; Felt sound; and Touch, Listen, and (Re)Act) intended to deliver an embodied musical experience for diverse bodies. The instruments are developed according to an inclusive participatory design and performance practices to create more inclusive music performances. The Bodyharp does not rely on d/Deaf participants, while the Felt sound instrument includes them in the evaluation process, and the Touch, Listen, and (Re)Act includes them in both the design and evaluation.

In order to classify the ADMIs according to the medical, social, and cultural models, we analyzed the eleven ADMI papers, applying four criteria (Table 1). Table 3 classifies the ADMIs

3 In this section the terms "Deaf" and "deaf" are presented in the same way they appear in each paper analyzed.

TABLE 3 Classification of eleven ADMIs directed to deaf people according to the medical (M), social (S), and cultural (C) models of disability.

Date	Project name	User participation	Disability view	Inclusion view	Impact view
2012	Mórimo	No information	C	C	M and C
2012	Mogat	M	M	M	M
2016	No title (Burn, 2016)	C and M	C and S	C	C
2016	Music Aid	S	S	S	S and C
2016	No title (Duarte and Tavares, 2017)	M and C	S and M	M and S	S and C
2019	Wearable Musical Sleeves	M	M	M	M
2019	Toc-Tum	S and M	S, M, and C	S and M	S and M
2021	Smartphone Drum	M and S	M	M	M
2022	Bodyharp	M	C and M	C	C
2022	Felt Sound	C and M	C and M	C and M	C
2022	Touch, Listen, and (Re)Act	C, S, and M	C and M	C and M	C

TABLE 4 Occurrence and dominance (medical, social, and cultural models) for the 11 ADMIs analyzed.

	User participation (Occur./Dom.)	Disability view (Occur./Dom.)	Inclusion view (Occur./Dom.)	Impact view (Occur./Dom.)	Total (Occur./Dom.)
M	9/5	8/3	7/4	5/4	29/16
S	4/2	4/3	3/2	3/3	14/10
C	4/3	6/5	5/5	7/4	22/17

TABLE 5 Occurrence of combined models.

	User participation	Disability view	Inclusion view	Impact view	Total
C and M	2	3	2	0	7
S and M	1	1	1	1	4
M and S	1	0	1	0	2
M and C	1	0	0	1	2
S and C	0	0	0	2	2
C and S	0	1	0	0	1
S, M and C	0	1	0	0	1
C, S and M	1	0	0	0	1

according to our criteria, with a more frequent appearance of the medical model (M) than the other two models (S and C).

From Table 4, we can observe a greater occurrence of the medical model; however, when referring to dominance, we notice that these models present close numbers. This indicates that although the projects are considering different models, the medical model is the one with the highest occurrence in general, being present in practically all projects, except for Music Aid (Søderberg et al., 2016).

Interestingly, the data in Table 5 show that most combinations include M (83,3% of the total); however, M is not the dominant model for the majority of the combinations, indicating that the research projects are analyzing social and cultural aspects of disability, although they tend to recur to medical ideas at some point.

4. Discussion

The recurrence of the cultural model in Table 2 highlights the fact that the ADMIs analyzed rely on a few d/Deaf individuals as reference for their reflections about music and deafness. For example, the Felt Sound and Touch, Listen and (Re)Act projects consider and incorporate aspects of Deaf culture by investigating the use of sign language and consulting with d/Deaf people in the design process. Otherwise, most of the papers analyzed do not take into account the music movements within Deaf culture (Maler, 2015; Holmes, 2017) or elements of sign language (Maler, 2015; Best, 2016a).

The frequent occurrence of the medical model in our analysis is in line with Partesotti's argument of that most ADMI proposals limit their technology development to user needs that could be incorporated in the design (Partesotti et al., 2018). Such an approach can be compared to medical treatment models that disregard cultural and social elements associated with different types of disabilities (Lucas et al., 2020). However, our classification criteria allow us to observe that in our sample, ADMIs also consider sociocultural aspects of disabilities, even if these aspects are not prevalent. Thus, our criteria facilitate our understanding of how different disability approaches are distributed in the ADMI design process, sometimes combined, as shown in Table 5, and other times characterized by the frequency of their occurrence, as shown in Table 4. In this way, our framework allows us to consider and measure the occurrence of the three models and to reflect on their impact on each ADMI.

The combination of different models can be a way to keep ADMI design open to different perspectives and promote participation according to different musicians needs. The social and cultural models may give disabled musicians an active voice in the ADMI design process by advocating the active participation of disabled people in issues that concern them. We observed that the cultural model is predominant in the three most recent ADMIs analyzed, but it still lags behind the medical model overall, as we can see with the other eight ADMIs (Table 3). We also observed that the ADMIs that use the social and cultural model give special attention to the participation of disabled individuals in the design process.

Frid (2019a) points out some key concepts for the success of ADMIs, including adaptability and customization, iterative prototyping, user participation, and interdisciplinary development teams. The author also argues for the huge potential of the field to diversify and for ADMIs to benefit larger groups of disabled musicians. These outcomes could be achieved by implementing more advanced sensing technologies and gesture acquisition, incorporating vibrotactile feedback, exploring more diverse sound synthesis methods, and using more diverse mapping strategies (Frid, 2019a). However, this process does not guarantee the promotion of inclusion and diversity. Discussions of diversity and inclusion in computer music should consider the users and developers of these technologies, as well as who has access to technology (Frid, 2019b).

Frid and Ilsar (2021) highlight the importance of disabled musicians to build their own custom instruments – instruments designed by disabled musicians for disabled musicians. However, we note that the works reviewed here, with the exception of Cavdir and Wang (2022), either neglect to provide space for opinions of disabled people or do not clearly mention the use of participatory methods intended to include them in the ADMI design process. In the next two paragraphs, we discuss participatory research methods and technology in the hope of providing new ways to think about participatory approaches in ADMI design.

The literature further highlights some initiatives intended to promote participatory technology design. For instance, Parke-Wolfe et al. (2019) encourage the development of instrument-building toolkits, showing that disabled people, music teachers, and music therapists are capable of ingenious instrument designs when they can access the right tools. The re-use of high-quality modular technical resources will likely facilitate the development of bespoke Accessible Music Technology (AMT). Following a modular structure in hardware, software, and coding procedures can make it easier to fix, update, and expand DMIs (Lucas et al., 2020). These sensing devices can be used to control completely different types of parameters or detect different movements, enabling the control of different musical functions through a mapping strategy (Marshall et al., 2009).

As evidenced in the literature, the design of ADMIs should remain open to a diverse range of perspectives and preferences (Wright, 2020). Some studies, such as Förster et al. (2020) and Lucas et al. (2021), propose participatory ADMI design and emphasize the importance of approaches that take sociocultural aspects of music into account. Incorporating the the views of disabled people is a vital part of the ADMI design process (Frid, 2019b). The use of participatory design methods significantly helps to enable inclusive sonic interaction through the design of custom instruments (Frid, 2019b). For instance, action-research methodologies promote the autonomy of disabled people and balance the relations of power between investigators and participants in the research process. Ward (2023) presents a modular toolkit that illustrates use of participatory and iterative methods to address the issue of accessibility to music-making with music technology. McMillan and Morreale (2023) argue that the incorporation of lived and cultural aspects related to the experiences of disabled musicians can lead to more effective and personalized instrument design. Also, Lindetorp et al. (2023) highlight the importance of ADMIs in music performances facilitated by assistants.

The term "action research" covers a diverse range of approaches to research, aiming to change social practices and make research more productive, sustainable, just, and inclusive (Kemmis et al., 2014). These approaches have the potential to promote the autonomy of disabled people in the ADMI development and design processes once action research recognizes that (1) people in particular settings are capable of participating actively in all aspects of the research process and (2) participants have special access to how social and educational life are conducted in their surroundings. Therefore, these approaches can lead to research directed at making improvements in practices and participant settings (Kemmis et al., 2014). Action-research methodologies (Kemmis et al., 2014) can also play an important role in ADMI design, development, and evaluation. These methodologies could potentially be used to ensure the active participation of disabled people in the process of ADMI design. A primary concern of action research is to recognize research participants as capable of engaging actively in all aspects of the research process and value participants as "insiders" who have privileged access to the functioning and social rules of their settings (Kemmis et al., 2014).

In action research, the investigator can play the role of an outside consultant who can provide valuable support to participant researchers when it is useful. The outside consultants do not need to be members of a community undertaking an action research initiative, but can become engaged participants alongside "insiders" as long as they remain critically alert to differentiate their own self-interests from those of other participants (Kemmis et al., 2014). As Frid (2019b) states, such an approach can be useful in ADMI research, as the quality of ADMIs can be improved by incorporating the views of disabled people (insiders) and working in multidisciplinary teams, including engineers, interaction designers, music teachers and so on (outside consultants).

Finally, the recent increase of available computational resources, miniaturization, and sensors further enables the development of ADMIs that use non-conventional interaction paradigms and interfaces, opening up new opportunities for inclusion by accommodating bodily differences, and valuing multiple ways of sensory accessibility (Davanzo and Avanzini, 2020). Sensors associated with computing elements can be used to capture and process gestures, expanding the range of possibilities for computer music instruments (Wessel and Wright, 2002). The introduction of new instruments or the adaptation of existing ones can be an opportunity to transform musical cultures and contribute to shifts in listening modes (Pinch and Bijsterveld, 2004).

5. Conclusion

The criteria proposed in this paper allowed the researchers to identify a lack of participatory approaches and an overall tendency toward the medical model of disability for the eleven surveyed ADMIs. Our analysis indicates that participants do not have room for active participation in the research process and that most research does not consider the social and cultural aspects related to disabilities and music. Indeed, most of the studies based on the medical model did not include disabled people in the research process, unlike studies focused on the social and cultural models, which generally included disabled people.

Our chosen criteria allowed the analysis of 11 ADMIs according to the medical, social, and cultural models of disability. Our criteria also enabled the observation of different models across the data analyzed and the potential impact of the different disability models on the ADMI design process. The approach proposed here brings several benefits to the ADMI field: 1) it provides a way to introduce researchers to different disability models in the context of music and technology; 2) by classifying ADMI papers according to the medical, social, and cultural models of disability, it increases ADMI researchers awareness of different models of disability; 3) by considering technology outside of the medical context, it generates insights for the design of new ADMIs through different approaches to disabilities and different perspectives on disabled people participation, inclusion, and evaluation. Our study also emphasizes the importance of participatory methodologies and the value of disability as a source of cultural diversity that can impact the development of new ADMIs.

Our use of the proposed criteria demonstrates the insights that researchers can gain from applying our framework. Our analysis allowed the classification of the selected ADMIs, confirming the medical model as the most frequent approach, while further showing the continued influence of the social and cultural models, either in combination or individually. Our method thus provided a way to examine previous studies and reflect on them from different perspectives, opening up opportunities for more inclusive, comprehensive, and diverse ADMI developments in the future.

The criteria proposed here can help researchers reflect on previous and future projects and designs through the lens of different disability models. The criteria contribute to the field by providing insights about the development of inclusive technology under different perspectives on disability and by valuing disability as a source of sociocultural diversity to be considered in the ADMI design process.

Although our criteria encourage the understanding of ADMI design from different disability perspectives, our analysis of a limited data set of instruments (eleven) minimizes the potential for generalizations. Most of the ADMIs analyzed here do not allow the participants to engage actively in all aspects of the research process. Usually, the participants are consulted after the design process is finished. Their roles are restricted to answering questionnaires about their experience with the instruments, or performing activities that outsiders have designed to evaluate how the ADMIs can address their needs. These processes do not allow participants to have an active role in the design of the instruments or the activities. We addressed these issue by arguing about the potential impact of action research on the ADMI design process. The use of such approaches would allow the collaboration between disabled participants and engineers or music teachers, while keeping participants needs and preferences in the foreground.

We hope that our study will help to connect music technology researchers with the literature related to modern conceptions of disability, helping to develop frameworks that look to find solutions from multiple disability perspectives and leading researchers to reflect on the importance of developing a literacy of disability models, stating and discussing the model they adhere to in their research. Our contribution is intended to promote ADMI design that goes beyond physiological needs related to the medical model, considers social and cultural aspects related to disabilities, and thinks about disability more deeply.

A natural progression of this work is to consult with the designers of the ADMIs analyzed to discuss how our approach impacts their critical reflections on their work. A further study could evaluate the implications of applying the proposed criteria to analyse ADMIs targeting participants other than d/Deaf individuals as well as to help develop new ADMIs.

Data availability statement

The data supporting this research can be found in the nine articles mentioned in Section 3, further inquiries can be directed to the corresponding author.

Author contributions

ED proposed the concept for the article. All the authors discussed the article content, structure and research results. ED wrote the manuscript with the support from IC and MW. All authors contributed to the final manuscript.

Funding

This work was funded by Schulich School of Music of McGill University and Fonds de recherche du Québec—Société et culture, grant number 304200.

Conflict of interest

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

Publisher's note

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

References

Adams, R., Reiss, B., and Serlin, D. (2015). *Introduction.*, *Chapter Introduction*. New York University Press, New York, USA.

Beaudry, J.-S. (2016). Beyond (models of) disability? *J. Med. Philos.* 41, 210–228. doi: 10.1093/jmp/jhv063

Best, K. E. (2016a). Musical belonging in a hearing-centric society: adapting and contesting dominant cultural norms through deaf hip hop. *Song Popul. Cult.* 60–61, 61–86.

Best, K. E. (2016b). We still have a dream: the deaf hip hop movement and the struggle against the socio-cultural marginalization of deaf people. *Song Popul. Cult.* 60–61, 61–86. doi: 10.2307/26538859

Bowman, W. D. (2009). The community in music. *Int. J. Commu. Music.* 2, 109–128. doi: 10.1386/ijcm.2.2-3.109_1

Burn, R. (2016). Music-making for the deaf: exploring new ways of enhancing musical experience with visual and haptic feedback systems," in *Proceedings of the 2016 International Conference on Live Interfaces* (Brighton; Falmer: Experimental Music Technologies (EMuTe) Lab and REFRAME Books), 206–209.

Cavdir, D., and Wang, G. (2022). Designing felt experiences with movement-based, wearable musical instruments: from inclusive practices toward participatory design. *Wearable Technol.* 3, 15. doi: 10.1017/wtc.2022.15

Chaves, E. M., Braga, P. B. d. A., Montenegro, Y. F. L., Rodrigues, V. B., Munguba, M. C., et al. (2021). Toc-tum mini-games: an educational game accessible for deaf culture based on virtual reality. *Expert Syst.* 38, e12470. doi: 10.1111/exsy.12470

Davanzo, N., and Avanzini, F. (2020). A dimension space for the evaluation of accessible digital musical instruments," in *International Conference on New Interfaces for Musical Expression*, 214–220.

Duarte, E., and Tavares, T. (2017). A tool for the musical education of deaf people," in *Proceedings of the Brazilian Symposia on Computer Music*, 152–153.

Fawkes, W. G., and Ratnanather, J. T. (2009). Music at the mary hare grammar school for the deaf from 1975 to 1988. *Visions Res. Music Educ.* 14, 4.

Förster, A., Komesker, C., and Schnell, N. (2020). SnoeSky and SonicDive - design and evaluation of two accessible digital musical instruments for a SEN school," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. Honolulu: Zenodo, 83–88.

Frid, E. (2018). Accessible digital musical instruments: a survey of inclusive instruments presented at the nime, smc and icmc conferences," in *Proceedings of the International Computer Music Conference 2018 : Daegu, South Korea, International Computer Music Conference Proceedings*. San Francisco: The International Computer Music Association, 53–59.

Frid, E. (2019a). Accessible digital musical instruments - a review of musical interfaces in inclusive music practice. *Multimodal Technol. Interact.* 3, 3. doi: 10.3390/mti3030057

Frid, E. (2019b). *Diverse Sounds: Enabling Inclusive Sonic Interaction* (PhD thesis). Stockholm: KTH Royal Institute of Technology.

Frid, E., and Ilsar, A. (2021). Reimagining (accessible) digital musical instruments: a survey on electronic music-making tools," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. Shanghai, China: International Conference on New Interfaces for Musical Expression, 171–220.

Garland-Thomson, R. (1997). *Extraordinary Bodies: Figuring Physical Disability in American Culture and Literature*. Columbia University Press., New York.

Haag, G. A. (2017). *The Role of HCI in the Construction of Disability* (Master's thesis) Uruguay: Universidad de la República (Uruguay).

Hammel, A., and Hourigan, R. M. (2011). *Teaching Music to Students with Special Needs: A Label-Free Approach*. New York: Oxford University Press.

Harder, M. K., Burford, G., and Hoover, E. (2013). What is participation? Design leads the way to a cross-disciplinary framework. *Design Issues* 29, 41–57. doi: 10.1162/DESI_a_00229

Hawley, R. (2016). *Frequalise: A Project Exploring the Use of Music Technology with Deaf and Hard of Hearing Children and Young People [pdf] An Evaluation Report by Ros Hawley* (London: Youth Music Network).

Holland, S., Mudd, T., Wilkie-McKenna, K., McPherson, A., and Wanderley, M. M. (2019). *Understanding Music Interaction, and Why It Matters, chapter Introduction, pages 1-20*. Cham: Springer International Publishing.

Holmes, J. A. (2017). Expert listening beyond the limits of hearing: music and deafness. *J. Am. Musicolog. Soc.* 70, 171–220. doi: 10.1525/jams.2017.70.1.171

Howe, B. (2015). 191Disabling Music Performance," in *The Oxford Handbook of Music and Disability Studies*. Oxford: Oxford University Press, 191–209.

Howe, B., Jensen-Moulton, S., Lerner, N., and Straus, J. (2015). Introduction: disability studies in music, music in disability studies," in *The Oxford Handbook of Music and Disability Studies*, Oxford: Oxford University Press, 1–12.

Iijima, R., Shitara, A., Sarcar, S., and Ochiai, Y. (2021). Smartphone drum: Gesture-based digital musical instruments application for deaf and hard of hearing people," in *Proceedings of the 2021 ACM Symposium on Spatial User Interaction, SUI '21*. New York, NY, USA: Association for Computing Machinery, 1–2.

Jones, J. D. (2015). Imagined hearing: music-making in deaf culture," in *The Oxford Handbook of Music and Disability Studies*. Oxford: Oxford University Press, 54–72.

Kemmis, S., McTaggart, R., and Nixon, R. (2014). *Introducing Critical Participatory Action Research*. Singapore: Springer Singapore.

Lindetorp, H., Svahn, M., Hölling, J., Falkenberg, K., and Frid, E. (2023). Collaborative music-making: special educational needs school assistants as facilitators in performances with accessible digital musical instruments. *Front. Comp. Sci.* 5. doi: 10.3389/fcomp.2023.1165442

Llewellyn, A., and Hogan, K. (2000). The use and abuse of models of disability. *Disabil. Soc.* 15, 157–165. doi: 10.1080/09687590025829

Lucas, A. (2023). *Exploring and Cultivating Connections Within Ecosystems of Inclusive Musicking: Moving Towards Sustained Access to Music* (PhD thesis) Belfast: Queen's University Belfast.

Lucas, A., Harrison, J., Schroeder, F., and Ortiz, M. (2021). Cross-pollinating ecological perspectives in admi design and evaluation," in *NIME*, 22.

Lucas, A., Schroeder, F., and Ortiz, M. (2020). Enabling communities of practice surrounding the design and use of custom accessible music technology. *Comp. Music J.* 44, 9–23. doi: 10.1162/comj_a_00567

Maler, A. (2015). Musical Expression among Deaf and Hearing Song Signers," in *The Oxford Handbook of Music and Disability Studies*. Oxford: Oxford University Press, 73–91.

Marshall, M. T., Hartshorn, M., Wanderley, M. M., and Levitin, D. J. (2009). Sensor choice for parameter modulations in digital musical instruments: empirical evidence from pitch modulation. *J. New Music Res.* 38, 241–253. doi: 10.1080/09298210903085865

McMillan, A., and Morreale, F. (2023). Designing accessible musical instruments by addressing musician-instrument relationships. *Front. Comp. Sci.* 5, 1153232. doi: 10.3389/fcomp.2023.1153232

Miranda, E. R., and Wanderley, M. M. (2006). *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. Middleton, WI: A-R Editions.

Oliver, M. (1983). *Social Work with Disabled People*. London: The Macmillian Press LTD.

Parke-Wolfe, S. T., Scurto, H., and Fiebrink, R. (2019). Sound control: supporting custom musical interface design for children with disabilities," in *New Interfaces for Musical Expression*, 192–197.

Partesotti, E., Peñalba, A., and Manzolli, J. (2018). Digital instruments and their uses in music therapy. *Nordic J. Music Ther.* 27, 399–418. doi: 10.1080/08098131.2018.1490919

Pinch, T., and Bijsterveld, K. (2004). Sound studies: new technologies and music. *Soc. Stud. Sci.* 34, 635–648. doi: 10.1177/0306312704047615

Shogren, K. A., Wehmeyer, M. L., Martinis, J., and Blanck, P. (2018). *Social-Ecological Models of Disability*. Cambridge: Cambridge Disability Law and Policy Series, Cambridge University Press.

Søderberg, E., Odgaard, R., Bitsch, S., Høeg-Jensen, O., Christensen, N., Poulsen, S., et al. (2016). Music aid: Towards a collaborative experience for deaf and hearing people in creating music," in *Proceedings of the International Conference on New Interfaces for*

Musical Expression (NIME 2016). Mexico City: New Interfaces for Musical Expression, NIME, 321–326.

Straus, J. N. (2011). *Extraordinary Measures: Disability in Music*. Oxford: Oxford University Press.

Swinbourne, C. (2016). *Music and the Deaf: Find Out About our Forte Ensemble of Deaf Musicians and Our Frequalise Project*!.

Toro, J., Kiverstein, J., and Rietveld, E. (2020). The ecological-enactive model of disability: Why disability does not entail pathological embodiment. *Front. Psychol*. 11, 01162. doi: 10.3389/fpsyg.2020.01162

Trivedi, U., Alqasemi, R., and Dubey, R. (2019). Wearable musical haptic sleeves for people with hearing impairment," in *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments, PETRA '19*. New York, NY, USA: Association for Computing Machinery, 146–151.

Waldschmidt, A., Berressem, H., and Ingwersen, M. (2017). *Disability Goes Cultural: The Cultural Model of Disability as an Analytical Tool*. Verlag: Bielefeld, 19–28.

Waldschmidt, A. (2018). Disability culture society: Strengths and weaknesses of a cultural model of dis/ability," in *Sixth Alter Conference, Lausanne, 2017 Disability, Recognition and "Community living". Diversity of practices and plurality of values* (Amsterdam: Elsevier).

Ward, A. (2023). The development of a modular accessible musical instrument (mami) technology toolkit using action research. *Front. Comp. Sci*. 5, 1113078. doi: 10.3389/fcomp.2023.1113078

Wessel, D., and Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Comput. Music J*. 26, 11–22. doi: 10.1162/01489260232058 2945

Wright, J. E. (2020). "The appropriation and utility of constrained admis," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Birmingham), 564–569.

Zhou, Y., Sim, K. C., Tan, P., and Wang, Y. (2012). Mogat: Mobile games with auditory training for children with cochlear implants," in *Proceedings of the 20th ACM International Conference on Multimedia*. New York, NY, USA: Association for Computing Machinery, 429–438.

Zubrycka, J., and Cyrta, P. (2012). Tactile sound aesthetic experience," in *Non-Cochlear Sound: Proceedings of the 38th International Computer Music Conference*. Ljubljana, Slovenia: Michigan Publishing.

Frontiers in Computer Science

Explores fundamental and applied computer science to advance our understanding of the digital era

An innovative journal that fosters interdisciplinary research within computational sciences and explores the application of computer science in other research domains.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact