Work, Talk, or Play Through Auditory Virtual Space

Zach Mason, Paul Coulton, David Green, Joseph Lindley

Design Research Works, Imagination, Lancaster University
[z.mason] [p.coulton] [d.p.green] [j.lindley] @lancaster.ac.uk

Abstract: Video conferencing has exploded into a prominent space for our work and social lives due to the Covid-19 pandemic. This move to virtual co-presence has been joyless for many, highlighting the shortcomings of video conferencing. In contrast, the video gaming community has already freely adopted online shared spaces for several decades. Gather Town, a web-based video conferencing platform we explored attempts to adapt some conventions of video games, aiming to produce a better video conferencing experience, including interactive spatial design and use of avatars. By considering Gather's design affordances, we uncovered the need for alternative design approaches for both games and video conferencing, which predominantly rely on visuals. This led us to design and create non-visual games through Co-Design with a visual impairment support charity as a steppingstone to facilitate wider use of audio to enhance accessibility and spatial presence across digital spaces.

Keywords: virtual co-presence, video conferencing, auditory design, non-visual games.

Videoconferencing – or, commonly, being talked at by a wall of faces – is a joyless activity for many due to "[t]he impact of bad design… multiplied across the billions of people striving to flourish online" (O'Toole, 2020). This leads to a feeling known as "Zoom Fatigue" (Bailenson, 2020, p.1), especially when it constitutes a major part of your daily activities. We note that video calls often fail to achieve the kind of interaction they aspire to because the more users expect from remote communication, the more they note that "mediated" (Nowak, 2017) interactions will never be on par with face-to-face communication (Hollan, 1992, p.121).

It is notable that 'Gamers' (people who participate in video game culture, such as by regularly playing video games) not only *use* virtual interactions resembling the video conference calls which many find unpleasant, but usually *enjoy* them. As Luskin notes, "The idea of socialising in a game is not new at all" (2020). Can insights from a thriving, established culture, which relishes online mediated interactivity, inform the design of systems to support more enjoyable virtual interactions among the general public?

Games are by no means a perfect model for all virtual interactions, having been tailored specifically for Gamers—who amongst other stereotyped traits, "have shown statistically significant gains in problem solving, spatial skills and persistence" (Barr, 2017, p.87). This can lead games to be inaccessible to those who struggle in digital spaces, however platforms like Gather Town[1] take cues from the gaming world and repurpose them in simpler ways for more general audiences. Aiming to dissolve some of the shortcomings of video conferencing, Gather Town is a game-like spatial video conferencing system with its 'calls' being initiated by positioning their avatars in space. When two or more users' avatars come into proximity, a traditional call is established on-the-fly, allowing them to converse as they would via any other video-conference platform. The spaces in question are presented to users as two-dimensional rooms which are viewed from a 'top down'

[1] https://gather.town

perspective, and navigated through an avatar, using the W, A, S, and D keys on a standard keyboard as is common for mainstream games. Once the avatars move apart, the call cuts out allowing them to move elsewhere in the virtual environment.

In short, Gather Town's functionality is of a standard video and audio-conferencing system, overlayed on top of a simple game environment. After designing and experiencing a variety of spaces within it, we realised it suffers from many of the same pitfalls of mainstream video conferencing (e.g., Zoom, Teams, Google Meet, etc) such as video feeds being arranged in a grid and audio signals being 'flattened' so that concurrent streams cannot all be heard at the same time, acting to degrade the freedom of movement and sense of space the platform aims to promote. The aesthetic qualities, while initially refreshing soon fade into the experience. Eventually the flaws of videoconferencing systems (lack of eye contact, personal image issues and having to feel presentable in our home spaces) re-emerge. When playing games online, people rarely share their video feed, unless streaming publicly, in which case the aim is to present oneself. Consequently, although Gather has taken attributes from games and their communities, the rich and customisable avatar system has become overshadowed by people's real-world webcam feeds. These are overlayed in traditional video conference-rectangles, which are not only superficial irrelevant but also generate a conflict of focus.

Our embodiment and participation, as humans within spaces, is something we took for granted before the pandemic. "The flow and changes in interpersonal distances between individuals in a shared space is an integral part of nonverbal communication." (Williamson et al., 2021, p.4). Video feeds in Gather create a shroud which reduces the sense of space. This, alongside the inability to recognise which avatars are talking due to the lack of vocal spatialisation begins to make Gather, and programs like it feel quite flat or non-spatial. By conceptualising spatial sound as a tool for users to feel further immersed in virtual space, we began to envisage how video game designers might better account for the bandwidth of sound as a sense. This crucial revelation, where a disillusionment with video in conferencing became an extreme eagerness to investigate the capacity of audio was fundamental due to it being largely overlooked in virtual co-presence design.

While "the game industry has been quick to exploit increased graphics capabilities to create often stunning virtual, visual environments, they remain leery of exploiting the potential of audio technology" Grimshaw and Schott (2007). Game space mirrors physical space in its lack of accessibility for blind and visually impaired people who engage with it. Spatial (binaural) audio can easily be employed to enhance the fidelity of the audio aspect of a game space, using a standard game engine such as *Unreal Engine,* without any peripherals other than a set of stereo headphones to create generative spatial audio. Feeling the need for these mechanics to be explored, we embarked on the creation of a game to realise and highlight the need for spatial sound systems in software like Gather. We perceived that by enabling full functionality without the need for any visuals in our game, the same could be done in systems like Gather. Working with SASL[2] (Sight Advice South Lakes), a sight loss charity who have connections with a diverse group of blind and visually impaired people, we will gain insights from those most attuned with sound as a navigational sense. This will steer the design of the mechanics and control systems for the game through a Co-Designed collaborative process.

[2] http://sightadvice.org.uk/

The limits of tangibility within interlayered sound channels are more challenging to accommodate than with visual depictions of space because "the amount of information conveyable through sound is limited" Röber and Masuch (2005). For that reason, our game's elements in initial testing stages resemble classic games like *Pong* (Atari, 1972) or *Pac-Man* (Namco, 1980), with complexity kept to a minimum. *A Blind Legend* (Dowino, 2015) and *Papa Sangre* (Somethin' Else, 2010) are two prominent non-visual binaural audio games. While they focus on story, using the lack of visuals to create horror themed games, we intend to focus on the mechanical aspects, aiming to test user control schemes to discover which best induce a sense of space without any visuals.

Applying systems that we develop through our Co-Designed approach alongside visual elements like Gather's avatar and movement system will encourage a spatial regard towards digital platforms. Rather than assuming the need for video due to conventions from standard video conferencing which can make users fatigued and uncomfortable, we believe future digital spaces for virtual and hybrid co-presence need to have increased sensory regard and usage beyond sight to immerse users without the need for costly or inaccessible equipment to be accessible and enjoyable for all audiences.

References

Andrist, S., Mutlu, B., & Gleicher, M. (2013). *Conversational gaze aversion for virtual agents.* In R. Aylett, B. Krenn, C. Pelachaud, & H. Shimodaira (Eds.), Intelligent virtual agents, 249–262. https://doi.org/10.1007/978-3-642-40415-3_22

Argyle, M., & Dean, J. (1965). *Eye-contact, distance and affiliation.* Sociometry, 28(3), 289–304. https://doi.org/10.2307/2786027

Bailenson, J. N. (2021). *Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue.* Technology, Mind, and Behavior, 2(1). https://doi.org/10.1037/tmb0000030

Barr, M. (2017). *Video games can develop graduate skills in higher education students: A randomised trial,* Computers & Education, Volume 113, 86-97. ISSN 0360-1315. https://doi.org/10.1016/j.compedu.2017.05.016

Bernieri, F. J., Davis, J. M., Rosenthal, R., & Knee, C. R. (1994). *Interactional Synchrony and Rapport: Measuring Synchrony in Displays Devoid of Sound and Facial Affect.* Personality and Social Psychology Bulletin, 20(3), 303–311. https://doi.org/10.1177/0146167294203008

Bulu, S. T. (2012). *Place presence, social presence, co-presence, and satisfaction in virtual worlds,* Computers & Education, Volume 58, Issue 1, Pages 154-161, ISSN 0360-1315,

https://doi.org/10.1016/j.compedu.2011.08.024

Chen, J. (2007). *Flow in games (and everything else).* Commun. ACM 50, 4 (April 2007), 31–34. https://doi.org/10.1145/1232743.1232769

Coulton, P. (2017). Sensing atoms and bits. in I Heywood (ed.), Sensory arts and design. London: Bloomsbury.

Erickson, T., & Kellogg, W. A. (2000). *Social translucence: an approach to designing systems that support social processes.* ACM Trans. Comput.-Hum. Interact. 7, 1 (March 2000), 59–83. https://doi.org/10.1145/344949.345004

Fiadotau, M. (2018) *Indie Game.* In: Lee N. (eds) Encyclopedia of Computer Graphics and Games. Springer, Cham. https://doi.org/10.1007/978-3-319-08234-9_188-1

Frauenberger, C., & Noisternig, M. (2003). 3D *Audio Interfaces for the Blind*. In *Proceedings of the 2003 International Conference on Auditory Display*. Boston, MA, USA. http://www.icad.org/websiteV2.0/Conferences/ICAD2003/paper/68%20Frauenberger.pdf

Gregory, R. L. (1997). *Eye and brain: The psychology of seeing.* Originally published in 1966 and updated of Princeton university press.

Grimshaw, M., & Schott, G. (2007). *Situating Gaming as a Sonic Experience: The acoustic ecology of First-Person Shooters*. In *Proceedings of the 2007 DiGRA International Conference: Situation Play.* Tokyo: The University of Tokyo. http://www.digra.org/digital-library/publications/situating-gaming-as-a-sonic-experience-the-acoustic-ecology-of-first-person-shooters/

Hollan, J., & Stornetta, S. (1992). *Beyond Being There.* In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92). Association for Computing Machinery, New York, NY, USA, 119–125. https://doi.org/10.1145/142750.142769

Hoogen, W. M. v.d., Wijnand A. I., & Kort, Y. A.W. d. (2009). *Effects of Sensory Immersion on Behavioural Indicators of Player Experience: Movement Synchrony and Controller Pressure*. In *Proceedings of the 2009 DiGRA International Conference: Breaking New Ground: Innovation in Games, Play, Practice and Theory*. London: Brunel University. http://www.digra.org/digital-library/publications/effects-of-sensory-immersion-on-behavioural-indicators-of-player-experience-movement-synchrony-and-controller-pressure/

Jia, A. L., Shen, S., Bovenkamp, R. D. V., Losup, A., Kuipers, F., & Epema, D. H. J. (2015). *Socializing by Gaming.* ACM Transactions on Knowledge Discovery from Data. https://doi.org/10.1145/2736698

Johannes, N., Vuorre, M., & Przybylski, A. K. (2021). *Video Game Play is Positively Correlated With Well-Being.* R. Soc. Open Sci. 8: 202049. https://doi.org/10.1098/rsos.202049

Lufkin, B. (2020). *How online gaming has become a social lifeline.* BBC. https://www.bbc.com/worklife/article/20201215-how-online-gaming-has-become-a-social-lifeline

International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems Proceedings, 1-6. https://doi.org/10.1109/VECIMS.2011.6053842

Molins-Ruano, P., Sevilla, C., Santini, S., Haya, P. A., Rodríguez, P., & Sacha, G. M. (2014). *Designing videogames to improve students' motivation,* Computers in Human Behavior, Volume 31, Pages 571-579, ISSN 0747-5632. https://doi.org/10.1016/j.chb.2013.06.013

Morris, D., Joshi, N., & Salisbury, K. (2004). *Haptic Battle Pong: High-Degree-of-Freedom Haptics in a Multiplayer Gaming Environment*. In *Experimental Gameplay Workshop, Game Developers Conference.* San Francisco, CA, USA. http://dmorris.net/projects/haptic.battle.pong/hbp_description.pdf

Nordvall, M. (2014). The Sightlence Game: Designing a Haptic Computer Game Interface. In *Proceeding of the 2013 DiGRA International Conference: DeFragging Game Studies*. http://www.digra.org/digital-library/publications/the-sightlence-game-designing-a-haptic-computer-game-interface/

Nowak, K. L., Watt, J., & Walther, J. B. (2005). *The Influence of Synchrony and Sensory Modality on the Person Perception Process in Computer-Mediated Groups,* Journal of Computer-Mediated Communication, Volume 10, Issue 3, JCMC1038. https://doi.org/10.1111/j.1083-6101.2005.tb00251.x

Oppezzo, M., & Schwartz, D. L. (2014*). Give your ideas some legs: The positive effect of walking on creative thinking.* Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(4), 1142–1152. https://doi.org/10.1037/a0036577

O'Toole, R., & Warburton, N. (2020). *Zoom and Gloom.* Aeon. https://aeon.co/essays/how-empathy-and-creativity-can-re-humanise-videoconferencing

Pan, Y., & Steed, A. (2017). *The impact of self-avatars on trust and collaboration in shared virtual environments.* PLoS ONE 12(12): e0189078. https://doi.org/10.1371/journal.pone.0189078

Pasch, M., Bianchi-Berthouze, N., Dijk, B., & Nijholt, A. (2009*). Immersion in Movement-Based Interaction.* 9. 169-180. https://doi.org/10.1007/978-3-642-02315-6_16

Reeves, B., Lang, A., Kim, E. Y., & Tatar, D. (1999) *The Effects of Screen Size and Message Content on Attention and Arousal,* Media Psychology, 1:1, 49-67. https://doi.org/10.1207/s1532785xmep0101_4

Röber, N., & Masuch, M. (2005). *Playing Audio-Only Games A Compendium of Interacting with Virtual, Auditory Worlds*. In *Proceedings of the 2005 DiGRA International Conference: Changing Views: Worlds in Play*. http://www.digra.org/digital-library/publications/playing-audio-only-games-a-compendium-of-interacting-with-virtual-auditory-worlds/

Rosedale, P. (2020). *What If Zoom Had Spatial Audio*? HighFidelity. https://www.highfidelity.com/blog/what-if-zoom-had-spatial-audio

Sánchez, J., Sáenz, M., Pascual-Leone, A., & Merabet, L. (2010). *Enhancing Navigation Skills through Audio Gaming*. In *Chi '10 Extended Abstracts on Human Factors in Computing Systems*. https://doi.org/10.1145/1753846.1754091

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). *Video Games and the Future of Learning.* Phi Delta Kappan. 2005;87(2):105-111. https://doi.org/10.1177/003172170508700205

Shute, J. V., Ventura, M., & Ke, F. (2015). *The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills,* Computers & Education, Volume 80, Pages 58-67, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2014.08.013.

Targett, S., & Fernström, M. (2003). *Audio Games: Fun for All? All for Fun?* In *Proceedings of the 2003 International Conference on Auditory Display*, Boston, MA, USA. https://www.icad.org/Proceedings/2003/TargettFernstroem2003.pdf

Tomprou, M., Kim, Y. J., Chikersal, P., Woolley, A. W., & Dabbish, L. A. (2021). *Speaking out of turn: How video conferencing reduces vocal synchrony and collective intelligence.* PLoS ONE 16(3): e0247655. https://doi.org/10.1371/journal.pone.0247655

Wilde, T. (2016). *How WASD became the standard PC control scheme.* PCGamer. https://www.pcgamer.com/uk/how-wasd-became-the-standard-pc-control-scheme/

Williamson, J., Li, J., Vinayagamoorthy, V., Shamma, D. A., & Cesar, P. (2021). *Proxemics and Social Interactions in an Instrumented Virtual Reality Workshop.* Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 253, 1–13. https://doi.org/10.1145/3411764.3445729

Atari. 1972. *Pong*. Arcade, Various consoles. Atari/Namco.

Dowino. 2015. *A Blind Legend*. iOS, Android, macOS, Windows. Plus In Digital.

Namco. 1980. *Pac-Man*. Arcade, Various Consoles. Namco/Midway.

Somethin' Else. 2010. *Papa Sangre*. iOS. Somethin' Else.