Improving Digital Accessibility Through Audio-Game Co-Design

Zach Mason
Lancaster University
z.mason@lancaster.ac.uk

David Philip Green,
Joseph Lindley, Paul Coulton
d.p.green@lancaster.ac.uk, j.lindley@lancaster.ac.uk
p.coulton@lancaster.ac.uk

ABSTRACT
Globally over 2.2 billion people are blind or visually impaired (World Health Organisation, 2022). Roughly half of these visual impairments cannot be prevented or addressed. Games are no longer just for fun, and as virtual environments often referred to as Metaverses and Meta-spaces continue to emerge, cyberspace is becoming increasingly spatial and important to everyday life. These virtual spaces do not afford tactile interaction. This presents an accessibility problem, particularly for visually impaired people, many of whom rely on haptic feedback to navigate the spatial environment of the real world. The games industry has a rich history of innovation in accessible virtual spaces, with games such as Sightlence (Nordvall, 2013), Papa Sangre (Somethin’ Else, 2010) and Blind Legend (Dowino, 2015) being exemplary of non-visual-first approaches. Games are excellent sandboxes for exploring virtual accessibility for people without a visual-first understanding. In this paper, we present an analysis of a workshop that engaged blind and visually impaired people in the first stage of a co-design process to develop a new audio-only game which will focus on spatial-audio (Frauenberger, Noistering, 2003) through generic low-end game hardware. The workshop was facilitated in collaboration with sight loss charities, with the objective of exploring how to create more inclusive virtual spaces from their conceptions (McElligott, Leeuwen, 2004). Through synthesis of these conversations, we bring to light a desire from blind and visually impaired people to include non-visual accessibility as a core consideration in the design of virtual space software and hardware to support the future of non-visual cyberspace access.

Keywords
Accessibility, Audio-Only, Co-Design, Inclusion, Visual Impairment

RESEARCH RATIONAL
As peoples’ need for a virtual sense of togetherness increases, cyberspace is becoming increasingly spatial in its configuration. Web conferencing systems like Gather (Gather, 2020), as well as other similar systems (HyHyve, 2022) have adopted spatial presence as a mechanism for initiating smaller conversations and attempting to make them ‘feel’ more natural. Such spatial approaches are still in their exploratory stages and, while facilitating fluid spatial conversations, this space-based approach can prove problematic for accessibility (Mason et al. 2022). This is due, in part, to the lack of tactile or haptic feedback available through standard digital interfaces. The combination of the spatial design metaphor that these platforms use, and the types of interfaces which are generally available to visually impaired people in the physical
world but not available in these platforms, makes such virtual spaces almost unusable for blind and visually impaired people. Video games as an evolving medium have always explored novel ways of evoking and utilising the notion of space inside the game world. However—perhaps most vividly demonstrated by the name video game—they have tended to have a visual-first approach. Most games are entirely playable without sound which is great for deaf or people with hearing loss. However, this means sound is often added purely to embellish and draw attention to certain events occurring visually onscreen. Haptics—mechanical systems that allow users to touch and feel virtual game elements—are utilised even less for core game mechanics because game developers want to produce equitable experiences across all platforms. Because most PC players use keyboard and mouse input, any game designed for PC must make haptics an optional sensory output, which proves problematic when designing games for platforms which blind and visually impaired people desire ready access to.

In this research we build on the premise that designing games as sandboxes can provide insights relating to emerging accessibility issues in digital spaces more generally. Specifically, we explore the use of sound to create a sense of space or Auditory Spatial Perception in an audio-only game. Our work seeks to understand an alternative to visual-first game design while also gaining insights that could be beneficial for the future of accessible metaverse design.

Virtual games that use sound or haptics as their primary output are starting to appear more frequently with titles such as *Sightlence*, *Papa Sangre*, *The Vale* (Falling Squirrel, 2021) and *Blind Legend* coming out in recent years. Other games like *The Last of Us Part 2* (Naughty Dog, 2020), *Mortal Combat 11* (NetherRealm Studios, 2019) and *Forza Horizon 5* (Playground Games, 2021), which are all AAA titles that use a visual-first approach but employ rich accessibility options (such as remapping controls, tying audio queues to game events and spoken menus) which are also contributing to the platforming of sensory accessibility in game design. Through facilitation of a Co-Design workshop in collaboration with sight loss charities, we aimed to amplify the voices of blind and visually impaired users who have an everyday need for accessibility in visually dominated media. People like *SightlessKombat* (a blind gamer) who helped with accessibility on *God of War Ragnarök* (Santa Monica Studio, 2022) are more commonly being included in the development of large-studio games. In contrast we saw it important to develop auditive games for people who are not experienced in non-visual digital navigation, in order to understand how we can maximise the accessibility of the tools we design.

Many of the audio games released in recent years have followed a trend of focusing on narrative elements within the game’s design due to the lack of spatial time sensitive movement required in genres like choose your own adventure, for example *Real Sound: Kaze no Regret* (Warp, 1997). While this makes the games accessible to the players at all ability levels, there is a danger that they can turn into ‘audio book’ versions of games. So-called ‘gamebooks’ can have rather limited interactivity and often don’t represent a gaming challenge for most players. Accordingly, there is a lowered replay potential or skill development. This means that for visually impaired players the passion for gaming that would emerge from re-playable and skill-based games rarely occurs. For this reason, we decided to focus our games design workshop on how the players were able to interact with the game space in ways that did not rely on sight (Targett, Fernstrom, 2003). Isolating the interaction mechanics of the game would allow narrative elements to be considered later in the game’s production and our focus was more specifically on movement, controls, Auditory Spatial-Perception (Grimshaw, Schott, 2007) and balanced difficulty. By tapping into increasingly advanced audio engines like the one used in *Demon Souls* (Bluepoint Games, 2020) we aimed to make
a game which is challenging for beginners, with an enjoyable learning curve that could translate to wider digital applications.

Limiting the output aspect of the game’s interactions to audio only enabled us to see how sound alone could generate a sense of space without any need for haptics. While we acknowledge haptics and movement-based interaction devices (Röber, 2005) are useful in providing sensory information, we wanted our experiences to be provided through generic controllers and output devices (such as headphones and arcade-joysticks, or keyboard and mouse). With a high possibility of audio games releasing on desktop computers where haptics feedback is unavailable through standard keyboard and mouse, we thought it important to explore audio games without haptics in the first instance. To further focus the feedback, the game experience avoided menus for the workshop, putting the players straight into the gameplay experience. There is an abundance of information around non-visual menu navigation available for us to draw upon for a potential future release of our game, such as avoiding infinite scrolling systems (i.e. when you reach the bottom of a menu, the cursor returns to the first option on one further downward input) and including dictated menus (Barlet, Spohn, 2012).

**DESIGNING GAMEPLAY**

The game’s instructions are segmented into three short statements that are dictated using a text-to-speech engine. The game’s concepts are introduced through the first few levels with all other information being conveyed through software-generated stereo audio that is played through a standard set of headphones. The game uses a joystick for workshop purposes but is also playable with keyboard alone (as mouse-based character movement can be hard to keep track of non-visually) and could be quickly adapted to function with a generic game controller. Using the joystick was intended to make the necessary peripheral easily accessible and affordable, as well as give the game a more familiar entry point that using a mouse and keyboard would not have achieved. We note that while any given game may be designed for sensory accessibility, it is also important to have monetary affordability. This is something that audio only games can achieve quite easily when compared to the resource-intensive visual fidelity of many modern games; with audio games there is no need for expensive graphic cards or games console hardware, with lower end computers being more than sufficient to play them.

![Figure 1: Diagram of game level elements](image)

The game was designed and produced using Unreal Engine 5 (UE5) which was in late-stage alpha at the time of the workshop. We used this engine due to its popularity in mainstream game production alongside rich features for generating sounds within the engine itself and off-the-shelf spatial sound functionality (although this has some limitations when compared to real world binaural audio recordings). Using UE5 not
only sped up production of the game but enabled us to explore the engine’s suitability for sound focused game design.

The game, created by the first author, was developed as a maze explorer game with linear level design and inspired by games like *Pac-Man* (Namco, 1980). This genre choice was supported by several comments made in virtual meetings prior to the in-person workshop, with many other potential audio games being envisioned during the workshop’s conversations. This genre of game proved great for pushing the player to move in varied ways, forcing them explore the environment with the aim of encouraging them to build a sense of game space throughout (in our case via sound feedback). Moreover, this game genre afforded a learnable playstyle, an important factor for entry level audio-only and accessible games (Andrade et al. 2019; Atkinson et al. 2006).

The game was produced with 15 levels for the Co-Design workshop with each level adding additional obstacles and elements to increase the difficulty. Through the testing process it was noted that the game played similarly to some titles in *bit Generations* (Skip Ltd, 2006) created for Game Boy Advanced and later released on WiiWare and DSiWare, a game which did rely on some sighted segmented and was heavily limited by hardware. For the first 2 levels of our game, the player is introduced to the goal sound in isolation. The aim of this was to enable players to get accustomed to positioning a singular sound while playing. Level 3 introduces the concept of walls as non-aggressive obstacles with a single corner to navigate around adding to a gradual increasing level of complexity and requiring the player to differentiate several sound elements. The fifth level introduced enemies. If a player bumps into these adversaries, the player is reset to the start point of the current level. Enemies are characterised by a ‘negative’ sound when they are nearby and/or bumped into.

The game features both proximity-based sounds within the space that play when close by and active sounds that play when the object is touched. The proximity sounds are spatialised using stereo audio and get louder as you approach them, playing louder in one channel (left or right) dependant on their position in virtual space. There is a limit to how quiet the endpoint can be as it acts as an audio beacon for wayfinding, so the player does not lose track of their position entirely within the level even if they move infinitely far away from it. This showcases how the effects of sound can be tweaked from those experienced in the physical world. Each active sound has a corresponding game-state change, with the endpoint sound starting the next level, the enemy sound restarting current level and the wall contact sound bouncing the player away from them. At the end of the final level, the player also receives a spoken notice telling them they beat the game, something which several players came very close to and lead them to express a strong desire to play the game more once published (even if it were to be a paid experience, even though we plan to release the game for free). Moving forward with the Co-Design process for this game we aim to implement FMod or Wwise (audio plugins) to increase the auditive fidelity further and increase the information perceivable through sound alone beyond what UE5’s built in sound engine can achieve.

**WORKSHOP RECRUITING AND STRUCTURE**

Workshop recruitment was a significant part of this research due to the importance of finding participants who had lived experience with significant visual impairments. It was highly valuable to get insight from people who not only had experience living with visual impairments, but also had an interest in games or broader virtual interactive experiences. We recruited participants by collaborating with *SASL* (Sight Advice South Lakes, 1956) who are a sight loss charity local to our research centre. Through our partnership they connected us with a group of people from across the UK, leading to the inclusion of *Galloways* (Galloways, 1867), another nearby sight loss charity.
The research was organised as a full day Co-Design workshop in Manchester with 12 participants (a range of intermediate to beginner visually impaired gamers) plus the workshop facilitators. Co-Design (Brewer, 2018; Feng, 2016) proved to be an incredibly powerful tool in developing our understanding of this design challenge by involving our participants more deeply within the design process rather than relying on our interpretation of their needs when deciding the output (Magnusson et al. 2018). The day’s focus revolved around getting the participants to take part in three different interactive experiences. Dividing up each of these interactive sections were three thirty-minute segments of open conversation around the participants’ opinions on the experiences they had just interacted with. The first of these segments had each of the participants trying a binaural haircut audio experience. This was intended to immerse them within the space of sound, and subtly nudge them towards thinking about distinguishing position of objects through stereo audio played through headphones.

The second section of the workshop was their first experience of playing the game, seen in Figure 2. We aimed to give as little as possible away about the game itself to capture their initial reactions to our gameplay (even though their discussion in several preliminary online calls had directed the games development from a very early stage). Participants were guided through the use of the controller (which was a traditional arcade joystick with 4 buttons). Through 4 different setups, each with 2 sets of headphones and a single joystick, 4 participants were able to play, and another 4 were able to listen in and help guide with the experience of playing. A third person was able to audibly perceive their reactions, giving each workshop attendee three perspectives on the gameplay experience and sufficient time to absorb the game’s elements and details. During their time playing through the game, 12 of the 15 levels were beaten by workshop participants with most players reaching around level 8. This meant all players were able to experience all the games mechanics, and the designed game was understood, played and then able to be reflected on. This was ideal for our research purposes suggesting the game was appropriately difficult and importantly learnable without sight. It is very reasonable to assume that the participants would have continued to beat further levels with more time due to the scaling of the levels in the game.
During the final section of the day workshop, we asked the participants to make a LEGO map of a future level they would want to play, either designed for the game they played, or another audio-centric game (seen in Figure 3) which they elaborated on in the following discussion. These were useful for seeing how the participants would like the game to change in further development and demonstrated how the participants understood our game as space by asking them to translate its mechanics and sensory modality from audio into a tactile and visual experience.

**WORKSHOP OBSERVATIONS & OUTPUTS**

After each session (the haircut audio, playing our game and making levels with LEGO) we asked the participants about their experience. Through recording of these vocal discussions, we identified comments ranging from their thoughts on the experiences to how best to develop the game going forward, and from new audio game ideas to their preferred controllers and consoles for accessibility purposes. This focused conversation around audio games, stimulated by our staggered interactive play sessions, proved to be extremely fruitful in generating feedback. Repeated ideas were taken note of and while many comments reaffirmed some of the assumptions we made during the design process, some revelations we hadn’t considered about designing audio games also emerged. We intend to implement these in future iterations of our audio game’s design, as discussed in Section 5.

**Figure 3:** A workshop participant building a LEGO level map

**Figure 4:** Level 2 heat-map using tank controls
The gameplay of players was recorded through a heat-map generation script within the game, represented in Figures 4, 5 and 6 (and at the end of the paper). This was made specifically for our game to capture the diverse ways in which players moved through the game space and was useful data for the researchers who otherwise couldn’t see the movement of players (Coulton et al. 2008). Each heat-map represents one attempt at a level, ending when the endpoint is reached, or an enemy is collided with (meaning the level is lost and reset). The maps highlight the difference in movement between the two controls schemes and successful movement strategies. These maps were visualised from a top-down viewpoint mimicking Pac-Man’s camera positioning, enabling rapid adaptation into physical depth maps through 3D printing for accessibility purposes (McDonald et al. 2014). The 2D perspective of our heatmaps was important to be able to convert them into 3D objects using displacement maps using colour as depth.

Places where the players spent little time are represented as dark blue, changing to cyan and then pink in parts of the heatmap where players crossed many times, or lingered longer in one place. On these maps the walls, player start-points, endpoints and enemies are also shown. These are displayed as white lines and yellow, green and red circles respectively. The maps visualise the confusion in understanding corners sonically. Giving walls varied audio dependant on their spatial orientation would alleviate this.

As well as the recordings of discussions being translated into transcripts from each of the 30-minute sections after each interactive experience, the discussions and comments
throughout the day were also converted into visual sketch notes (which could be converted to audio spatial notes) to give access to the outputs of the workshop in a variety of different sensorial media to increase accessibility options.

**IMPROVING AUDIO GAME DESIGN**

From the workshop, several takeaways emerged from repeat comments from the participants. All the participants expressed a strong positivity either towards the game itself or about further experiences within digital audio interactive spaces. They expressed enjoyment towards the difficulty and learnability a level-based maze game afforded them. When asked what they thought of it, the initial response was ‘It was really good’ from one participant, and when questioned further:

‘Because the further you go the more complex it gets with the introduction of the walls and the monsters. But I thought the monsters were quite easy to bypass. It was the walls that were a more difficult challenge.’

Some other participants seemed to share this feeling:

‘as you get through the levels it’s getting more and more difficult, so you’ve got something like a challenge to work through, rather than making like a narrative roleplaying kind of game which once you’ve done it you’ve done it, you can only really change the story, you can’t kind of get to the next level kind of thing. So that’s why I prefer this kind of game.’

The enjoyment and enthusiasm of the participants during the Co-Design workshop towards the experiences we presented to them was clear, and so furthering the development of our research within this space is paramount.

The data from the heat-maps, along with players reaching the later levels within the game proves that the sound alone was enough for them to perceive and understand the game environment. However, the methods of navigation visible in the heat-maps appeared bound to trial and error, meaning collision with walls (especially at corners) was extremely common and enemies were also hit frequently. This issue could potentially be rectified in this game and avoided in future audio games using a form of haptic mini-map (Yuan, Folmer, 2008). Rich audio not bound to stereo (dual channel) alone could also be employed which would need further user testing and runs the risk of falling into the problematic issue of making hardware required for these games’ play non-accessible.

It was also clear from the workshop that participants were excited about the potential of varied audio and accessible games, with them suggesting many different audio games including racing, alley shoot ‘em ups and mystery solving games:

‘I’d love to play a racing game. That’s one of the genres of game that I miss the most.’

‘if you have like a tin pan alley kind of thing, and you had different sounds. It may be a fun game if you did get it to work with a directional thing.’

‘For me it was like I thought it would fit maybe well with something like where you’re a jewel thief trying to break into something, or you’re trying to do something delicate and tricky’

The majority of participants appeared pleased by the move away from narrative focus within this audio game, but most still had some desire for narrative elements or framing in some capacity. The idea of creating an iconic audio game and how our game
compared stylistically to classic visual games like Tetris (Pajitnov, 1984) (which have no narrative elements at all and uses electronic 8-bit audio (Zappi, 2020)) was discussed: ‘I think what would be good as an abstract game, if you just went all out, like Tetris is an abstract game’. This point was very noteworthy, highlighting ideas around replay-ability of non-narrative games (Adellin et al. 2019) and how the different kinds of players they attract would translate when designing audio only games.

**INCREASING ACCESSIBILITY BEYOND GAMES**

Beyond the differentiation between different genres of games, the game’s movement and output mechanics were also seen as useful by the participants with some suggesting audio maps as a calmer experience for them to enjoy or use before going to a new place:

‘I was thinking I would love to take that element (audio navigation), make it quite easy, and put some really nice sounds in it, like create environments. I’m more into that.’

‘Say if there was like an audio map that you can access beforehand and figure out how to get from this place to that place. You could use it in so many different settings, even like in shopping centres and know which shops are where.’

These could be used before attending public spaces such as museums, airports, or train stations so that a blind or visually impaired person could practice navigating that space to alleviate stress on the day of attendance or travel (Sánchez et al. 2010). With anyone who has played the game, there has been a period of time needed to get accustomed to the audio navigation. This is something we as designers also noticed as each time we tried the game during development we found it easier to play and comprehend. This was also the case for informal testing with colleges as well as the workshop participants suggesting ‘Distinguishing between three sounds in a field took me a bit of getting used to.’ Due to this it seems important to standardise systems for audio navigation in digital spaces, such as Metaverse environments and games. This would reduce the learning required for each new virtual space entered, similar to the standardising of controls and controllers across visual games which has occurred naturally over recent decades.

The workshop proved that sound navigation systems are sufficient on their own, but also that as additional accessibility options alongside visual output in existing spatial web conferencing systems like Gather or Mozilla Hubs, they would only increase usability if properly implemented. This functionality could also be used in similar ways to subtitles being turned on by wider audiences who don’t have a reliance on them as an accessibility feature but have improved ease of access when watching programs or films with them (Davies, 2019). But it is also important to be aware that sound can be used in ways that doesn’t mimic normal (or human) perception of physical spaces when designing for it digitally (Gualeni, 2011) and this can actually increase the amount of information conveyed.

While there was clear interest in haptics as a secondary sense within our game or others, it didn’t appear as important to the workshop participants as we expected: ‘I do play games at home occasionally and I do like the kind of haptic feedback that you can get from game controllers’.

While this could be in part due to the framing of our workshop, it was commented that if haptics were implemented it should be as a sensory output which isn’t required for gameplay: ‘It should be optional. For that reason, I suppose you can’t have it as a main feature’. This is similar to how sound is often not required in visual games: ‘you can enhance it but it shouldn’t deprive the game if it’s not switched on.’ Haptics may be overlooked due to a lack of hardware that supports it so exploring them further in a very similar fashion to how we did with sound (in isolation) could help to clarify usability, something that has been explored in one aspect by Sightlence: a haptic-only adaptation for Pong (Atari, 1972) for Blind and Deaf users.
Figure 7: Sketch-notes from the workshop
ISSUES IN ACCESSING ‘ACCESSIBLE’ MEDIA

While our research into non-visual and audio games has given us many starting points to explore further avenues, the most pressing issue it has highlighted are the struggles visually impaired users often face even before gaining access to these game’s main menus. Once blind and visually impaired users have managed to launch their chosen games, all accessibility features which have been researched, designed and implemented are available to them. However, on the desktops, home-screens or dashboards of their computers, mobile devices and games consoles, the accessibility focused game developer hasn’t yet managed to take control. During our workshop there was justified advocacy from our participants for chances to break free from their mobile devices:

‘I think it would be nice to actually put the iPhone away and go to something else for a specific thing. Because I think as a blind person, you’re constantly using that device because of its accessibility, and therefore do you want to really implement something that’s meant to be a sort of escapism to a day-to-day device?’

While mobile devices usually are the most accessible devices for visually impaired people, this rationalised desire to have separation or escape from a singular device for all purposes, an issue most digital device users who rely on visuals may be completely oblivious to, is really significant to the design of accessible games.

Because games consoles are becoming more expensive, and these extra costs are generally going directly to improving the visual fidelity of the games we play, games consoles themselves aren’t considered worthwhile for visually impaired users if they aren’t benefiting from the visual components. If non-visual games were to release on the newest consoles, not only would the hardware be very underutilised, but the variety of other non-visual games available to play would be very low. While some games have released with a sound output focus on Windows/MacOS (The Vale for example), actually making an account to purchase and download these games can often be almost impossible as a blind or visually impaired person. One of our participants informed us about the incompatibility of Steam (the most common place to download games on computers) when using screen reader software (a tool used by visually impaired people to convert on screen content into sound) after the workshop inspired him to try out other non-visual games available at the time.

With this research project winning the Visionary award (Visionary, 2022) alongside the feedback from the workshop and many larger studios rapidly adopting accessibility options, there is clearly great desire and promise for the future of digital accessibility for non-visual audiences. However, regardless of the games and software solutions we and others develop, if the hardware itself isn’t tailored for rich sensory experiences other than through visual output (Hoogen et al. 2009) (however immersive this may be for visual users), the sensory bandwidth (Coulton, 2020) will always be limited. This is important to consider when striving to enable accessibility in mainstream games rather than as niches purely for accessibility purposes (Metatla et al. 2020). With this in mind, designing a console or device specifically for non-visual audiences could be the way to promote accessibility with the greatest impact. The limitations of current hardware and software in non-visual respects has been made abundantly clear during the design of this game even while using a mainstream game engine, as well as during the workshop itself. Controllers for early game consoles were less defined, leading to differentiation per game and enabling massive potential for accessibility (Montfort, Bogost, 2020). Nowadays, peripherals for accessibility are often hard to come by, both in terms of monetary cost and scale of production (Parisi, 2015) and are generalised to fit a wide range of accessibility needs. If a non-visual console was envisaged, it could save the users money by cutting out the need for graphics cards which are becoming
increasingly expensive. Non-visual consoles could provide lower resolution visuals while running the same games as mainstream hardware, trading visuals for higher fidelity audio and haptic outputs. Through focusing on haptics and sound output, the future of accessible games could skyrocket, and while games would benefit from this first, all digital spatial environments would be advantaged over time. While we will continue to develop our game through Co-Design, we are limited by software and hardware designed for visual purposes. By creating devices tailored for accessibility rather than hacking together accessibility features with inappropriate hardware, spatial virtual environments, both games and otherwise can extend their accessible reach much further. Through a hardware and software combined approach, accessibility becomes a core consideration at all stages of games and wider digital design. This approach is paramount to people who are only limited by our devotion to visual-first design.

ACKNOWLEDGMENTS
This template (used originally in DiGRA 2011 conference) was developed based on a similar template for the CHI conference (Doe and Smith 2011) and the template from DiGRA 2005. Some of the references cited in this paper are included for illustrative purposes only. Special thanks to Annika Waern and José Zagal.

BIBLIOGRAPHY


Frauenberger, C. and Noistering, M. 2003. “3D audio interfaces for the blind.” *Georgia Institute of Technology*. [Website](http://hdl.handle.net/1853/50428)

Galloway’s. 1867. [Website](https://www.galloways.org.uk/)

Gather. 2020. [Website](https://www.gather.town/)


HyHyve Meetup & Online Networking. 2022. [Website](https://www.hyhyve.com/%20alternative/gather-town/)


**HEATMAPS**

An array of heatmaps for each level 1-9 are displayed on the following pages to allow the reader to see more variety in the heatmaps. 4 heatmaps are shown for each level:
LEVEL 3
LEVEL 6