Proceedings Paper:

Geary, David, Francombe, Jon, Hentschel, Kristian et al. (1 more author) (2022) A survey of co-located multi-device audio experiences. In: AES 2022 International Audio for Virtual and Augmented Reality Conference (August 2022). Audio Engineering Society Conference: AES 2022 International Audio for Virtual and Augmented Reality Conference, 15-17 Aug 2022 Audio Engineering Society , USA

A SURVEY OF CO-LOCATED MULTI-DEVICE AUDIO EXPERIENCES

PRESENTED AT THE 2022 INTERNATIONAL CONFERENCE ON AUDIO FOR VIRTUAL AND AUGMENTED REALITY (AVAR 2022)

David Geary *
Department of Electrical Engineering
University of York
York, YO10 5DD, UK
drg519@york.ac.uk

Jon Francombe
Bang & Olufsen a/s
7600 Struer, Denmark
jofr@bang-olufsen.dk

Kristian Hentschel
BBC Research & Development
Salford M50 2LH, UK
kristian.hentschel@bbc.co.uk

Damian Murphy
Department of Electrical Engineering
University of York
York, YO10 5DD, UK
damian.murphy@york.ac.uk

August 24, 2022

ABSTRACT

Complimentary multi-device audio experiences are becoming increasingly common through the proliferation of mobile computing and the creation of bespoke multi-device audio production tools. However, the best use cases and design practices for these experiences remain less well understood. A review of multi-device audio experiences is therefore necessary to capture and consolidate the knowledge in this research area, and to move toward a set of design guidelines for creating more effective and engaging experiences. In this study, the range of applications of co-located multi-device audio experiences is explored and documented through a review of the literature and a survey, resulting in a dataset containing 31 individual experiences and 11 enabling tools or platforms. An initial analysis of the survey data revealed the frequency of types of devices and the forms of interaction in the experiences and platforms of the dataset. The full dataset is available at `https://doi.org/10.5281/zenodo.6839250`.

Keywords multi-device · audio · design · augmented-audio reality · applications · review

1 Introduction

Mobile computing devices are now ubiquitous in modern society, where different devices are used depending on the particular social and environmental context. The range in device form factors and capabilities continues to expand, illustrated by the uptick of 'smart' and wearable devices (1), and supported by the popular research fields of the Internet of Things (IoT) and Extended Reality (XR). In addition, multiple devices are now commonly used simultaneously, usually for separate activities but increasingly in connected, collaborative applications. Exploring the use of how multiple connected devices can be utilised to achieve a common goal is now a desired avenue of research (2; 3).

New opportunities now exist for the application of multiple devices for the delivery of audiovisual media, with benefits which include greater interaction possibilities and provision of complementary material. For audio applications, the use

*Corresponding Author. Preprint version. Original version to be found on: https://www.aes.org/e-lib/browse.cfm?elib=21834

of multiple commodity devices can be an economical method for spatial audio playback, compared with traditional multi-channel loudspeaker systems. Additionally, interaction options such as personalisation can be made easily available and accessible, especially when combined with object-based audio methods. Investigation into these audio applications has begun through individual prototypes and case studies; however, little research has been conducted which reviews and consolidates the collective efforts of these experiences. This paper sets out to address this issue by obtaining and reviewing information on the range of applications and associated technologies of co-located multi-device audio experiences.

A survey has been carried out to obtain a dataset which includes descriptions of co-located multi-device audio experiences and technologies. This dataset was analysed to retrieve information on types of devices used, interaction modes and other relevant design aspects.

2 Related work

2.1 Multi-device experiences

Research efforts are ongoing to better understand multi-device ecosystems and their associated challenges in terms of interaction and user behaviours, which has led to the development of design frameworks and taxonomies (3; 4; 5). A large proportion of this effort has a focus on the visual modality, centered around multi-display type applications such as augmented reality TV (6; 7) and second screen viewing (8; 9; 10).

The most prominent application of using multiple standalone devices for audio reproduction is creating arrays of speakers that allow for the spatial distribution of audio content, in a manner similar to traditional multi-channel audio experiences. This form of experience has been experimented with since the 1970s, where analogue radios and cassette players have been used to deliver experimental music installations (11; 12). The advent of personal mobile devices such as smartphones and laptops has since accelerated speaker array experience research, illustrated by the arrival of mobile and laptop orchestras (13; 14). These orchestras not only provide a performance platform but also have been used in an educational setting to teach music computing and human-computer interaction (14). More recently, speaker array experiences have become more accessible and interactive utilising modern web technologies such as HTML5 and the Web Audio API (15). Much of this work has been covered by Taylor (16) in his review of speaker arrays.

Aside from music-based speaker array experiences, the medium has been explored for broadcast type content such as audio dramas (17) and sports programmes (18). , for example, is a science-fiction audio drama designed to bring spatial audio into audiences' homes (19). A listener can connect multiple devices to a listening session through visiting a unique web URL, where a different mix is presented depending on the number of devices in the session. The experience was evaluated using a questionnaire and logged interaction data, and was reviewed positively. Out of 210 respondents, 72% liked using phones as loudspeakers with 80% saying they would use the technology again (20). Other applications include augmenting TV audio using acoustically transparent audio devices by providing additional content such as audio description (21).

In the last decade, a number of software tools have been developed with the purpose of creating multi-device audio experiences. *Soundworks* is a modular JavaScript framework which enables developers and artists to configure multi-device ecosystems and create and deliver interactive audio experiences through the web (15). Similarly, *HappyBrackets* is a Java-based programming environment that enables creative coding of multi-device ecosystems and can be used as a performance tool for multi-device audio experiences (22). *Audio Orchestrator* is another production tool for creating multi-device audio experiences that allows flexible routing of audio material to any number of devices, which can be customised through definable metadata and 'behaviours' (23). Each of these tools offer different features and cater to different users, and as a result have been responsible for the increased diversity observed in co-located multi-device audio experiences.

Generally, an issue within the multi-device audio experience literature is that there is a lack of consensus on the use of terminology, where there are a few different terms which describe similar concepts. For example, *mobile multi-speaker audio* (24), *media multiplicities* (5) and *device orchestration* (19) all refer to the use of multiple standalone devices for audio playback. As a result, traversing the literature can be quite difficult, justifying the need for a high level review of multi-device audio experiences.

2.2 Audio Augmented Reality

Audio augmented reality (AAR) can be described as the addition of virtual or digital audio material to a real auditory environment (25) and has seen increased interest due to the development of acoustically transparent wearable devices or 'hearables' (26). These devices are excellent candidates for AAR applications as they not only allow delivery of digital audio without acoustically isolating the listener, but also have a number of other useful affordances such as tactile interaction and head tracking for spatial audio rendering (27). However, AAR applications are not limited to wearable devices and can be achieved with loudspeakers in the right circumstances.

AAR applications and experiences can take many different forms, illustrated by Krzyzaniak's application taxonomy (28) where six different types have been identified. Co-located multi-device audio experiences can exhibit AAR characteristics, and overlap particularly well with two of these experience types. *Enchanted Objects* describes the sonification of objects with digital audio, that would not ordinarily make sound. Bown (29) has explored the use of multiple enchanted objects in a game of lawn bowls by using bespoke embedded audio devices as replacements for the balls, augmenting a typically quiet game with musical elements. The second type is *overlay of extra audio information*, where the listener is provided with additional audio information or content, specific to the present activity. McGill et. al. (21) have investigated such applications with multiple devices using acoustically transparent headsets to augment TV audio by overlaying additional speech content through the headset. The two experiences demonstrate two differing multi-device AAR approaches, the former where two or more devices augment the real surroundings and the latter where one or more devices augment the output of another device.

Perhaps the largest benefit of using multiple devices in AAR experiences is enabling of both personal and shared audio, where audience members can listen to personalised audio, whilst still being able to easily interact with other people. Such benefits are starting to be reported in case studies. The evaluation of *Please Confirm You Are Not A Robot*, a multiplayer audio augmented reality game using Bose Frames (30), revealed that participants thought the social aspect of the experience was the best feature. Building on these promising results will take further research and consideration as to how AAR and multi-device audio experiences intersect and benefit each other.

3 Methodology

This body of work looks to tie together the (largely dissociated) range of research into co-located multi-device audio experiences, with the objective of providing researchers with a dataset which can be utilised in further research. To achieve this, a survey and a systematic literature search were employed to obtain as much information as possible on these experiences. These methods are described in detail in the following section.

3.1 Scope of the research

This study is focused on co-located experiences, where the devices are positioned within the same physical space or group of adjoining spaces as the listener(s). In theory, the listener can perceive the auditory output from each device. Therefore, experiences where the devices are remote, such as in streaming parties or teleconferencing, are not considered in this work. In addition, experiences where one or more devices control the audio output of another device, without reproducing audio themselves (31) are also outside the scope of this paper.

3.2 Survey

To better understand co-located multi-device audio experiences, a survey was conducted to collect information on relevant experiences and technologies. The survey was delivered through the *Qualtrics* platform (32). Academic networks, social media and mailing lists were utilised to distribute the survey to knowledgeable researchers and creative practitioners. A set of inclusion criteria were outlined at the start of the survey to direct responses.

1. *The platform/experience must employ multiple devices with loudspeakers.*
2. *The audio content must be distributed across the devices.*
3. *The devices must be co-located in a single space, or group of adjoining spaces.*

The survey comprised of some questions with free-text fields, requesting participants to describe the experience and its audio content. In addition, a few multiple choice questions were asked, capturing the types of devices in the ecosystem; the modes of interaction for the experience; whether the number of devices was variable; and what the roles of the devices within the ecosystem were.

3.3 Systematic literature search

As a supplementary method to the survey, a systematic search of the literature was undertaken to identify relevant papers and obtain more instances of multi-device audio experiences and technologies. For this, the PRISMA methodology for systematic reviews and meta-analyses was adopted (33). The Association for Computing Machinery (ACM) Digital

Library (34) and the Audio Engineering Society (AES) E-Library (35) were targeted for the search. These two were chosen as they were likely to contain the highest number of relevant papers. Paper titles, abstracts and key words were queried in both databases using the following query string.

((“multi-device" OR “cross-device" OR “distributed" OR “orchestra" OR “multiple devices") AND (“audio" OR “sound" OR “music"))*

The search was completed by the first author and restricted to journal and conference publications between 2000 and 2021. The initial pool of 361 papers was initially screened using the same inclusion criteria set out in the survey, by reading the title and abstracts. This process resulted in the exclusion of 290 papers, leaving 71 papers remaining. These papers were then subsequently screened by reading the full texts. Through this process, 11 papers were identified which contained application instances of co-located multi-device audio platforms and experiences, as defined by the survey inclusion criteria. The first author then added these instances to the survey, and the resulting dataset was analysed (see Section 4).

4 Results

The survey received 42 responses (including responses from the authors and those discovered during the literature search described in Section 3). Of these responses, 11 were classed as platforms and 31 were classed as experiences. The full dataset of platforms and experiences is available at: `https://doi.org/10.5281/zenodo.6839250`, and at reference (36).

The frequency of the types of devices present within each platform/experience ecosystem in the dataset is shown in Figure 1. It can be seen that the prevalence of each type relates to the size and portability of the device type, with mobile phones being the most common devices (33%) and TV/radio being least common. This could be attributed to the advent of multi-device audio experience production tools such as *Soundworks* and *Audio Orchestrator*, which enable fast experience prototyping with mobile devices. Additionally, these results approximately reflect the market saturation of each device type, with the exception of TV and radio which are underrepresented (37). The ‘Other’ category, receiving 23% of selections, contained custom embedded audio devices utilising micro-controller/processors, ‘hearables’ such as the Bose Frames (38) and cassette players. This category contains bespoke instances of experiences rather than platforms, and includes micro-controller/processor based enchanted object AAR experiences such as SoundWear (39) and Bloom (40). Additionally, it includes social audio games and augmenting TV audio applications using acoustically transparent headsets, as described in Section 2.

Devices that are audio only, such as smart speakers, are less common than devices with visual displays. These devices are only observed within the platform ecosystems of Alexa (41), Sonos (42) and Google Home (43), where the application of usage is mostly limited to streaming the same audio content on each device, perhaps in different spaces within a home. This could be due to the challenges around developing for, and integrating, audio only devices, which are still relatively young in comparison to other device types. Many modern experiences are delivered through a web URL that can be typed or accessed via a QR code; however, an additional solution will be required to integrate audio only devices.

Figure 2 shows the frequency of the different forms of device interaction. The most common forms are tactile input and touch controls (63%), followed by motion gesturing (19%), which are possessed most commonly by mobile phones and tablets. It is unclear at this stage whether the most common devices and interaction methods observed are due to a primary desire for these particular interaction methods, or whether the interactions are a consequence of the desired devices. In this chart, the ‘Other’ category includes, but may not be limited to, proximity-based interaction using RFID, or Bluetooth technologies. Most platforms/experiences were labelled as interactive, with only 16% being considered not interactive. The number of involved devices within an experience was reported to be variable in 84% of responses. While the question was insufficient to clearly elicit whether devices numbers could vary *during* or *between* instances of experiences, the data still demonstrates the flexibility and modularity of many co-located multi-device audio experiences, which is in contrast to traditional multi-channel audio experiences. For the final question about the differences in importance for involved devices, 58% of entries indicated that all the devices were of equal importance, whereas 26% reported some form of responsibility hierarchy for the devices in the ecosystem. This highlights that the individual roles of each device in an experience may be an important consideration for experience design.

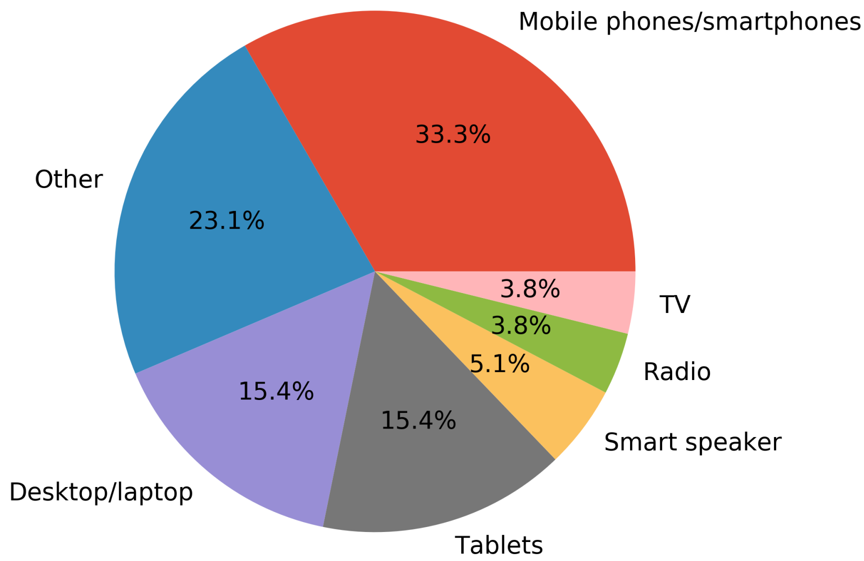

Figure 1: The frequency of types of devices for all survey responses.

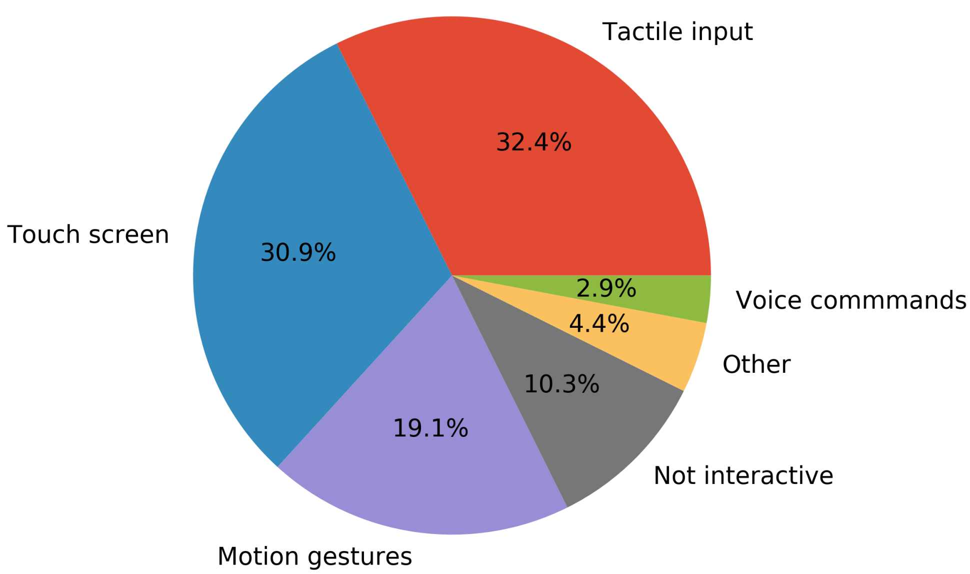

Figure 2: The frequency of modes of interaction for all survey responses.

5 Conclusion and future work

As multi-device experience research continues to develop, a larger proportion of it is required to explore the benefits of using multiple devices for audio reproduction. Currently, work in this particular area seems to be disconnected with a lack of consensus on the use of terminology. The creation of experience production tools such as *Audio Orchestrator* and *Soundworks* has lowered the barrier to development of such experiences; however, there remains a gap in knowledge around the best use cases and design principles. This work has made initial steps toward addressing this gap through collection and evaluation of current co-located multi-device audio experiences and technologies.

A survey and structured literature review were used to produce a dataset which contains device and interaction information and descriptions of co-located multi-device audio experiences and platforms. Out of the 42 entries, 31 were classified as experiences and 11 were classified as platforms. Personal mobile devices are most commonly utilised for these types of experiences, with tactile input and touch gestures being the most common form of interaction. Smart speakers and voice interaction are uncommon outside of established smart device ecosystems. In 84% of entries, the number of devices was reported as variable, demonstrating modularity and flexibility.

Further work will include analysing the free-text data from the survey using qualitative methods and using that information to form a design framework for co-located multi-device audio experiences.

Acknowledgements

This work is supported in part by the UK Arts and Humanities Research Council (AHRC) XR Stories Creative Industries Cluster project, grant no. AH/S002839/1, in part by a University of York funded PhD[or MSc Research] studentship, with additional support from BBC Research & Development. Thank you to everyone who contributed to the co-located multi-device audio dataset.

References

[1] C. Rieger and T. A. Majchrzak, "A taxonomy for App-Enabled devices: Mastering the mobile device jungle," in *Web Information Systems and Technologies*, pp. 202–220, Springer International Publishing, 2018.

[2] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, pp. 2787–2805, Oct. 2010.

[3] F. Brudy, C. Holz, R. Rädle, C.-J. Wu, S. Houben, C. N. Klokmose, and N. Marquardt, "Cross-Device taxonomy: Survey, opportunities and challenges of interactions spanning across multiple devices," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–28, New York, NY, USA: Association for Computing Machinery, May 2019.

[4] M. Levin, *Designing Multi-Device Experiences: An Ecosystem Approach to User Experiences across Devices*. "O'Reilly Media, Inc.", Feb. 2014.

[5] O. Bown and S. Ferguson, "Understanding media multiplicities," *Entertain. Comput.*, vol. 25, pp. 62–70, Mar. 2018.

[6] R.-D. Vatavu, P. Saeghe, T. Chambel, V. Vinayagamoorthy, and M. F. Ursu, "Conceptualizing augmented reality television for the living room," in *ACM International Conference on Interactive Media Experiences*, IMX '20, (New York, NY, USA), pp. 1–12, Association for Computing Machinery, June 2020.

[7] P. Saeghe, G. Abercrombie, B. Weir, S. Clinch, S. Pettifer, and R. Stevens, "Augmented reality and television: Dimensions and themes," in *ACM International Conference on Interactive Media Experiences*, IMX '20, (New York, NY, USA), pp. 13–23, Association for Computing Machinery, June 2020.

[8] V. Vinayagamoorthy, R. Ramdhany, and M. Hammond, "Enabling Frame-Accurate synchronised companion screen experiences," in *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '16, (New York, NY, USA), pp. 83–92, Association for Computing Machinery, June 2016.

[9] T. Neate, M. Jones, and M. Evans, "Cross-device media: a review of second screening and multi-device television," *Pers. Ubiquit. Comput.*, vol. 21, pp. 391–405, Apr. 2017.

[10] V. Lohmüller and C. Wolff, "Towards a comprehensive definition of second screen," in *Proceedings of Mensch und Computer 2019*, MuC'19, (New York, NY, USA), pp. 167–177, Association for Computing Machinery, Sept. 2019.

[11] J. Maceda, *Ugnayan*. Tzadik, 2009.

[12] E. Team, "Cassettes 100 by josé maceda, re-staged at national gallery singapore | national gallery singapore." https://www.nationalgallery.sg/magazine/cassettes-100-jose-maceda-at-national-gallery-singapore, January 2020. (Accessed on 04/12/2022).

[13] D. Trueman, P. R. Cook, S. Smallwood, and G. Wang, "Plork: the princeton laptop orchestra, year 1," in *ICMC*, 2006.

[14] G. Wang, N. J. Bryan, J. Oh, and R. Hamilton, "Stanford laptop orchestra (slork)," in *ICMC*, Citeseer, 2009.

[15] S. Robaszkiewicz and N. Schnell, "Soundworks–a playground for artists and developers to create collaborative mobile web performances," in *WAC-1st Web Audio Conference*, wac.ircam.fr, 2015.

[16] B. Taylor, "A history of the audience as a speaker array," in *NIME*, pp. 481–486, pdfs.semanticscholar.org, 2017.

[17] E. Young and J. Francombe, "Monster: A hallowe'en horror played out through your audio devices - bbc r&d." https://www.bbc.co.uk/rd/blog/2020-10-audio-drama-monster-interactive-sound, October 2020. (Accessed on 03/28/2022).

[18] J. Francombe, "Get to the centre of the scrum - our immersive six nations rugby trial - bbc r&d." https://www.bbc.co.uk/rd/blog/2021-02-synchronised-audio-devices-sound-immersive, 2021. (Accessed on 03/07/2022).

[19] J. Francombe, J. Woodcock, R. J. Hughes, K. Hentschel, E. Whitmore, and A. Churnside, "Producing audio drama content for an array of orchestrated personal devices," tech. rep., 2019.

[20] J. Francombe and K. Hentschel, *Evaluation of an immersive audio experience using questionnaire and interaction data*. Universitätsbibliothek der RWTH Aachen, 2019.

[21] M. McGill, F. Mathis, M. Khamis, and J. Williamson, "Augmenting TV viewing using acoustically transparent auditory headsets," in *ACM International Conference on Interactive Media Experiences*, IMX '20, (New York, NY, USA), pp. 34–44, Association for Computing Machinery, June 2020.

[22] A. Fraietta, O. Bown, S. Ferguson, S. Gillespie, and L. Bray, "Rapid composition for networked devices: HappyBrackets," *Comput. Music J.*, vol. 43, pp. 89–108, June 2020.

[23] K. Hentschel and J. Francombe, "Framework for web delivery of immersive audio experiences using device orchestration," *Adjunct Proceedings of ACM TVX*, 2019.

[24] H. Kim, S. Lee, J.-W. Choi, H. Bae, J. Lee, J. Song, and I. Shin, "Mobile maestro: enabling immersive multi-speaker audio applications on commodity mobile devices," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '14, (New York, NY, USA), pp. 277–288, Association for Computing Machinery, Sept. 2014.

[25] M. Lawton, S. Cunningham, and I. Convery, "Nature soundscapes: an audio augmented reality experience," in *Proceedings of the 15th International Conference on Audio Mostly*, pp. 85–92, New York, NY, USA: Association for Computing Machinery, Sept. 2020.

[26] P. Crum, "Hearables: Here come the: Technology tucked inside your ears will augment your daily life," *IEEE Spectrum*, vol. 56, pp. 38–43, May 2019.

[27] J. Plazak and M. Kersten-Oertel, "A survey on the affordances of "hearables"," *Inventions*, 2018.

[28] M. Krzyzaniak, D. Frohlich, and P. J. B. Jackson, "Six types of audio that DEFY reality! a taxonomy of audio augmented reality with examples," in *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound*, AM'19, (New York, NY, USA), pp. 160–167, Association for Computing Machinery, Sept. 2019.

[29] O. Bown and S. Ferguson, "A musical game of bowls using the DIADs," in *NIME*, pp. 371–372, nime.org, 2016.

[30] A. N. Nagele, V. Bauer, P. G. T. Healey, J. D. Reiss, H. Cooke, T. Cowlishaw, C. Baume, and C. Pike, "Interactive audio augmented reality in participatory performance," *Front. Virtual Real.*, vol. 1, Feb. 2021.

[31] L. Turchet and M. Barthet, "An ubiquitous smart guitar system for collaborative musical practice," *Journal of New Music Research*, vol. 48, pp. 352–365, Aug. 2019.

[32] Qualtrics, "Qualtrics xm // the leading experience management software." https://www.qualtrics.com/uk/?rid=ip&prevsite=en&newsite=uk&geo=GB&geomatch=uk, 2022. (Accessed on 04/19/2022).

[33] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and P. Group, "Preferred reporting items for systematic reviews and meta-analyses: the prisma statement," *Annals of internal medicine*, vol. 151, no. 4, pp. 264–269, 2009.

[34] ACM, "Acm digital library." https://dl.acm.org/, 2022. (Accessed on 04/19/2022).

[35] Audio-Engineering-Society, "Aes e-library." https://www.aes.org/e-lib/, 2022. (Accessed on 04/19/2022).

[36] D. Geary, "Co-located multi-device audio experiences dataset," Jul 2022.

[37] C. Cieciura, R. Mason, P. Coleman, and M. Paradis, "Survey of media device ownership, media service usage, and group media consumption in UK households," in *Audio Engineering Society Convention 145*, aes.org, 2018.

[38] Bose, "Wearables by bose—classic bluetooth® audio sunglasses." `https://www.bose.co.uk/en_gb/products/frames/bose-frames-alto.html`, 2022. (Accessed on 04/19/2022).

[39] J. Hong, H. Yi, J. Pyun, and W. Lee, "SoundWear: Effect of Non-Speech sound augmentation on the outdoor play experience of children," in *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, DIS '20, (New York, NY, USA), pp. 2201–2213, Association for Computing Machinery, 2020.

[40] S. J. Ferguson, A. Rowe, O. Bown, L. Birtles, and C. Bennewith, "Sound design for a system of 1000 distributed independent audio-visual devices," *New Interfaces for Musical Expression 2017*, 2017.

[41] A. Charlton, "Amazon echo multi-room music with alexa: How to set up - gearbrain." `https://www.gearbrain.com/amazon-echo-multi-room-music-2528315228.html`, February 2020. (Accessed on 06/30/2022).

[42] Sonos, "Sonos | wireless speakers and home sound systems." `https://www.sonos.com/en-gb/home`, 2022. (Accessed on 05/05/2022).

[43] Google, "Smart home automation from google | google home." `https://home.google.com/welcome/`, 2022. (Accessed on 07/19/2022).