An Interaction-Centered Perspective on Audio-Based Gameplay for Blind Players

Nicoleta Radu[1], Stine Froulund Astorp[2], Max Væhrens[3], Antonia Krummheuer[1], Thomas Bjørner[2], Flemming Christensen[3] & Dorte Hammershøi[3]

[1]Department of Culture and Communication, Aalborg University, Denmark | [2]Department of Architecture, Design and Media Technology, Aalborg University, Denmark | [3]Department of Electronic Systems, Aalborg University, Denmark

[1]ndr@ikk.aau.dk, antonia@ikk.aau.dk, [2]sfas@create.aau.dk, tbj@create.aau.dk, [3]maxv@es.aau.dk, fc@es.aau.dk, dh@es.aau.dk

DOI: 10.54337/aau.icdvrat26.42

Abstract

Audio is a key modality for blind players to navigate and interact with digital games. This paper presents a qualitative study of blind players engaging with commercial games to examine how audio supports the interaction process. We introduce an interaction-centered perspective structured around a four-stage loop: affordance, actionability, execution, and feedback. Our analysis demonstrates that interaction breakdowns occur when this loop is disrupted, including failures in perception, unclear actionability, problematic execution, and ambiguous feedback. We further identify two design tensions: accessibility vs. challenge and automation vs. player agency. Together, these findings provide an interaction-centered account of audio-based gameplay and highlight how actionable design supports functional presence in immersive virtual environments, forming a foundation for a broader framework of audio-based interaction.

1. Introduction

In non-visual gaming, audio serves as an important link between the player and the virtual world, facilitating everything from basic spatial navigation to complex system engagement. While current research often focuses on overcoming sensory barriers through feature-level accessibility [Yuan et al. (2011)], designing effective interaction remains a challenge in dynamic gameplay where players must continuously interpret cues and act in real time. In virtual reality (VR) research, the sense of presence (the subjective experience of being there) is tied to the player's ability to act meaningfully within the environment [Slater (2009)]. This suggests that accessibility is not only about information access, but about maintaining the continuity of the interaction process, from perceiving affordances to acting and interpreting system feedback.

This study adopts an interaction-centered perspective to examine how audio supports (or fails to support) blind players across diverse gameplay situations. Based on think-aloud and observational data from blind players engaging with commercial titles, we analyze how players move from perception to action and understanding. As a step toward a broader framework, we introduce a four-stage interaction loop structured around **affordance, actionability, execution, and feedback.** Our analysis shows that breakdowns frequently stem from disruptions within this loop, including gaps in actionability, missing or conflicting cues, and ambiguous feedback. We further identify two design tensions: between accessibility and challenge, and between automation and player agency. We argue that accessibility in immersive systems depends on supporting actionable understanding and functional presence, ensuring the player remains a capable agent within the virtual world.

2. Related work

Game accessibility research primarily addresses barriers in navigation and progression through feature-level adaptations [Yuan et al. (2011), Westin et al. (2020)]. While essential, this remedial approach often overlooks the preservation of core gameplay loops and the holistic player experience [Heron (2016), Power et al. (2012)]. Audio-specific research, including sonification and audio-only games, demonstrates that sound can effectively convey spatial and event-based information in non-visual contexts [Vickers (2011), Hermann et al. (2011)]. However, this work typically focuses on how information is encoded and perceived, rather than how it supports ongoing interaction processes.

Interaction theory offers a broader perspective by examining how users move between perception, action, and interpretation through the gulfs of execution and evaluation [Norman (2013)]. While this model captures the general structure of interaction, it does not account for how users derive actionable input mappings from non-visual cues, a challenge that becomes important in audio-based gameplay. Although immersion in virtual environments is tied to meaningful action [Slater (2009)], these processes are rarely examined in terms of how they break down in non-visual contexts. We draw on embodied interaction [Dourish (2001)] and the concept of *actionability*: the knowledge required to translate perceived affordances into meaningful action [Gaver (1991)]. Building on these perspectives, we model interaction as a structured process spanning multiple stages and examine how breakdowns emerge across this process. This interaction-centered account addresses a gap in current literature by explaining how non-visual interaction unfolds as a coherent system rather than a set of isolated features.

3. Method

3.1. Participants and Data Collection

The study is based on 4.5 hours of video recordings of six blind gamers (aged 29–58), five of whom were born blind and one who had been blind for the last 20 years as a consequence of an accident. The sessions followed a semi-structured format in which participants were invited to play games of their own choice to showcase their typical gaming experiences and the specific interaction challenges they encounter. Participants were encouraged to think aloud, sharing real-time reflections and reasoning as they interacted with the game environment. A sighted researcher was present, occasionally asking clarifying questions and providing minimal assistance when necessary. In Table 1, we present the specific games that the six gamers chose to demonstrate, categorized by their respective genres.

Genre	Games Represented
Action-Adventure	The Last of Us, Assassin's Creed, God of War
RPG / Adventure	Final Fantasy, Brok the Investigator
Specialized Audio Games	The Vale: Shadow of the Crown, A Hero's Call
Strategy / Simulation	Stardew Valley, RimWorld
Rhythm / Casual	Bits & Bops

Table 1: Commercial Games Selected and Played by Participants.

3.2. Data Preparation and Analysis

The recordings were segmented into discrete interaction episodes and documented as two types of logs: think-aloud (TA) logs capturing players' reflections, and observed (Obs) logs capturing interaction sequences and break-downs observed by the co-present sighted person. In total, the dataset comprised 163 think-aloud logs and 64 observed logs. Throughout this paper, participants are referred to as G1–G6, while specific data excerpts are indexed as TA1–TA163 for think-aloud reflections and Obs1–Obs64 for observed interaction sequences.

The analysis followed an iterative and inductive approach inspired by thematic analysis [Braun & Clarke (2021)]. Interaction events from the think-aloud and observed recordings were manually coded by one researcher in separate coding schemas and logged in Excel sheets. The coded material was grouped into clusters, abstracted into higher-level patterns, and finally synthesized into cross-cutting dimensions. These dimensions were later refined in a workshop with five researchers, where they were checked against the original data and adjusted where necessary.

4. Preliminary findings

Our analysis identifies a set of dimensions describing the structure of audio-based interaction for blind players, relating to a fundamental loop:

Affordance → Actionability → Execution → Feedback

Stage	Definition	Data-Backed Breakdown Example
Affordance	The player perceives via audio that an interaction is possible.	A required affordance is not perceived: a hole in the floor enabling progression is missed(G1, TA14).
Actionability	The player understands how to execute the intent using specific inputs.	Player presses buttons at random to find a combat action for a door (G4, Obs38).
Execution	The physical act of performing the command.	Player hears the "Press R1" assassination prompt, but cannot process and execute the input before the time window closes (G1, TA25).
Feedback	Non-visual confirmation that the action succeeded or failed.	After picking up an item, the player is unsure if the action was registered or if progress was saved (G1, TA24).

Table 2: The Interaction Loop: Evidence-Based Interaction Breakdowns.

Breakdowns emerge when one or more stages of this loop are disrupted, including failures in perceiving affordances, interpreting required actions, executing inputs, and understanding system feedback. As shown in Table 2, these disruptions manifest across all stages of the interaction process. In some cases, players are unable to perceive an affordance due to missing or misleading cues. For example, a player did not detect a hole in the floor required for progression, as the character automatically jumped over it without providing a non-visual cue (G1, TA14). In other cases, players perceive an affordance but lack guidance on how to act, resulting in an actionability gap. This often leads to trial-and-error behavior, such as pressing buttons at random and accidentally striking non-player characters (G4, Obs38). Execution can also break down when the required action is known but cannot be performed within the available timing window, such as when a player hears the "Press R1" assassination prompt but cannot execute the input before the opportunity disappears (G1, TA25). Finally, breakdowns also occur at the level of feedback, where players are unable to interpret whether an action has succeeded or how the system state has changed (G1, TA24). Conversely, when audio supports actionable understanding, players move through the environment with significantly higher independence.

Key Tensions in Audio-Based Interaction

The interaction loop is mediated by two central tensions: **Accessibility vs. Challenge** and **Automation vs. Agency**. Our data shows that players value *meaningful challenge* when it is grounded in interpretable information. In *The Last of Us*, the player described the difficulty of finding hidden trophies as highly rewarding precisely because the items are not explicitly marked (G2, TA54). If these challenges were removed via automation, the sense of achievement would disappear. The difficulty transitions into frustration only when the interaction loop is broken by missing information, turning a test of skill into an unfair external barrier.

A second tension exists between *automation and agency*. While features like GPS-guided navigation and auto-aim are essential for progression, they risk shifting the player from an active agent to a passive follower. Players frequently utilize automation to manage low-level tasks, such as a horse automatically following a path (G1, TA32). However, they also express a need to regain control for high-level decision-making, such as turning off navigation assistance to explore optional areas. Successful design must therefore ensure that accessibility features function as a support for the player's intent rather than a total substitution of their agency.

5. Discussion

While our study examines mainstream titles rather than VR-specific environments, the results have implications for non-visual immersion in VR. Presence is traditionally defined by Mel Slater through two components: Place Illusion (PI), the sensory feeling of being in a space, and Plausibility Illusion (Psi), the belief that events in that space are actually occurring [Slater (2009)]. Accessibility research has largely focused on PI by ensuring that blind players can perceive environmental elements [Yuan et al. (2011)]. However, our data suggests that for blind users, presence is a functional achievement rooted in interactional fidelity: the seamless completion of the interaction loop.

When this loop is disrupted, presence breaks down. Failures in perceiving affordances, forming actionable understanding, executing inputs, or interpreting feedback shift the player's focus away from the virtual environment and toward troubleshooting the interaction itself. As a result, players transition from situated actors within the virtual world to managing the interface, undermining the plausibility of the experience.

We argue that the design of inclusive immersive systems must extend beyond high-fidelity 3D audio. Psi is sustained not only by how a world sounds, but by how reliably it responds to a player's agency. Supporting a continuous and responsive interaction loop is therefore important for enabling blind players to maintain a consistent and meaningful sense of presence in virtual environments.

6. Conclusion

This paper presented an exploratory qualitative analysis of how audio supports interaction for six blind players in digital games. By examining think-aloud and observational data, we identify a four-stage interaction loop: affordance, actionability, execution, and feedback. Our findings show that breakdowns arise not only from missing sensory information, but also from failures in actionability, execution, and feedback.

The tensions between accessibility and challenge, and between automation and agency, further show that meaningful interaction depends on balancing system support with player independence. Players do not seek the removal of difficulty, but the actionable information needed to overcome it. We introduce interactional fidelity as a pre-requisite for maintaining plausibility in non-visual immersive systems, arguing that the virtual world remains meaningful when the interaction loop is continuous and responsive to the player's intent.

As ongoing work, these findings provide a foundation for a broader interactional framework for audio-based interaction. We will next analyze interviews with the six recorded participants and three additional blind participants who participated in interviews only. This analysis will be compared with the findings from the gameplay recordings to examine whether the current dimensions are reinforced or whether additional dimensions emerge. Finally, the resulting framework, with dimensions mapped onto the interaction loop, will be translated into a questionnaire-based instrument to support systematic evaluation and design of audio-based interaction in games and immersive systems.

References

Braun, V. & Clarke, V. (2021), *Thematic Analysis: A Practical Guide*, SAGE Publications, London, UK.

Dourish, P. (2001), *Where the Action Is: The Foundations of Embodied Interaction*, MIT Press.

Gaver, W. W. (1991), Technology affordances, in '*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*', CHI '91, Association for Computing Machinery, New York, NY, USA, p. 79–84. https://doi.org/10.1145/108844.108856

Hermann, T., Hunt, A. & Neuhoff, J. G. (2011), *The Sonification Handbook*, Logos Verlag Berlin.

Heron, M. J. (2016), 'Cultural integration and the accessibility of gaming', *The Computer Games Journal* 5(3), 91–94.

Norman, D. (2013), *The Design of Everyday Things: Revised and Expanded Edition*, Basic Books. https://books.google.dk/books?id=qBfRDQAAQBAJ

Power, C., Freire, A., Petrie, H. & Swallow, D. (2012), Guidelines are only half of the story: accessibility problems encountered by blind users on the web, in '*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*', CHI '12, Association for Computing Machinery, New York, NY, USA, p. 433–442. https://doi.org/10.1145/2207676.2207736

Slater, M. (2009), 'Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments', *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1535), 3549–3557.

Vickers, P. (2011), Sonification for process monitoring, in T. Hermann, A. Hunt & J. G. Neuhoff, eds, '*The Sonification Handbook*', Logos Publishing House, Berlin, Germany, chapter 18, pp. 455–491.

Westin, T., Hamilton, I. & Ellis, B. (2020), Game accessibility: Getting started, in R. Dillon, ed., '*The Digital Gaming Handbook*', CRC Press, pp. 37–52.

Yuan, B., Folmer, E. & Harris, F. C. (2011), 'Game accessibility: a survey', *Universal Access in the Information Society* 10(1), 81–100.