Article

Design Dimensions of Co-Located Multi-Device Audio Experiences

David Geary [1,*], **Jon Francombe** [2,3,†], **Kristian Hentschel** [3] **and Damian Murphy** [1]

1   Audiolab, Deparment of Electronics, University of York, York YO10 5DD, UK; damian.murphy@york.ac.uk
2   Bang & Olufsen a/s, 7600 Struer, Denmark; jofr@bang-olufsen.dk
3   BBC Research & Development, Salford M50 2LH, UK; kristian.hentschel@bbc.co.uk
*   Correspondence: drg519@york.ac.uk
†   Jon Francombe was affiliated with BBC R&D. He is now affiliated with Bang & Olufsen.

Abstract: The widespread distribution of mobile computing presents new opportunities for the consumption of interactive and immersive media experiences using multiple connected devices. Tools now exist for the creation of these experiences; however, there is still limited understanding of the best design practices and use cases for the technology, especially in the context of audio experiences. In this study, the application space of co-located multi-device audio experiences is explored and documented through a review of the literature and a survey. Using the obtained information, a set of seven design dimensions that can be used to characterise and compare experiences of this type is proposed; these are *synchronisation*, *context*, *position*, *relationship*, *interactivity*, *organisation*, and *distribution*. A mapping of the current application space is presented where four categories are identified using the design dimensions, these are *public performances*, *interactive music*, *augmented broadcasting*, and *social games*. Finally, the overlap between co-located multi-device audio and audio-augmented reality (AAR) experiences is highlighted and discussed. This work will contribute to the wider discussion about the role of multiple devices in audio experiences and provide a source of reference for the design of future multi-device audio experiences.

Keywords: multi-device; audio; applications; augmented-audio reality; design

1. Introduction

For many individuals in modern 21st century society, daily life includes many interactions with a number of different mobile computing devices. This is a result of the proliferation of such devices, including smartphones and tablets, over the last two decades. Moreover, an increasing number of individuals now own multiple such devices. New types of 'smart' devices are also emerging which diverge from traditional form factors, such as smart-watches and smart-eyewear [1]. While these devices appear physically different, many have similar hardware features which can enable a wide array of applications, especially through software 'apps'. As a consequence, there is an increasing number of devices that have different form factors but overlap in their possible applications.

The device that is selected for use at any given time depends on the chosen activity and various contextual factors such as location and social surroundings. However, it is increasingly common to use multiple devices at the same time, either for separate activities or in a connected, collaborative capacity. Research exploring the use of many connected devices to accomplish a task has become popular, most notably under the banner of the Internet of Things (IoT) [2], and aims to harness the power of intentional multi-device networking, which can bring valuable applications to users [3].

One such application is the use of multiple connected devices to consume and interact with audiovisual media content. This approach has benefits which includes the provision of additional content and greater interaction possibilities, especially in social contexts. For audio experiences, using multiple commodity devices can be an accessible, cost-effective

method for spatial or immersive audio playback and can also enable straightforward participatory interaction within experiences. In the last five years, the way these devices can be utilised in the design of audio experiences has started to be established through case studies and early demonstrations [4,5]. However, in order to develop the application space further, more work is needed on experience characterisation and design to identify new experience opportunities and to increase our conceptual understanding of the experiences. This is necessary, as much of the research on multi-device experience design until now has focused on the visual modality. Furthermore, the work across multi-device audio experiences suffers from an inconsistent use of terminology, making navigating the literature quite difficult. Consequently, this work addresses the following questions.

1. What are the defining characteristics of co-located multi-device audio experiences?
2. How do these characteristics manifest in different applications?

A survey was carried out to obtain a dataset which included descriptions of co-located multi-device audio experiences. The survey was then analysed to retrieve information on design aspects and attributes. A set of resulting design dimensions was generated and presented to aid the understanding and characterisation of co-located, multi-device audio experiences. In addition, the application space was reviewed and mapped into example categories using a method in which use-cases can be categorised and new instances can be discovered.

This paper is organised as follows. In Section 2, a review of the current state of multi-device experience design is given, alongside a brief summary of the development of the multi-device audio medium, including relevant production tools and applications. Within Sections 3 and 4, the methodology for the survey and the results are detailed. The design dimensions are presented and discussed in Section 5, followed by an analysis and review of the applications in the dataset in Section 6. In Section 7, the use of the design dimensions, and the overlap between the applications and audio-augmented reality (AAR) applications is discussed. Finally, a summary of the work is provided and future work discussed in Section 8.

2. Related Work

Multi-device ecosystems and their applications are complex and encounter many additional challenges compared to their single-device counterparts. To tackle this complexity, and to enable the creation of more meaningful multi-device experiences, research has been conducted to increase our understanding, more broadly, of multi-device interactions and user behaviours.

Levin [6] presents a set of high-level design approaches to multi-device experiences, named the 3Cs, which consist of *consistent*, *continuous*, and *complementary* methods. These act as a practical guide for designing multi-device interactions for media experiences and provide a useful introduction for newcomers to the research space. As this research area has progressed and new prototypes have emerged, efforts have been made to consolidate the work through surveys and taxonomies. An excellent example of this is presented by Brudy et al. [3], who report on design characteristics, applications, tracking systems, and interaction techniques for multi-device systems. Moreover, for contemporary multi-device audiovisual systems, Bown and Ferguson [7] propose the term 'media multiplicities' to describe coordinating systems of devices that interact to manifest media content and present four key affordances: *spatial*, where the devices are distributed in physical space; *scatterable*, where devices can move around and be reconfigured; *sensing*, devices can interact and coordinate using sensors; and *scalable*, where devices can be added and removed freely.

While many of the discussed concepts can be applied in audio-only applications, most of the discussion is centered around multi-screen applications. This follows a common trend of research focusing on the visual modality, exploring applications such as 'second screen' viewing [8–10] and augmented-reality TV [11,12]. In addition, while many experiences are comprised of both visual and auditory stimuli, it is also valid to consider audio and visual

experience aspects separately, due to the differences via which both of these modalities are perceived by the human body.

The use of multiple devices for audio reproduction is most commonly applied in music-based performances, for instance, where the audience's devices are used as an array of loudspeakers. The bulk of this research occurred over the last decade with the rapid uptake of smartphones; however, this practice can be observed as far back as the 1970s, where analogue radios and cassette players have been used in experimental sound–art installations and exhibits [13–15]. Since then, mobile/laptop orchestras have been a popular avenue for speaker-array experimentation [4,16]. Dahl [17] described designing these experiences as a 'wicked problem', where creators must design the instrument and the interactions between devices, as well as composing the music. To address this complexity, a prototyping and iteration design approach is commonly chosen [18].

With the proliferation of personal computing devices, a number of tools have been developed to reduce the difficulty in creating effective, cross-platform, multi-device audio experiences. *Soundworks*, for example, is a JavaScript framework for creating multi-device audio experiences for consumer devices and has been used in many demonstrations of participatory sound and music experiences [19]. *Audio Orchestrator* is a production tool which allows producers to create multi-device audio experiences by facilitating audio routing to any number of devices through an allocation algorithm, which can be controlled through defined behaviours [20]. Using this tool, a number of trial productions have been made ranging from immersive audio dramas [21] to sports programs [22]. A final example is *HappyBrackets*, a programming environment for the creative coding of multi-device systems, that can also be used as a composition and performance tool for audio experiences [23].

Similar to many other developing research areas, the disjointed use of terminology is also a problem in this field, where many different terms are used to describe similar ideas or concepts. For example, *device orchestration* [21], *media multiplicities* [7], and *mobile multi-speaker audio* [24] can all refer to the use of multiple devices for audio reproduction. Finding some consensus on this issue would help to improve literature accessibility and may lead to more fruitful collaboration across this discipline. As noted above, the research into multi-device experiences has focused mainly on multi-display applications and includes a number of review papers capturing the breadth of the work [3,9,25]. In contrast, far fewer similar studies have been conducted in the audio domain, aside from Taylor's review of speaker array applications [15]. More effort is therefore necessary to capture and consolidate the knowledge around the applications of multiple devices for audio experiences.

3. Methodology

3.1. Scope of the Research

This paper is concerned only with co-located experiences, where the active devices are within the same physical space or group of adjoining spaces as the listener/audience. This excludes typically remote applications such as teleconferencing and streaming parties. Furthermore, a subset of multi-device audio experiences, where one or more devices control the audio output of a single device (such as the use of an additional controller with a smart instrument [26] or streaming media from one device to another) are also outside the scope of this paper.

During this period of work, attempts have been made to specifically define the word 'device' in the context of this work. However, no satisfactory definition was found; they were either too encompassing or too specific. Nonetheless, the loose criteria for devices in this work are that they are standalone, reproduce audio, and offer some form of interface for interaction. For this reason, multi-channel surround experiences have not been considered as co-located, multi-device audio experiences for this study, as they are essentially deemed to constitute a single device. A single loudspeaker may be used in isolation; however, they are designed to be used in multiples. A particular use case, somewhere near the definition boundary of these experiences, that was included in this study was silent films

or discos. This type of experience differs from multi-channel audio experiences in two ways. Firstly, the loudspeakers in a multi-channel system are specifically designed to be used in a collective, where the disruption of the number of speakers or the arrangement would likely diminish the experience. In the silent film/disco case, each set of headphones is independent, where the individual listening experience is approximately the same, regardless of whether there are 50 listeners or 500 listeners. Secondly, multi-channel speakers are purely static output objects with no intentional interaction patterns, whereas in silent films/discos, there is typically some form of device interaction, namely personal volume adjustment or selection of a different radio channel. For other research, there may be a legitimate case for including a different range of experiences in this study.

3.2. Survey

To address the first research question *What are the defining characteristics of co-located, multi-device audio experiences?*, a short survey was conducted to obtain information on both experiences and production tools and systems. A survey was selected as the chosen methodology in order to obtain as large a sample size as possible and was deemed more suitable for purpose versus other qualitative methods such as interviews and focus groups. Researchers, technologists, and creative practitioners were contacted through academic networks, mailing lists, and social media to contribute their knowledge. The survey was delivered using the *Qualtrics* platform [27]. To specify the types of experiences desired for this study, a set of inclusion criteria was defined at the beginning of the survey.

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.*

'Platform' refers to the underlying technologies and systems which support and deliver multi-device audio experiences, whereas 'experiences' are the single applications of those technologies which include the audio content. The survey consisted of open questions with free-text fields where participants could describe the experience and its audio content, as well as multiple choice questions capturing the types of devices used, the modes of interaction, whether the number of devices is variable, and the importance of the devices in the ecosystem. Participants were able to respond to the questionnaire multiple times if necessary. The full questionnaire can be viewed in Appendix A.

3.3. Systematic Literature Search

To supplement the survey, a systematic literature search was completed to locate relevant papers and capture more instances of co-located, multi-device audio platforms and experiences. This was achieved using the PRISMA methodology [28]. Both the Association for Computing Machinery (ACM) Digital Library [29] and the Audio Engineering Society (AES) E-Library [30] were targeted for the search. For both databases, the following query string was used to search paper titles, abstracts, and key words.

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

The search was conducted by the first author and constrained to papers published between 2000 and 2021. Using the same experience-inclusion criteria as the survey, the initial pool of 361 papers was preliminarily screened through reading the titles and abstracts. This screening stage resulted in 290 paper exclusions and 71 inclusions. These 71 papers were further screened by reading the full texts. Following this process, 11 papers were identified as containing application instances of multi-device audio platforms and experiences that fit into the survey inclusion criteria mentioned above. These instances were then added to the survey by the first author, and the resulting dataset was analysed (see Section 4).

3.4. Thematic Analysis

The free-text data was analysed using a thematic analysis approach, based upon Braun and Clarke's methods [31] and qualitative descriptive analysis [32], due to the descriptive nature of the data. The purpose of this analysis was to uncover defining characteristics for these types of experiences and to use them to develop a framework for understanding and comparing these seemingly related experiences. The analysis was conducted in three distinct stages. In the first stage, a panel of three experts (including two of the authors) individually assigned codes to each of the text responses. Themes were generated in the second stage, which involved the experts discussing and clustering the codes over two 1-hour group sessions, using a shared *Miro* board, and which were facilitated by the first author. In the final stage, three non-experts were recruited and given the same text responses as given to the experts to assign codes to. The non-expert codes were then compared and placed into the themes generated by the experts as a form of validation exercise.

4. Results

The survey received 43 responses in total (including responses from the authors and those discovered during the literature search described in Section 3). All multiple choice questions and responses can be seen in Appendix B, and a full list of experiences in the dataset can be seen in Appendix C. Figure 1a shows the frequency of device types from the platforms and experiences described in the responses. An observable trend is that the occurrence of each type largely relates to the size and mobility of the devices, with mobile phones being the most common (33%), and TV/radio being the least common (8%). This result roughly reflects the relative market saturation of each type of device [33] but is also due to the increasing number of experiences created through specific multi-device audio production tools such as *Soundworks* and *Audio Orchestrator*, which enable the easy inclusion of mobile devices. The 'Other' category, which received 23% of the selections, included Raspberry-Pi-based embedded audio devices and cassette players. Audio-only devices are far less common than those with a visual display. This can be attributed in part to integration challenges for these devices, where many of the modern experiences implement delivery through a web URL that can be typed or accessed via a QR code.

The frequency of various interaction modes is shown in Figure 1b. Tactile input and touch interfaces are most common (63%), followed by motion-gesturing interaction (19%), both of which smartphones and tablets commonly possess. The 'Other' category includes, but is not limited to, proximity-based control through the use of RFID or Bluetooth technologies. Out of all 43 entries, 16% of them are labelled as non-interactive. The number of devices was reported to be variable in 84% of entries. The data collected from the responses to the question are insufficient to determine whether the experience could vary the number of devices *during* or *between* instances of the experiences. Despite this, the results still demonstrate the flexible nature of these experiences and tend to display configuration agnosticism, unlike traditional multi-channel audio experiences. Finally, in answer to the question about device importance, 58% of responses indicated that the devices were perceived to be equally important, whereas 26% reported some form of hierarchical device structure in the ecosystem.

Through the thematic analysis process described in Section 3, 17 themes were generated (Table 1) which were captured under four parent themes: *devices*, *listeners*, *content*, and *other*. These themes were validated by a set of 300 non-expert codes, with 92% of the non-expert codes being represented within the 17 themes defined by the experts.
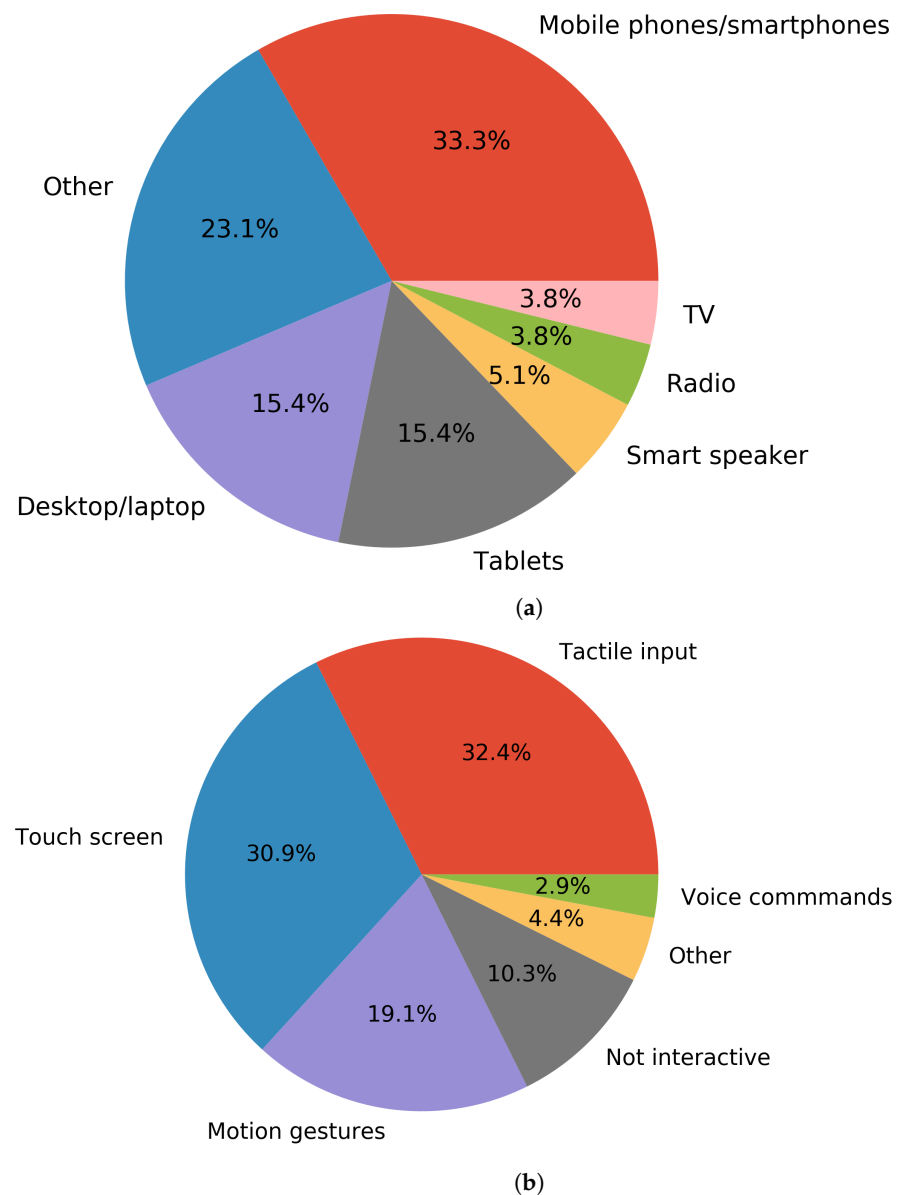
(**a**)

(**b**)

Figure 1. Pie charts representing the frequency of types of devices and modes of interaction for the multiple choice response data. (**a**) The frequency of types of devices for all survey responses. (**b**) The frequency of modes of interaction for all survey responses.

Table 1. The 17 themes generated from the thematic analysis of the free-text survey responses, their descriptions, and associated parent themes.

Parent Theme	Theme	Description
Devices	Device position	The position of the devices
	Synchronisation	The existence of and the accuracy of synchronisation in use
	Device hierarchy	The existence of and the configuration of the device hierarchy
	Types of device	The types of devices in use
	Number of devices	The number of devices in use
Listeners	Number of listeners	The number of listeners present
	Listener position	The position of the listeners
	Listener role	The role of the listener in the experience, including the extent of contribution
	Social interaction	The type and extent of social interaction during the experience

Table 1. *Cont.*

Parent Theme	Theme	Description
Content	Content interactivity	The extent of interaction with the content, including device input modalities
	Content hierarchy	The existence of and the configuration of the content hierarchy
	Content distribution	How the content is distributed across the devices
	Content genre	What does the audio content consist of i.e., music, soundscape
Other	Scale	The overall size and scale of the experience
	Flexible distribution	The ability for reconfiguration of devices and redistribution of audio
	Immersive	Experiences aiming to possess an immersive quality
	Type of experience	The format of the experience, i.e., performance, game, etc.

5. Design Dimensions

This section describes a set of seven design dimensions, derived from further analysis and refinement of the themes from the thematic analysis, with reference to relevant literature. This process involved re-analysing the themes and aggregating related themes into encompassing design aspect groups, which are different from the parent themes. For example, the *Listener role*, *Social interaction*, and *Content interactivity* are combined into an *Interactivity* design aspect, despite all being under the same parent theme. In other cases, larger individual themes were assigned their own design aspect. Once all the high-level design aspects were established, the design dimensions were formed by identifying all the distinct design choices associated with that aspect. Figure 2 illustrates the process of generating the design dimensions from the original data.

These dimensions aim to represent the key design considerations for co-located multi-device audio experiences. Within each dimension are a small number of categories which constitute the different design choices available for that dimension. For each dimension, examples from the literature are provided, and how the dimension affects various aspects of audio experiences is discussed. A few of the following dimensions have been adapted from or are related to Brudy's cross-device taxonomy [3], as many of the important design considerations are applicable to multi-device experiences beyond audio-only. Nevertheless, any adaptations made are intended to make the considerations more relevant to audio experiences. These dimensions relate to:

- the degree of audio *synchronisation*,
- the social and environmental *context* of the experience,
- the physical *position*ing of the devices,
- the interaction *relationship* between audience members and devices,
- the degree of audience *interactivity*,
- the roles of the devices and how they are *organised* in the ecosystem, and
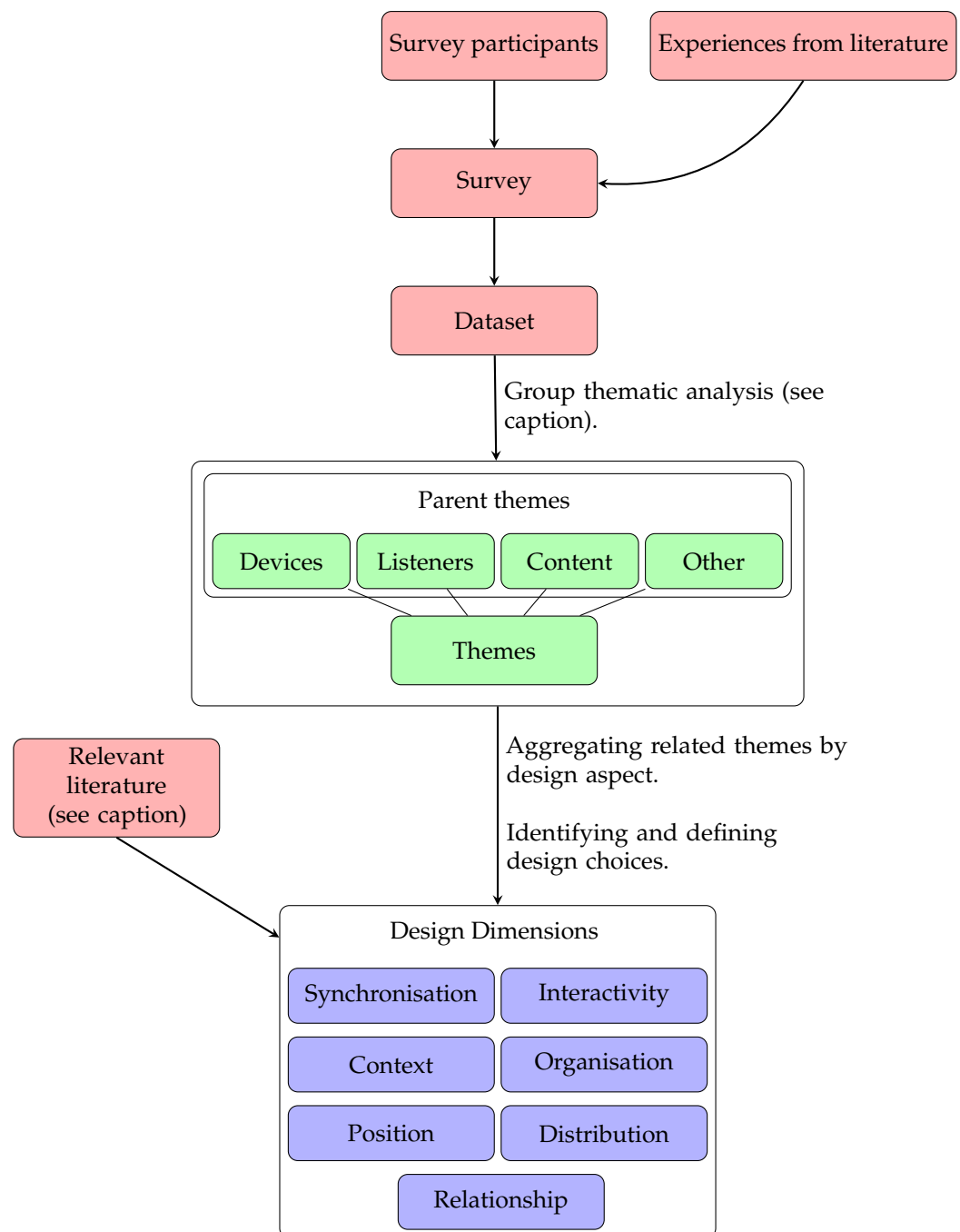- the *distribution* of audio content among the devices.

Figure 2. Diagram illustrating the methodology of obtaining the design dimensions from the original data. The thematic analysis followed Braun and Clarke's methods [31]. The relevant literature [3,7,34] was considered when forming the design dimensions.

5.1. Synchronisation

Synchronisation is defined as the extent to which the devices in the ecosystem are synchronised in their reproduction of audio. This dimension appropriates and extends Brudy's *temporal* dimension [3] for synchronous, multi-device interactions through three categories of synchronisation for audio content.

1. **Asynchronous** —In this instance, the devices are not synchronised in any form, acting completely independently from each other. The applications are limited, but asynchronous systems can be used for experimental art installations [35] and enchanted object experiences [36].

2. **Loose synchronisation**—Loosely synchronised audio can be achieved through artistic composition, without the use of device networking. The audio material is designed to be played at approximately the same time but without the need for precise timing. Many early multi-device audio experiences were limited to this approach, which typically involved listeners pressing play on devices at the same time [37] or other manual triggering of audio sources [38]. The synchronisation accuracy between devices is typically between 100 ms and 1 s.

3. **Tight synchronisation**—This category captures experiences where devices are connected and synchronised using a network. This approach is common in more modern experiences and enables musical performance applications [39] and narrative-driven experiences such as audio dramas [21]. For such applications, experiences in this category require device synchronisation of <100 ms or from word-level accuracy down to approximately frame-level synchronisation [8]. Typical synchronisation solutions for these experiences obtain a sync accuracy of around 20 ms for heterogeneous devices after calibration [40].

5.2. Context

This dimension captures the social and environmental context of the experience, such as the scope for social interaction and the physical space it takes place in. The three categories generally relate to the number of people and devices present. According to Bown's media multiplicities and their "significance of numbers" [7], the relationship between the number of loudspeakers and individual audio elements also affects how the audio is perceived.

1. **Personal**—Personal multi-device audio is an emerging class of experience involving a single listener and a small number of devices. The experiences usually have limited interactivity due to the impracticality of individually controlling multiple devices. Applications therefore tend to consist of overlaying additional audio content to a program [22] or enhancing the spatial audio image through placing devices in different positions [41].

2. **Social**—Experiences comprised of small to medium-sized groups of listeners with multiple devices are classified as *Social*. In this context, rich social interaction is possible and audio scenes with increasing complexity can be created, while it is still possible to perceive the individual contribution of the devices in the system [42]. Experiences of this kind can take place in both small domestic environments and larger spaces, depending on the exact number of involved devices and listeners.

3. **Public**—*Public* represents the larger-scale experiences with many listeners and devices. Experiences in this category commonly have more than 20 listeners and a large number of devices. In this context, communication or coordination is more difficult and commonly requires a central composer or figure to orchestrate the experience. Individual device contributions are typically displaced by the larger aggregate audio image. Many use-cases, such as art installations [35] and performances [13], fall into this category and tend to take place in large public venues.

5.3. Position

Position is concerned with the physical configuration and positions of the devices in the experience, and is similar to Brudy's *dynamics* dimension [3]. Three categories are identified.

1. **Fixed**—Device positions are typically static and pre-determined, similar to traditional channel-based speaker systems. The number of devices in the system tends to be known, and device placements are optimised for the focal aspect of the particular experience, i.e., the spatial audio image. For example, Tsui et al. designed a one-person orchestra experience [43], where device positions are carefully calibrated with motion controllers to enable audio control through motion gestures.

2. **Arbitrary**—Many modern tools and frameworks enable flexible, ad hoc device ecosystems where devices can join and leave at will. This ambiguity in the number of devices

makes the control of device positions very difficult. With this flexible positioning, devices can be placed in arbitrary positions. Often, these positions are the same as the listener position if personal devices are used [44,45]. However, this tends to remove the control of the position of individual sound sources in the spatial audio image.

3. **Dynamic**—In a few cases, experiences are designed with intentional device movement by utilising Bluetooth or RFID technologies, which can result in ever-changing spatial audio images, which can change over time [46,47].

5.4. Relationship

A number of different people-to-device interaction relationships exist with multi-device ecosystems. These are represented by three categories, similarly adopted from Brudy's *relationship* dimension [3].

1. **One-to-many**—The first category captures use-cases where an individual utilises many devices to expand or augment a single device listening experience [48,49]. This is aligned with the 'personal' value of the *context* dimension.
2. **Multiple one-to-one**—Many experiences implement a 'use your own device' approach, which is an easy and accessible method for delivering shared audio experiences to a group. Each listener interacts with their own device, therefore known as a multiple one-to-one relationship [5,38,50]. This relationship lends itself well to interactive experiences, where audience members can influence the experience.
3. **Many-to-many**—Other shared or collaborative settings where the numbers of devices and people are not equal can be associated under this category. In this case, there is no link between the number of listeners and the number of devices, where individuals may interact with different devices over the duration of the experience. Examples of this are commonly seen in art installations [35] and music performances [37].

5.5. Interactivity

While traditional, passive media experiences, such as TV and film, still have their place, modern technology now allows for experiences with varying degrees of audience interactivity. Consequently, this affects and broadens the possible roles of audience members in an experience, where an audience member can be described as a *bystander*, *spectator*, *customer*, *participant*, or *player* [34]. This dimension is split into three categories, a simplification of Striner's [34] spectrum of audience interactivity for entertainment domains.

1. **Passive**—*Passive* experiences capture more traditional media experience formats, where audiences are distinctly passive and have no influence on the media content. Many multi-device audio performances are of this nature [51] but so are storytelling experiences typical of radio or TV content [52].
2. **Influence**—This category includes application instances where the audience has some limited form of influence or control over the media content, but the experience creator retains all creative control over the experience. This could be personalisation through the selection of audio assets [22] or other influence, perhaps through low-impact manipulation of the audio signal [36].
3. **Create**—Highly interactive experiences where audience members could be better described as players or performers, actively creating the audio content of the experience, are situated in this final category [44,53]. For these experiences, the experience creator passes some of the creative control onto the audience members.

5.6. Organisation

Ecosystems of multiple devices introduce new questions around device organisation, from technical, social, and content delivery points of view. This dimension is focused on how the devices are configured for audio reproduction, and what the roles of each of the devices in the delivery of the experience are. This is linked to the *interaction topologies* presented by Matuszewski et al. for mobile-based music systems [54].

1. **Non-hierarchical**—The simplest approach is to assign all devices to have equal responsibility in the experience. In this configuration, each device contributes to the audio experience equally and is regularly observed in social experiences, where personal devices are used [44,47,50]. This can be achieved in practice through the disconnected graph or circular interaction topologies [54].

2. **Hierarchical**—In some cases, it may be better to organise devices into two or more tiers, with differing responsibilities and affordances. This is effective where the audio content can be organised into an importance hierarchy [21,55], e.g., dialogue and sound effects in TV and film, or where mobile devices cannot adequately reproduce certain frequencies or amplitudes and need support from an additional device [5,19]. A common implementation of this design choice is assigning a single device as the 'main' device and assigning other devices as 'auxiliary' devices [20]. These device types can be controlled separately and can be assigned different audio assets if necessary. Other configuration examples include the star and forest interaction topology graphs [54].

5.7. Distribution

The *distribution* dimension captures different approaches for the distribution of audio to the devices in an experience. This can be linked to the previous dimension, where audio distribution can be determined based on the roles of each of the devices in the experience. There are three categories.

1. **Mirrored**—The simplest case is where the same audio content is allocated to every device in the system. This mirrored approach is useful when a simple amplification or duplication of audio is required, e.g., in a silent disco or silent film [56], or in commercial multi-room speaker ecosystems such as Sonos [57].

2. **Distinct**—Here, the audio assets on each device are unique, where no two devices are playing the same audio. This approach is prevalent in interactive, music-based experiences where each device is allocated a unique instrument or sample, allowing each participant to have an equal yet distinct contribution to the experience [39].

3. **Hybrid**—In this instance, devices are allocated a mixture of shared audio material as well as unique assets for individual or groups of devices. Audio allocation here is determined through a pre-defined set of rules, which can take into account many factors such as device numbers, types of devices, and device locations. Examples of this include larger audio performances with many devices, where there are a number clusters of devices playing different content [13,38].

5.8. Summary

The themes presented in Section 4 provide an overview of the low-level design aspects of co-located, multi-device audio experiences; however, the number of themes and the complexity of the relationships between them can be overwhelming and, therefore, less practical. These design dimensions aim to provide a simplified framework for the core design considerations for these experiences by abstracting away some of the aforementioned complexity. As a result, the dimensions are likely best utilised in the early stages of experience design.

6. Application Patterns

The dataset obtained from the survey was re-analysed to explore the various current applications of co-located, multi-device audio experiences. For this exercise, each experience was assessed against each design dimension and given a respective category. The experiences and their categories were then compared to see how the different design choices affected the type of experience. Using this method, four example application categories were identified, where the experiences in each category share a common combination of categories. For each experience in the dataset, their descriptions, and corresponding

application types, please refer to Appendix C. Additionally, in the dimension breakdown table for each of the experience types, the key dimensions are highlighted in bold.

6.1. Public Performances

Key attributes

Synchronisation	Asynchronous/Loose/Tight
Context	Public
Position	Fixed/Flexible/Dynamic
Relationship	Many one-to-one/Many-to-many
Interactivity	Passive
Organisation	Non-hierarchical
Distribution	Mirrored/Hybrid

The largest category in the dataset consists of large-scale performances/installations in a public space. These experiences are akin to more traditional media experiences with largely passive audiences. These types of experiences mark the early experimental years of multi-device audio as a medium, many occurring prior to the 21st century, before 'personal' devices [13,37]. However, this is not always the case, notably where new forms of audio devices are being explored for their creative potential [51].

6.2. Interactive Music

Key attributes

Synchronisation	Asynchronous/Loose/Tight
Context	Social
Position	Flexible/Dynamic
Relationship	Many one-to-one/Many-to-many
Interactivity	Influence/Create
Organisation	Non-hierarchical
Distribution	Distinct

Interactive music contains interactive experiences that involve audience participation in music composition or performance. Most of the experiences exhibit either loose or tight synchronisation, as expected for the content medium, although a small number do display asynchronicity [53,58], where the musical content has less of an emphasis on rhythm and structure; for example, ambient music or soundscapes. Additionally, most of the experiences employ a non-hierarchical organisation of devices, ensuring that all participating listeners have the opportunity to make equal contribution to the experience. Where a hierarchical structure is utilised, it is typically in the form of a supporting loudspeaker system to provide better bass frequency reproduction, as seen, for example, in *Drops* [19]. Many of these experiences have been developed using the Soundworks framework [19], which provides tools for developing interactive experiences.

6.3. Augmented Broadcasting

Key attributes

Synchronisation	Tight Synchronisation
Context	Personal/Social
Position	Arbitrary
Relationship	One-to-many/Many one-to-one
Interactivity	Passive/Influence
Organisation	Hierarchical
Distribution	Hybrid/Distinct

A relatively new set of experiences has emerged over the last decade, that tend to extend a conventional, domestic TV or radio listening experience with additional devices for

more immersive spatial audio [21,22,52,55] or to allow the personalisation of content [48,49]. These experiences leverage the ubiquity of personal devices in the home, such as smartphones, tablets, and even newer wearables such as the Bose Frames [59]. Many of these experiences have been created through the *Audio Orchestrator* production tool, which allows separate audio allocation to the 'main' and 'auxiliary' devices via a configurable set of 'behaviours'. Elsewhere, German broadcaster Beyerischer Rundfunk (BR) has developed a smartphone app which communicates with smart TVs to allow the selection of audio description for selected programs, which can be heard through headphones connected to the smartphone [49].

6.4. Social Games

Key attributes

Synchronisation	Asynchronous/Loose
Context	Social/Public
Position	Dynamic
Interactivity	Influence/Create
Relationship	Many one-to-one
Organisation	Non-hierarchical
Distribution	Mirrored/Distinct/Hybrid

Finally, the last and smallest application category seen in the dataset contains shared, interactive audio games. These experiences generally involve listeners or players moving around with devices in a physical space and demonstrate audio-augmented reality (AAR) characteristics through the augmentation of a game with synthesised sound. Audio assets can be triggered and modified through positional and motion-based interactions. These experiences are limited to research studies for now, utilising bespoke, custom-made devices, but represent a vision for future possibilities. Examples include Bown and Ferguson's work with DIADS [36]; SoundWear [47], a study investigating the effect of non-speech sound augmentation on children's play; and *Please Confirm you are not a Robot* [60], a narrative-focused AAR game.

7. Discussion

7.1. Dimension Categories and Combinations

When characterising experiences using these dimensions, it is sometimes difficult (and counterproductive) to assign one given category. Many modern experiences are built with technologies that enable them to traverse between the categories of given dimensions, even during the experience. For example, some experiences could equally be assigned any of the categories in the *Context* and *Relationship* dimensions, due to the affordance of being able to add and remove devices. It may or may not be better to consider each of these variations separately; however, it is valuable to recognise the possibility that experiences can move along dimensions and to understand the capabilities of the technologies chosen to build a given experience that may enable this behaviour.

The design dimensions also contain a number of invalid category combinations, of which five have been identified. These are outlined below, where the dimension category is given alongside the dimension name in brackets.

1. One-to-many (Relationship) and Social (Context)
2. One-to-many (Relationship) and Public (Context)
3. Personal (Context) and Multiple one-to-one (Relationship)
4. Personal (Context) and Many-to-many (Relationship)
5. Mirrored (Distribution) and Create (Interactivity)

Combinations one to four exist due to both dimensions containing categories which represent a single listener experience and are incompatible with the shared experience categories of the comparative dimension. For combination five, if the audio content is

dependent on the creative process of an audience member, it is highly unlikely that the audio output of each device will be identical. Knowledge of these invalid combinations can help steer users of the design dimensions toward more useful outcomes.

7.2. Relationship with Audio-Augmented Reality

Another point of interest is the apparent application overlap between co-located, multi-device audio experiences and audio-augmented reality (AAR) experiences, when comparing with Krzyzaniak's six categories of AAR [61]. Specifically, experiences in the *Augmented broadcasting* and *Social games* experience groups illustrate AAR through two forms. *Augmented broadcasting* experiences tend to exhibit AAR through the provision of extra context-relevant audio content. In contrast, *Social games* use an enchanted objects approach to AAR by adding extra purposeful sound to typically silent objects. Given the understanding that AAR can be defined as real-world experiences that are accompanied by additional layers of sound under either form mentioned above, there is an inherent requirement for two or more sound sources, where either, some, or all of those sources could be audio-capable devices. The question then is: When are co-located multi-device audio experiences considered to be AAR experiences?

A good starting point is to evaluate the *Organisation* dimension, to consider the roles of the devices in the experience. In most 'non-hierarchical' experiences, where each device contributes to the audio equally, the experience content can be considered to be a sum of the output of all the devices. Another perspective is that the experience cannot be 'complete' without the array of devices, so there is not augmentation but equal complementation. Conversely, in many 'hierarchical' experiences, the core experience can be listened to on a single device, and the addition of new devices can introduce enhancements such as spatialised audio and additional audio content, which is consistent with AAR characteristics. An example of this is demonstrated in the BBC trial 'Immersive Six Nations' where the original program can be viewed on a main device, but additional devices grant access to extra audio assets such as on-pitch sound effects [22]. This is applicable for *Augmented Broadcasting* use-cases but is less useful for *Social games*, as is illustrated by Soundwear [47], where the device organisation is 'non-hierarchical', yet the purpose of the devices is to add additional audio to an otherwise ordinary outdoor play experience. Thus, when trying to characterise new experiences, it is proposed that the *Organisation* dimension should be used as a guide to determine whether a co-located, multi-device audio experience is an AAR experience. Consider, however the roles of the devices on an individual experience basis, as well as evaluating which form of AAR is exhibited in the experience.

7.3. Research Limitations

It is important to note that the dimensions presented here should not be treated as a fixed model, accounting for all possible design considerations, but should be considered as a helpful starting point for further discussion and refinement, through highlighting the most important current design factors. As research in this area continues, it is possible that new categories of experiences will emerge that cannot be sufficiently characterised by the proposed model, and so the design dimensions will have to be revised accordingly. This is in part due to the fact that the dataset of experiences obtained in this study is not an exhaustive list, and it is likely that there are other experiences, academic, commercial, or otherwise, that might be considered to be multi-device audio experiences under the inclusion criteria used in this research. However, this dataset represents the largest available sample of these specific experiences to date.

8. Conclusions

Research on multi-device experiences is steadily maturing, although, only a small proportion of it is focused on how multiple devices can be used to facilitate more immersive and engaging audio experiences. In addition, research efforts concerning audio tend to be disconnected, which is shown by the lack of consensus on the use of terminology. In this

research, an attempt has been made to link these research threads and find common ground through the evaluation of current co-located, multi-device audio experiences, driven by two research questions.

1. What are the defining characteristics of co-located, multi-device audio experiences?
2. How do these characteristics manifest in different applications?

A survey and structured literature review were used to produce data on which a thematic analysis was performed to determine a set of seven design dimensions. These are *synchronisation*, *context*, *position*, *relationship*, *interactivity*, *organisation*, and *distribution*. Furthermore, these dimensions were used to characterise and compare experiences in the dataset, leading to the identification of four example use-case categories of co-located multi-device audio experiences. These are *public performances*, *interactive music*, *augmented broadcasting*, and *social games*. More generally, this work provides an overview of research efforts from a practical use-case perspective, as well as acting as a guide for researchers and creative practitioners in furthering the development of these experiences.

Future Work

More work is necessary to test and validate the design dimensions and is likely to be an ongoing process. This can be achieved through assessing more experiences against the design dimensions and using the dimensions to explore new edge-case and uncharted applications by investigating unexplored combinations of the design dimension categories.

One of the major benefits of using personal devices for multi-device audio experiences is the accessibility afforded for shared, interactive audio experiences. In many instances of these, the in-built speakers of the devices are used, whereas headphones are rarely used. This could be due to the acoustic isolation effect on the listener, which may inhibit social interaction. In experiences with headphones, they are used predominantly for this acoustic isolation property but can be used with good effect in some social situations, where collaboration is involved, as evidenced in *Schminky* [50] and *ProXoMix* [46]. Future work could explore the implications of the use of different types of headphones (i.e., closed-back, open-back, bone conduction, transparent-mode) and their effect on social interaction, for various types of multi-device audio experiences.

Another potential area yet to be explored is the use of smart speakers, and voice interaction more generally, in these experiences. This may be due to the difficulties encountered by speech-recognition software when in the presence of multiple, competing sound sources. There may also be more development difficulties associated with smart speakers than compared to an open-source ecosystem such as Raspberry Pi. Nevertheless, future work could explore the creative use and effectiveness of voice interaction in co-located, multi-device audio experiences.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Multi-Device Audio Survey

Appendix A.1. Introduction

Welcome, and thank you for taking part in the following survey. We greatly appreciate the time and effort you're taking to complete this survey, and your contribution will significantly aid our research in exploring the landscape of multi-device audio experiences.

Appendix A.2. Aim

The aim of this survey is to use crowdsourcing to obtain a dataset of multimedia experiences and technologies that utilise multiple devices to reproduce audio, referred to in this study as 'multidevice audio experiences'. The dataset will act as a supplementary resource for further research into the area including supporting characterisation and analysis of multi-device audio experiences, as well as aiding the development of a set of design criteria for multi-device audio experiences. The resulting dataset may be openly published at an unspecified time after the conclusion of the survey.

Appendix A.3. Instructions

We would like you to tell us about a multi-device audio experience that you are aware of, whether you have experienced it yourself or not. We will ask you some questions about this in a short survey, which should take no longer than 5 min to complete. Your task is to answer each of the questions relating to your specific entry. You are encouraged to provide as much detail about your entry as possible using the free text boxes. If you are aware of more than one multi-device audio experience, please feel free to complete the survey for each entry.

Appendix A.4. Criteria

The term 'multi-device audio experiences' is a broad term which encompasses a wide range of technologies and use-cases. We're collecting experiences that use multiple devices for playing audio, where all playing devices can be perceived by the listener. We would like to obtain information on:

- products, platforms or systems that support and deliver multi-device audio experiences (Example: a SONOS sound system)
- single experiences, or pieces of content which include multiple devices playing audio (Example: A silent disco)

The inclusion criteria in more detail are provided below. Ideally, your entry must fit all three criteria.

1. The Platform/Experience must employ multiple devices with loudspeakers.
   Explanation: Essentially, any platform or experience that includes at least two devices that have loudspeakers (including any type of headphones). This aims to encapsulate any devices capable of playing audio, this includes but is not limited to smartphones, tablets, laptops, TV's, Bluetooth speakers and smart speakers.
2. The audio content must be distributed across the devices. Explanation: The experience must be capable of playing audio over multiple devices, whether that's simultaneously or intermittently. An example that does not fit this criterion is streaming audio content from one device to another using AirPlay or Chromecast.
3. The devices must be co-located in a single space, or group of adjoining spaces.
   Explanation: This final specification aims to make a distinction between a loudspeaker

array where devices are within the same vicinity, and a group of individual devices remotely connected over the Internet. For example, the devices can be located either in a single room, or across multiple rooms in a building. This criterion excludes teleconferencing platforms such as Zoom or Skype and traditional phone calls.

The criteria will be displayed at the bottom of each page of questions for your reference. This test has been designed to follow the ethical codes and practices laid out by the University of York. This test is anonymised and all data, personal or otherwise, is confidential. This is a confidential and anonymous study, and you can withdraw at any time for any reason. This research is being undertaken by David Geary (drg519@york.ac.uk). If you have any queries, please let them know by email.

I have read and acknowledged all of the above and agree to participate:

- Yes
- No

Appendix A.5. Questions

1. Would you describe your entry as a platform, or a single experience?

    - Platform—a system which supports and delivers experiences
    - Experience—a single experience which involves multiple devices playing audio
    - Don't know

2. What is the name of your entry?
3. Please provide a brief description of your entry.
4. Describe the audio content of your entry. Specifically, what does the audio content consist of? How is it distributed across the devices?
5. Which devices are a part of the platform/experience ecosystem?

    - Mobile Phones/Smartphones
    - Tablets
    - Desktop/Laptop
    - Smart speaker
    - Radio
    - TV
    - Other

6. How do participants interact with the platform/experience? For example, how are the devices interacted with?

    - Tactile Input (Keyboard/Mouse/Buttons)
    - Touch screen
    - Motion gesturing
    - Voice commands
    - Not interactive
    - Other

7. For your platform/experience, is the number of involved devices variable? For example, can devices be added/removed from the platform/experience freely?

    - Yes
    - No
    - Don't know

8. Do some devices in the platform/experience have a greater importance than others? For example, one device may be responsible for playing most of the audio content, while the other devices play a supporting role.

    - Yes
    - No
    - Don't know

9. Can you provide any links to further information about your entry? (e.g., videos, webpages, publications, etc.)
10. Is there anything else about your entry that you'd like to tell us?

Appendix B

Table A1. Multiple choice questions and responses for the multi-device audio experiences and technologies survey. The numbers and percentages of responses are given in parentheses.

Question	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6	Option 7
Would you describe your entry as a platform, or a single experience? (43)	Platform (14, 33%)	Experience (28, 65%)	Don't know (1, 2%)	—	—	—	—
Which devices are a part of the platform/experience ecosystem? (78)	Mobile phones/smartphones (26, 34%)	Tablets (12, 15%)	Desktop/laptop (12, 15%)	Smart speaker (4, 5%)	Radio (3, 4%)	TV (3, 4%)	Other (18, 23%)
How do participants interact with the platform/experience? (68)	Tactile input (22, 32%)	Touch screen (21, 31%)	Motion gesturing (13, 19%)	Voice commands (2, 3%)	Not interactive (7, 10%)	Other (3, 5%)	—
For your platform/experience, is the number of involved devices variable? (43)	Yes (36, 84%)	No (5, 12%)	Don't know (2, 4%)	—	—	—	—
Do some devices in the platform/experience have a greater importance than others? (43)	Yes (11, 26%)	No (25, 58%)	Don't know (7, 16%)	—	—	—	—

Appendix C. Full Experiences Table

Table A2. All the individual experiences in the dataset, with descriptions, the type of experience, and the associated reference.

Name	Description	Experience Type	Reference
Fields	Collaborative ambient-style music performance using mobile devices.	Public performances	[5]
Ugnayan	Numerous audio tapes transmitted over different radio channels, replayed on personal radios across an urban space.	Public performances	[13]
Cassettes 100	100 performers walk around with cassette platyers playing different instrumental and environmental sound.	Public performances	[14]
So Predicatable?!	An improvisational dance piece using two Raspberry Pi devices with loudspeakers (DIADs).	Public performances	[23]
Babel	An art installation of a tower of analogue radios all playing simultaneously.	Public performances	[35]
Boombox Experiment	An experiment conducted by the Flaming Lips where 100 audience members, each with their own tape player, played music tapes in synchrony.	Public performances	[37]
Dialtones	Art installation involving mobile phones which played a selection of ringtones.	Public performances	[38]
Bloom	Hundreds of Raspberry Pi devices with loudspeakers (DIADs) in outside installation at Kew Gardens.	Public performances	[51]
Movies on the Meadows	Silent film screening on large screens in fields, with audience members wearing headphones.	Public performances	[56]
Parking Lot Experiments	An experiment by the Flaming Lips where 30 individual tapes where played at the same time on 30 separate car tape deck systems.	Public performances	[62]
Manifolds	Audience in a gallery use their cellphones to project the sound installation's voices into the room. The audience becomes a moving loudspeaker orchestra.	Public performances	[63]
Collective Loops	Participatory music experience where each player controls a segment of a sequencer and can add musical elements to their segment.	Interactive music	[39]
Mesh Garden	Exploratory music-making game for smartphones, creating a piece of distributed ambient music between participants.	Interactive music	[42]
One-Man Orchestra	A human conductor uses motion gestures to control individual smartphones playing music.	Interactive music	[43]
Echobo	Participatory music experience where a central performer and audience members collaborate to create music.	Interactive music	[44]
Pick A Part	An experience built using Audio Orchestrator which enables individual devices to choose different instruments of a classical music piece.	Interactive music	[45]

Table A2. *Cont.*

Name	Description	Experience Type	Reference
ProXoMix	Participatory music experience created with Soundworks. Uses bluetooth beaconing so when devices move closer to each other, you can hear other devices music tracks.	Interactive music	[46]
Schminky	Interactive musical game using HP iPaqs and headphones.	Interactive music	[50]
Crowd in c[loud]	Participatory music game where audience members compose melodies on their own devices. These melodies can be shared via profiles and other members can listen and like them, like a social media experience.	Interactive music	[53]
Drops	Participatory music experience created using the Soundworks framework. Triggered sound sources echo across multiple devices.	Interactive music	[54]
GrainField	An individual performer is recorded, every second of recording is sent randomly to audience devices and replayed using a granular synthesizer to create an ambient complimentary backing track.	Interactive music	[58]
The Vostok-K Incident	An immersive 3D audio drama created using Audio Orchestrator.	Augmented broadcasting	[21]
Immersive Six Nations Rugby	Sports programme experience using multiple devices to add additional audio to a highlights show of a six nations rugby match.	Augmented broadcasting	[22]
Spectrum Sounds	Distributed musical experience where short pieces of classical music are spread across a minimum of 3 connected devices.	Augmented broadcasting	[41]
Augmented TV Viewing	A study exploring the application of acoustically transparent Bose Frames to improve TV viewing by intermixing TV and headset audio.	Augmented broadcasting	[48]
BR Audiodeskription App	A mobile app that can deliver audio description to the users connected headphones, while the TV plays the standard broadcast mix.	Augmented broadcasting	[49]
Decameron Nights	Multi-device halloween themed audio story programme.	Augmented broadcasting	[52]
Monster	A halloween themed drama episode over a minimum of three devices.	Augmented broadcasting	[55]
Musical Bowls	Musical game of bowls using DIADs, bespoke Raspberry Pi devices with loudspeakers and sensors.	Social games	[36]
SoundWear	Devices consist of a wearable wrist device which can play audio, designed as a play toy for children. Sounds can be transferred to other devices by aligning them.	Social games	[47]
Please Confirm You Are Not A Robot	Multiplayer AAR game experience using Bose Frames to deliver personalised audio.	Social games	[60]

References

1. Rieger, C.; Majchrzak, T.A. A Taxonomy for App-Enabled Devices: Mastering the Mobile Device Jungle. In Proceedings of the Web Information Systems and Technologies, Seville, Spain, 18–20 September 2018; Springer International Publishing: Berlin/Heidelberg, Germany, 2018; pp. 202–220.
2. Atzori, L.; Iera, A.; Morabito, G. The Internet of Things: A survey. *Comput. Netw.* **2010**, *54*, 2787–2805.
3. Brudy, F.; Holz, C.; Rädle, R.; Wu, C.J.; Houben, S.; Klokmose, C.N.; Marquardt, N. 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, Glasgow, UK, 4–9 May 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1–28.
4. Wang, G.; Bryan, N.J.; Oh, J.; Hamilton, R. Stanford laptop orchestra (slork). In Proceedings of the 2009 International Computer Music Conference (ICMC), Montreal, QC, Canada, 16–21 August 2009.
5. Shaw, T.; Piquemal, S.; Bowers, J. Fields: An Exploration into the Use of Mobile Devices as a Medium for Sound Diffusion. In Proceedings of the International Conference on New Interfaces for Musical Expression, NIME 2015, Baton Rouge, LA, USA, 31 May–3 June 2015; The School of Music and the Center for Computation and Technology (CCT), Louisiana State University: Baton Rouge, LA, USA, 2015; pp. 281–284.
6. Levin, M. *Designing Multi-Device Experiences: An Ecosystem Approach to User Experiences across Devices*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2014.
7. Bown, O.; Ferguson, S. Understanding media multiplicities. *Entertain. Comput.* **2018**, *25*, 62–70.
8. Vinayagamoorthy, V.; Ramdhany, R.; Hammond, M. Enabling Frame-Accurate Synchronised Companion Screen Experiences. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX'16), Chicago, IL, USA, 22–24 June 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 83–92.
9. Neate, T.; Jones, M.; Evans, M. Cross-device media: A review of second screening and multi-device television. *Pers. Ubiquit. Comput.* **2017**, *21*, 391–405.

10. Lohmüller, V.; Wolff, C. Towards a Comprehensive Definition of Second Screen. In Proceedings of the Mensch und Computer 2019 (MuC'19), Hamburg, Germany, 8–11 September 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 167–177.

11. Vatavu, R.D.; Saeghe, P.; Chambel, T.; Vinayagamoorthy, V.; Ursu, M.F. Conceptualizing Augmented Reality Television for the Living Room. In Proceedings of the ACM International Conference on Interactive Media Experiences (IMX'20), Barcelona, Spain, 17–19 June 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–12.

12. Saeghe, P.; Abercrombie, G.; Weir, B.; Clinch, S.; Pettifer, S.; Stevens, R. Augmented Reality and Television: Dimensions and Themes. In Proceedings of the ACM International Conference on Interactive Media Experiences (IMX'20), Barcelona, Spain, 17–19 June 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 13–23.

13. Maceda, J. *Ugnayan*; Tzadik: New York, NY, USA, 2009.

14. Team, E. Cassettes 100 by José Maceda, Re-Staged at National Gallery Singapore | National Gallery Singapore. 2020. Available online: https://www.nationalgallery.sg/magazine/cassettes-100-jose-maceda-at-national-gallery-singapore (accessed on 12 April 2022).

15. Taylor, B. A history of the audience as a speaker array. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2017), Copenhagen, Denmark, 15–18 May 2017; pp. 481–486.

16. Trueman, D.; Cook, P.R.; Smallwood, S.; Wang, G. PLOrk: The Princeton laptop orchestra, year 1. In Proceedings of the 2016 International Computer Music Conference (ICMC), New Orleans, LA, USA, 6–11 November 2006.

17. Dahl, L. Wicked Problems and Design Considerations in Composing for Laptop Orchestra. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2012), Ann Arbor, MI, USA, 21–23 May 2012.

18. Lind, A.; Yttergren, B.; Gustafson, H. Challenges in the Development of an Easy-Accessed Mobile Phone Orchestra Platform. 2021. Available online: https://webaudioconf2021.com/wp-content/uploads/2021/06/MPO-paper_WAP21.pdf (accessed on 20 August 2021).

19. Robaszkiewicz, S.; Schnell, N. Soundworks—A playground for artists and developers to create collaborative mobile web performances. In Proceedings of the WAC—1st Web Audio Conference, Paris, France, 26–27 January 2015.

20. Hentschel, K.; Francombe, J. Framework for web delivery of immersive audio experiences using device orchestration. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX'19), Salford, UK, 5–7 June 2019.

21. Francombe, J.; Woodcock, J.; Hughes, R.J.; Hentschel, K.; Whitmore, E.; Churnside, A. Producing Audio Drama Content for an Array of Orchestrated Personal Devices. In Proceedings of the Audio Engineering Society Convention 145, New York, NY, USA, 17–20 October 2018.

22. Francombe, J. Get to the Centre of the Scrum—Our Immersive Six Nations Rugby Trial—BBC R&D. 2021. Available online: https://www.bbc.co.uk/rd/blog/2021-02-synchronised-audio-devices-sound-immersive (accessed on 7 March 2022).

23. Fraietta, A.; Bown, O.; Ferguson, S.; Gillespie, S.; Bray, L. Rapid composition for networked devices: HappyBrackets. *Comput. Music J.* **2020**, *43*, 89–108.

24. Kim, H.; Lee, S.; Choi, J.W.; Bae, H.; Lee, J.; Song, J.; Shin, I. 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), Seattle, WA, USA 13–17 September 2014; Association for Computing Machinery: New York, NY, USA, 2014; pp. 277–288.

25. do Nascimento, F.J.B.; de Souza, C.T. Features of Second Screen Applications: A Systematic Review. In Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web (Webmedia'16), Teresina, Brazil, 8–11 November 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 83–86.

26. Turchet, L.; Barthet, M. An ubiquitous smart guitar system for collaborative musical practice. *J. New Music Res.* **2019**, *48*, 352–365.

27. Qualtrics. Qualtrics XM//The Leading Experience Management Software. 2022. Available online: https://www.qualtrics.com/uk/?rid=ip&prevsite=en&newsite=uk&geo=GB&geomatch=uk (accessed on 19 April 2022).

28. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann. Intern. Med.* **2009**, *151*, 264–269.

29. ACM. ACM Digital Library. 2022. Available online: https://dl.acm.org/ (accessed on 19 April 2022).

30. Audio-Engineering-Society AES E-Library. 2022. Available online: https://www.aes.org/e-lib/ (accessed on 19 April 2022).

31. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101.

32. Kim, H.; Sefcik, J.S.; Bradway, C. Characteristics of Qualitative Descriptive Studies: A Systematic Review. *Res. Nurs. Health* **2017**, *40*, 23–42.

33. Cieciura, C.; Mason, R.; Coleman, P.; Paradis, M. Survey of media device ownership, media service usage, and group media consumption in UK households. In Proceedings of the Audio Engineering Society Convention 145, New York, NY, USA, 17–20 October 2018.

34. Striner, A.; Azad, S.; Martens, C. A Spectrum of Audience Interactivity for Entertainment Domains. In Proceedings of the International Conference on Interactive Digital Storytelling, Little Cottonwood Canyon, CT, USA, 19–22 November 2019; Springer International Publishing: Berlin/Heidelberg, Germany, 2019, pp. 214–232.

35. Barson, T. 'Babel', Cildo Meireles, 2001 | Tate. 2011. Available online: https://www.tate.org.uk/art/artworks/meireles-babel-t14041 (accessed on 24 March 2022).

36.   Bown, O.; Ferguson, S. A Musical Game of Bowls Using the DIADs. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2016), Brisbane, Australia, 11–15 July 2016; pp. 371–372.

37.   Kaufman, G. Flaming Lips' Bizarre 'Boombox Experiment'—MTV. 1998. Available online: https://www.mtv.com/news/3439/flaming-lips-bizarre-boombox-experiment/ (accessed on 24 March 2022).

38.   Levin, G.; Shakar, G.; Gibbons, S.; Sohrawardy, Y.; Gruber, J.; Semlak, E.; Schmidl, G.; Lehner, J.; Feinberg, J. *Dialtones (A Telesymphony)*; Ars Eletronica Festival; Ars Eletronica: Linz, Austria, 2001.

39.   Schnell, N.; Matuszewski, B.; Lambert, J.P.; Robaszkiewicz, S.; Mubarak, O.; Cunin, D.; Bianchini, S.; Boissarie, X.; Cieslik, G. Collective Loops: Multimodal Interactions Through Co-Located Mobile Devices and Synchronized Audiovisual Rendering Based on Web Standards. In Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction (TEI '17), Yokohama, Japan, 20–23 March 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 217–224.

40.   Lambert, J.P.; Robaszkiewicz, S.; Schnell, N. Synchronisation for distributed audio rendering over heterogeneous devices, in html5. In Proceedings of the WAC 2016: 2nd Annual Web Audio Conference, Atlanta, GA, USA, 4–6 April 2016.

41.   BBC Research and Development. Spectrum Sounds: A Lockdown Audio Composition in Seven Colours—BBC R&D. 2021. Available online: https://www.bbc.co.uk/rd/blog/2021-09-hearing-synaesthesia-audio-composition-music (accessed on 12 April 2022).

42.   Carson, T. Mesh Garden: A creative-based musical game for participatory musical performance. In Proceedings of the NIME 2019, Porto Alegre, Brazil, 3–6 June 2019.

43.   Tsui, C.K.; Law, C.H.; Fu, H. One-Man Orchestra: Conducting Smartphone Orchestra. In Proceedings of the SIGGRAPH Asia 2014 Emerging Technologies (SA'14), Shenzhen, China, 3–6 December 2016; Association for Computing Machinery: New York, NY, USA, 2014;

44.   Lee, S.W.; Freeman, J. echobo: A mobile music instrument designed for audience to play. In Proceedings of the 2013 International Conference on New Interfaces for Musical Expression (NIME 2013), Daejeon, Korea, 27–30 May 2013.

45.   Francombe, J. Conduct Your Own Orchestra? Pick a Part with the BBC Philharmonic!—BBC R&D. 2020. Available online: https://www.bbc.co.uk/rd/blog/2020-09-proms-synchronised-audio-music-interactive (accessed on 24 March 2022).

46.   Schnell, N.; Matuszewski, B.; Poirier-Quinot, D.; Audouze, V. ProXoMix, Embodied Interactive Remixing. In Proceedings of the Audio Mostly, London, UK, 23–26 August 2017.

47.   Hong, J.; Yi, H.; Pyun, J.; Lee, W. 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), Eindhoven, The Netherlands, 6–10 July 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 2201–2213.

48.   McGill, M.; Mathis, F.; Khamis, M.; Williamson, J. Augmenting TV Viewing using Acoustically Transparent Auditory Headsets. In Proceedings of the ACM International Conference on Interactive Media Experiences (IMX'20), Barcelona, Spain, 17–19 June 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 34–44.

49.   Rundfunk, B. BR Audiodeskription. Available online: https://apps.apple.com/gb/app/br-audiodeskription/id1568128729 (accessed on 25 March 2022). (In German)

50.   Reid, J.; Hull, R.; Melamed, T.; Speakman, D. Schminky: The design of a café based digital experience. *Pers. Ubiquit. Comput.* **2003**, *7*, 197–202.

51.   Ferguson, S.J.; Rowe, A.; Bown, O.; Birtles, L.; Bennewith, C. Sound design for a system of 1000 distributed independent audio-visual devices. In Proceedings of the International Conference on New Interfaces for Musical Expression 2017, Copenhagen, Denmark, 15–18 May 2017.

52.   Young, E. Decameron Nights: Immersive and Interactive with All Your Audio Devices—BBC R&D. 2020. Available online: https://www.bbc.co.uk/rd/blog/2020-08-audio-drama-surround-sound-spatial (accessed on 26 March 2022).

53.   Lee, S.W.; de Carvalho, A.D., Jr.; Essl, G. Crowd in c[loud]: Audience participation music with online dating metaphor using cloud service. In Proceedings of the WAC 2016: 2nd Annual Web Audio Conference, Atlanta, GA, USA, 4–6 April 2016.

54.   Matuszewski, B.; Schnell, N.; Bevilacqua, F. Interaction Topologies in Mobile-Based Situated Networked Music Systems. *Wirel. Commun. Mob. Comput.* **2019**, *2019*, 9142490.

55.   Young, E.; Francombe, J. Monster: A Hallowe'en Horror Played out through Your Audio Devices—BBC R&D. 2020. Available online: https://www.bbc.co.uk/rd/blog/2020-10-audio-drama-monster-interactive-sound (accessed on 28 March 2022).

56.   Movies on the Meadows 2019 | Cambridge Film Festival. 2019. Available online: https://www.cambridgefilmfestival.org.uk/news/movies-meadows-2019 (accessed on 12 April 2022).

57.   Sonos. Sonos | Wireless Speakers and Home Sound Systems. 2022. Available online: https://www.sonos.com/en-gb/home (accessed on 5 May 2022).

58.   Matuszewski, B.; Schnell, N. GrainField. In Proceedings of the Audio Mostly, London, UK, 23–26 August 2017.

59.   Bose. Wearables by Bose—Classic Bluetooth® Audio Sunglasses. Available online: https://www.bose.co.uk/en_gb/products/frames/bose-frames-alto.html (accessed on 19 April 2022).

60.   Nagele, A.N.; Bauer, V.; Healey, P.G.T.; Reiss, J.D.; Cooke, H.; Cowlishaw, T.; Baume, C.; Pike, C. Interactive Audio Augmented Reality in participatory performance. *Front. Virtual Real.* **2021**, *1*, 610320.

61.   Krzyzaniak, M.; Frohlich, D.; Jackson, P.J.B. 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), Nottingham, UK, 18–20 September 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 160–167.

62. Lips, F. 1996–1997: The Parking Lot Experiments. Available online: https://christophercatania.com/2009/03/18/a-sxsw-tribute-flaming-lips-parking-lot-experiment-no4/ (accessed on 12 April 2022).
63. Tammen, H. Manifolds. 2020. Available online: https://radiopanspermia.space/manifolds/ (accessed on 12 April 2022).