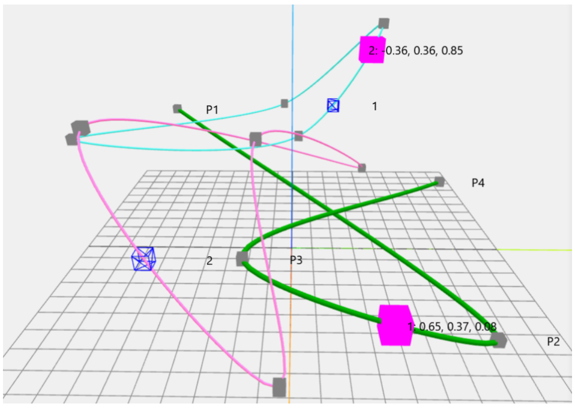
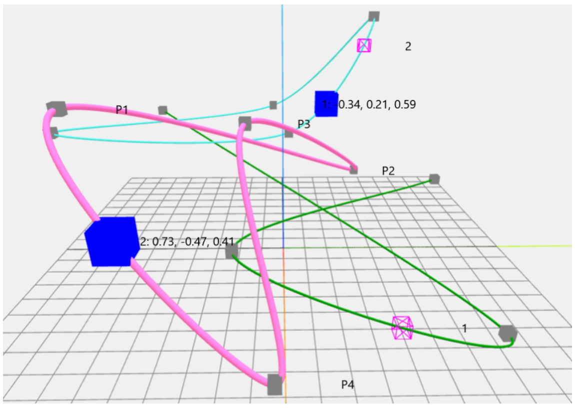
# Shared Agency through Collaborative Trajectory Editing in Immersive Audio


Lennart Sailer
lennartsailer@googlemail.com
Audio Communication Group, TU Berlin
Germany, Berlin

Henrik von Coler
hvc@gatech.edu
School of Music, Georgia Tech
USA, Atlanta


**(a) First user's view, focusing the green trajectory.**




**(b) Second user's view, focusing the pink trajectory.**

**Figure 1: Shared 3D-Scene for two connected users, each focusing a different trajectory.**


## Abstract

WeSp is a browser-based interface for real-time collaborative trajectory editing in immersive audio performance and production. Trajectories are defined as 3D splines and traversed by playheads that represent sound objects moving along these paths with adjustable speed and direction. Multiple users can edit the same set of trajectories simultaneously, enabling shared agency over a common virtual acoustic scene. A collaborative user study was conducted with four groups of four participants, using two tablets and two laptops within an immersive sound environment. Participants completed structured tasks and engaged in open exploration of the interface. Results from established usability questionnaires confirm usability and learnability. A thematic analysis of open-ended responses further reveals how groups coordinate interaction in live spatialization contexts. In particular, the shared and dynamic visual representation supported spatial orientation and made collaborative actions, such as synchronizing movements or following gestures, more apparent. Verbal communication emerged as an effective strategy for coordination among participants.



## Keywords

Collaborative Control, Spatial Audio, Web Technologies, User Experience Evaluation


## 1 Introduction

Trajectories are a common approach for organizing the spatial movement of virtual sound sources in immersive audio systems. They enable precise control over spatio-temporal attributes such as position, speed, and direction, while also providing a visual representation of both past and anticipated motion. This concept also aligns with the principle of parameter automation that is ubiquitous in today's electronic and computer-based music production. The use of trajectories in spatialization is specifically suited for object-based approaches which treat sound sources as discrete entities positioned and moved in space [10, 17]. Trajectory-based interfaces extend the automation paradigm by defining paths along which sound sources evolve over time. This approach has a wide range of applications that create virtual acoustic scenes, as in music, film scores, video games and interactive media.

However, trajectory-based composition – let alone performance – has certain usability constraints and limitations. Previous research shows that listeners have limited ability to distinguish and track complex spatial trajectories using auditory cues alone. Visual representations are thus necessary to improve localization and trajectory perception [14, 18]. In addition to this, trajectory-based workflows are not inherently expressive. While precise and perfectly suited for a studio environment, trajectory-editing is not a favorable performance practice. Most importantly for the scope of this paper, trajectory-based workflows are typically designed for single-user scenarios and do not support shared editing or collaborative control. This is a significant drawback, since the complexity of immersive audio environments is ideal for shared agency and creative interaction.

This work addresses the collaborative limitations by introducing WeSp, a browser-based interface for real-time collaborative trajectory editing in immersive audio. Designed as a research tool, WeSp separates shared trajectory structures from individually controlled sound objects, enabling collaborative editing while preserving individual agency in performance. A shared visual representation of the spatial scene, as shown in Figure 1, allows users to manipulate the shared trajectories, while each



*NIME '26, June 23–26, 2026, London, UK*

participant retains control over their own sound objects. Groups can thus collaboratively create immersive audio experiences in co-present and distributed settings.

A group user study presented in this paper aims at investigating the core principles and challenges of interaction in collaborative spatialization with WeSp. It follows the notion of *usability opportunities* – as proposed in prior work on immersive and AR-based interfaces [19, 20] – aiming to reveal promising directions for interaction design in collaborative spatialization.

# 2 Related Work

## 2.1 Object-Based Spatialization

Object-based spatialization represents sound sources as discrete entities that can be positioned and moved in space [10, 17]. In contrast to channel-based practices, which work with experimental algorithms on arbitrary loudspeaker systems, object-based approaches depend on spatial rendering algorithms such as Vector-Based Amplitude Panning (VBAP), Wave Field Synthesis (WFS), or Higher-Order Ambisonics (HOA). These systems are typically implemented as standalone applications or plugins for Digital Audio Workstations (DAWs), providing graphical interfaces for positioning and automating sound objects. While such interfaces support detailed control, interaction is often constrained to timeline-based automation or parameter editing, offering limited affordances for real-time manipulation and live performance [15].

## 2.2 Trajectory-Based Interfaces

Trajectory-based control extends object-based spatialization by defining paths along which sound sources move over time. While early approaches, like Chowning's work in the 70s [6], created *sound paths* algorithmically – without any visual representation – later systems such as Holo-Edit (1998)[1], IanniX (2003)[2] and Zirkonium (2006)[3] introduced graphical environments for designing and executing spatial trajectories.

IanniX, for example, formalizes trajectories through curves, triggers, and cursors, enabling both deterministic and generative behaviors, as well as integration with external media and control systems [7]. More recent tools such as Grapes[4] emphasize live interaction, offering editable trajectory presets and real-time control over speed, direction, and transitions between paths.

In addition to predefined trajectories, some systems incorporate dynamic behaviors such as random motion, flocking, or gesture-based manipulation. Despite these developments, trajectory-based interfaces remain primarily single-user systems, with limited support for collaborative editing or shared control of spatial structures.

## 2.3 Collaborative Spatial Audio Applications

Collaboration has been extensively researched in the NIME- and HCI-Community through tabletop interfaces (see for example [12] or [22]). Xambó et al. integrated Ambisonic spatialization in a tabletop system called SoundXY4 (2014) [23], where the rotation and position of white cubes represent virtual sound sources of the spatial scene. Tabletop interfaces offer a shared interaction design where the users are co-located in the same space performing on one interface together.

Virtual Environments offer new possibilities of collaboration without the need to be co-located. Çakmak et al. [24] developed a networked multimedia instrument based upon the idea of optical discs. Remote users have control of synthesis parameters during the performance. To compensate for the loss of bodily gestures between the performers in musical collaboration, they implemented a chat system. Virtual Reality (VR) and Augmented Reality (AR) are other interesting domains of virtual environments that have been studied in the scope of collaborative music performances.

Invoke (2023) [8] by Thomas Deacon and Mathieu Barthet is a VR-tool that allows users to sketch trajectories based on voice. Multiple users are able to spatialize sound samples in real-time using a combination of a virtual pen and the volume of the voice, which corresponds to the amplitude of a sample once it is attached to a trajectory.

A web-based example that deals with spatialization is INVISO (2017) [25] which enables users to construct sonic environments. It integrates binaural rendering of distances between a navigable head model and objects, which can be assigned to drawable trajectories. INVISO also comes in versions for VR and AR [26].

As previous studies on collaboration in virtual space point out identifying oneself and monitoring each other's actions facilitate a sense of a collaborative shared space [16].

## 2.4 Evaluation of Live-Interaction and Collaboration

Although the concept of collaboration is mentioned in several documentations of tools developed in the scope of novel musical interfaces, it often remains a perspective for further investigation in the development [5]. Nick Bryan-Kinns points out the importance of a shared representation to enhance mutually engaging interactions in the context of collaborative music making. Referring to previous research on creativity and group flow theory in the field of Human-Computer Interaction (HCI) he mentions the following aspects of collaboration: *mutual awareness of actions, shared and consistent representations, mutual modifiability, and shared annotation* [2, 3].

Related studies emphasize qualitative methods to explore the experiential dimension of interaction. Stowell et al. promote Discourse Analysis (DA) as a method of evaluating qualitative data such as transcribed interviews and comments to "extract a detailed reconstruction of users' conceptualization of a system" [21]. DA is a comprehensive but nevertheless demanding method of analysis that does not always appear practical when determining experience constraints. Grounded Theory and Thematic Analysis have emerged as standard approaches for investigating user experience and collaboration with musical interfaces [4, 8, 11].

# 3 WeSp

## 3.1 System and Interface Overview

WeSp is a browser-based application that lets multiple users interact with a shared spatial audio scene. The software is implemented in JavaScript, using HTML and CSS for structure and styling. It follows a client–server architecture in which multiple users connect to a central server via a web interface, enabling shared interaction within a common session.

Communication between clients and server is handled through bidirectional WebSocket connections, allowing low-latency exchange of interaction data. User actions within the interface, such

[1]https://gmem.org/holophon
[2]https://www.iannix.org/en/
[3]https://zkm.de/en/about-the-zkm/organization/hertz-lab/software/zirkonium
[4]https://grapes-3d.com/

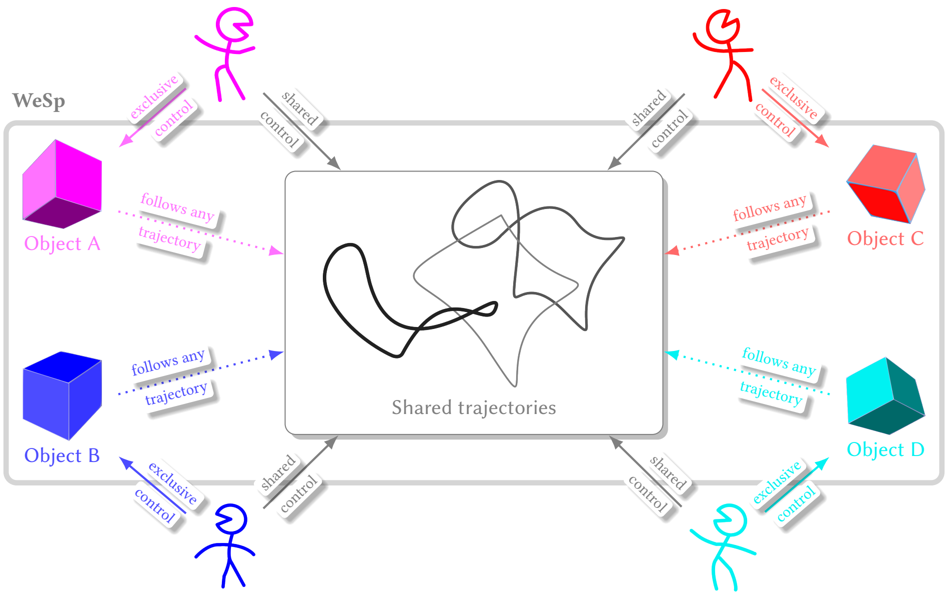



**Figure 2: Interaction concept of WeSp: all shared trajectories can be manipulated by any user (gray arrows); each user controls only their individual play-head (solid colored arrows); each play-head can follow any trajectory (dashed arrows).**

as editing trajectories or controlling sound objects, are transmitted to the server and redistributed to all connected clients to maintain a consistent shared state.

The system distinguishes between two types of data with different update characteristics. Shared data elements are synchronized on change, minimizing network load. In contrast, user-specific data are updated continuously to support real-time interaction and animation.

On the output side, interaction data is translated into OSC messages on the server and forwarded to external audio rendering environments. This enables integration with a wide range of spatial audio systems while keeping the interface independent from a specific rendering implementation. A Reaper[5] session has served as a rendering and playback system in the experiments presented in this paper. Further implementation details are available in the accompanying repository.[6]

## 3.2 Interaction and Collaboration Design

WeSp combines shared structural elements with individual control to support collaborative spatial performance. The system is based on two primary interaction primitives: *trajectories* and *sound objects*.

The interface is centered around a 3D scene that visualizes trajectories and sound objects, as shown in Figure 1. Trajectories are defined as editable spline curves and serve as spatial paths. They can be modified through direct manipulation in the 3D view as well as through parameter controls such as sliders and buttons, designed to support both precise adjustment and performative interaction on touch-based devices.

Sound objects act as playheads moving along these trajectories. Their behavior can be controlled in real time, either by adjusting playback speed along a trajectory or by manually positioning them along arbitrary trajectories. This establishes a separation between the spatial structure, defined by the trajectories, and the temporal behavior, defined by movement of the objects.

Collaboration in WeSp builds directly on this distinction between shared trajectories and individually controlled objects, as visualized in Figure 2. All users operate within a common 3D scene in which trajectories are mutually modifiable. Changes made by one user – indicated as solid gray arrows – are synchronized in real time, resulting in a shared and consistent spatial layout. Trajectories therefore function as a collectively authored layer that supports and demands coordinated planning and shaping of the performative space.

In contrast, sound objects are not shared. Each user controls only their own objects, preventing conflicts and discontinuities during performance. This establishes individual agency within a shared structure. To support coordination, objects are visually differentiated by user-specific colors. A user's own objects are emphasized through size and additional real-time information, while objects of others are visually reduced through transparency. This enables performers to identify their own actions while maintaining awareness of the overall scene.

This design supports shared agency and mutual awareness without requiring direct control over others' actions. Performers can adapt their behavior in response to the spatial positioning and movement of other objects, enabling coordinated group dynamics while preserving individual control.

# 4 Collaborative Interaction Study

## 4.1 Setup and Procedure

A group-based user study was conducted to examine collaborative interaction and user experience in WeSp. The study followed a mixed-methods approach, with particular emphasis on qualitative data from open-ended responses. It was guided by the following research question:

**How does the interaction design of WeSp support shared agency?**

[5] https://www.reaper.fm/
[6] https://github.com/koyi8/WeSp

**Figure 3: Participants in the studio during the study.**

The study took place in a controlled multichannel studio equipped with a 21-loudspeaker Ambisonic dome. Four groups of four participants (N = 16) took part in the study. Participants included students, musicians, and engineers with an interest or experience in spatial audio.

During the study, groups were seated around a table inside the loudspeaker system in the studio, as shown in Figure 3. Each group used four devices – two laptops and two tablets – all running the browser-based interface. Sessions lasted approximately one hour. The procedure consisted of four stages:

- Introduction to the interface (~6 min)
- Individual exploration (~4 min)
- Collaborative tasks and improvisation (~30 min)
- Post-study survey (~20 min)

During the collaborative phase, participants controlled a shared spatial soundscape consisting of four looped samples. Each participant controlled one sound source. Tasks progressed from individual familiarization to coordinated group tasks, followed by a five-minute improvisation.

## 4.2 Group Tasks

The collaborative phase consisted of a sequence of structured tasks with increasing complexity, followed by an open improvisation. The tasks can be organized into four phases:

**Phase 1: Initialization and Static Spatialization**

- *Object initialization:* Participants sequentially created one sound object each on a default trajectory without modifying interface parameters, first listening individually and then collectively.
- *Static distribution:* With animation disabled, objects were spread along the trajectory to explore spatial separation and global layout.
- *Group-based positioning:* Participants formed subgroups (tablet vs. laptop) and moved objects to designated spatial regions, requiring coordinated action.

**Phase 2: Dynamic Coordination**

- *Convergence:* With animation enabled, participants synchronized movement by converging to a shared point and acting as a single sound source.
- *Distributed motion:* Objects were redistributed and kept at approximately equal distances while participants reacted to each other through changes in speed and direction.

**Phase 3: Trajectory Editing**

- *Trajectory creation:* New trajectories were created and extended with additional control points.
- *Trajectory modification:* Trajectories were modified (e.g., shape, closure, tension), and objects reassigned.

**Phase 4: Creative Application**

- *Improvisation:* Groups developed and performed a short collaborative piece using all available features.

## 4.3 Survey Instruments and Analysis

After the session, participants completed a browser-based survey in LimeSurvey[7], consisting of open-ended questions and standardized questionnaires. Quantitative measures included the original User Experience Questionnaire (UEQ) [13] and a modified Sonic Interaction Design Questionnaire (SIDQ) [9]. The following modified SIDQ items were used:

(1) I was connected to the sound changes through my actions.
(2) I could relate my actions to specific sonic effects.
(3) I could plan spatial events using the interface.
(4) I felt part of a creative process.
(5) I felt in control.
(6) It was challenging.
(7) The feedback in the system made sense.
(8) I liked the spatialization I created.

Open-ended responses to structured survey questions were analyzed using Thematic Analysis (TA) [1], focusing on patterns related to collaboration, interaction strategies, and user experience. The questions are based on a related user study on VR interaction [9] and refined with AI assistance to address WeSp as a collaborative spatialization system. The open-ended questions were:

(1) How would you describe your experience collaborating with others within the WeSp interface?
(2) To explore ideas for spatial composition, how did you create shared agency?
(3) Did you feel engaged or disengaged in the group tasks? Explain why.
(4) Did you encounter any unexpected benefits or challenges while collaborating with others in WeSp? If so, what were they?
(5) Can you share any strategies or tactics you found effective for fostering collaboration within WeSp?
(6) Which features within WeSp did you find most helpful?
(7) Did you encounter any challenges while collaborating with others on the tasks?
(8) In what ways do you feel that WeSp's collaborative features differentiate it from other spatial audio tools?
(9) To what extent do you imagine using WeSp in a spatial live performance?
(10) What would you like to see improved in the interface, input methods, and/or interaction possibilities in WeSp?

[7] https://www.limesurvey.org/

# 5 Results

## 5.1 Participants

A total of 16 participants took part in the study. Participants ranged in age from 25 to 68 years ($M = 33.3$, $SD = 10.0$). Most participants reported a background in music (75%) or acoustics (62.5%).

## 5.2 User Experience

UEQ results indicate a positive user experience across all scales (Figure 4).

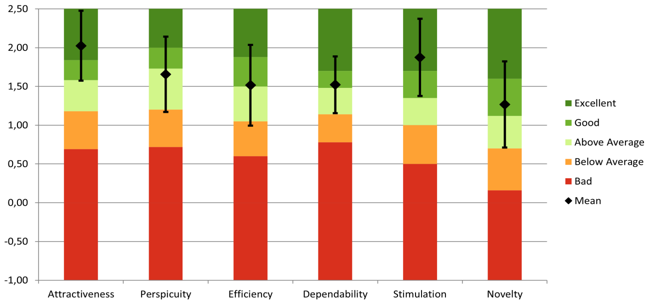



**Figure 4: UEQ results: Scale means with confidence intervals in relation to the benchmark dataset.**

All dimensions were rated above average relative to the UEQ benchmark dataset. Attractiveness and Stimulation reached the "excellent" range. Mean values above 0.8 indicate a clearly positive evaluation across all scales. The Novelty scale showed comparatively higher variability across participants.

## 5.3 SIDQ

SIDQ responses for tablet and laptop users are shown in Figure 5. Across all participants, responses were generally high for items related to perceived connection between action and sound and the ability to relate actions to sonic effects. Differences between device groups (laptop, tablet) were analyzed using Mann–Whitney U tests, with results summarized in Table 1.

**Table 1: Summary of Mann–Whitney U test results for device comparison.**

| Item | U | p | r |
|---|---|---|---|
| Connected | 28.00 | 0.704 | 0.1250 |
| Sonic Effects | 29.00 | 0.783 | 0.0938 |
| Plan Events | 8.50 | 0.011 | 0.7344 |
| Creative Process | 22.50 | 0.310 | 0.2969 |
| Control | 24.00 | 0.405 | 0.2500 |
| Challenging | 28.50 | 0.744 | 0.1094 |
| Feedback | 25.00 | 0.729 | 0.1071 |
| Liked Spatialization | 25.00 | 0.462 | 0.2188 |

Responses were largely consistent across device types. A statistically significant difference was observed only for the item *Plan Events* ($p = 0.011$), with higher ratings for laptop users.

## 5.4 Collaborative Interaction Themes

The thematic analysis revealed several recurring patterns in how participants interacted and coordinated within the shared environment.

1: **Shared Awareness Through Visual Representation.** Participants frequently relied on the visual representation of trajectories and sound objects to maintain awareness of others. The shared visual space enabled users to observe and adapt to each other's actions, supporting coordination without explicit role assignment.
However, this awareness was less consistent in the auditory domain. Some participants reported difficulty distinguishing their own sound source from others.

2: **Coordination Through Speech, Observation, and Restraint.** Collaboration was primarily achieved through verbal communication and visual observation. Talking was frequently described as an effective strategy for establishing shared intent.
In addition, several participants reported deliberately reducing their own activity to observe others. This form of restraint functioned as a coordination strategy.

3: **Engagement, Overload, and Loss of Individual Presence.** Most participants reported high levels of engagement, particularly during structured tasks. In more open-ended situations, some participants experienced reduced engagement due to perceptual masking and increased system activity.
This suggests that shared control can lead to attentional challenges in dense interaction scenarios.

4: **Trajectory-Based Interaction and Requested Extensions.** Participants responded positively to the trajectory-based interaction model, particularly the ability to manipulate and share spatial paths in real time.
At the same time, users requested extended functionality, including additional trajectory operations and alternative interaction modes beyond trajectory-based control.

# 6 Discussion

## 6.1 Usability Considerations

The quantitative results indicate that WeSp provides a usable interaction framework across devices. UEQ scores show consistently positive evaluations across all dimensions, with particularly high ratings for Stimulation and Attractiveness. In relation to the UEQ benchmark dataset, which aggregates results from a broad range of interactive products, all scales are rated above average, with several dimensions reaching the *excellent* range. This situates the user experience of WeSp within or above the range of typical interactive systems, despite its collaborative and spatial interaction context.

SIDQ responses indicate that participants were generally able to relate their actions to resulting sonic changes and reported a sense of control over the spatialization process. Differences between tablet and laptop users were limited, with a statistically significant effect observed only for the item *Plan Events*, where laptop users reported higher ratings. This suggests that tasks requiring precise spatial planning may be more effectively supported by pointer-based interaction than by touch input. In addition, the larger display size of laptops may support interaction with the visual representation, which participants relied on for coordination.

At the same time, qualitative responses point to limitations in the current interaction design. Participants reported difficulties in distinguishing their own sound sources, particularly in

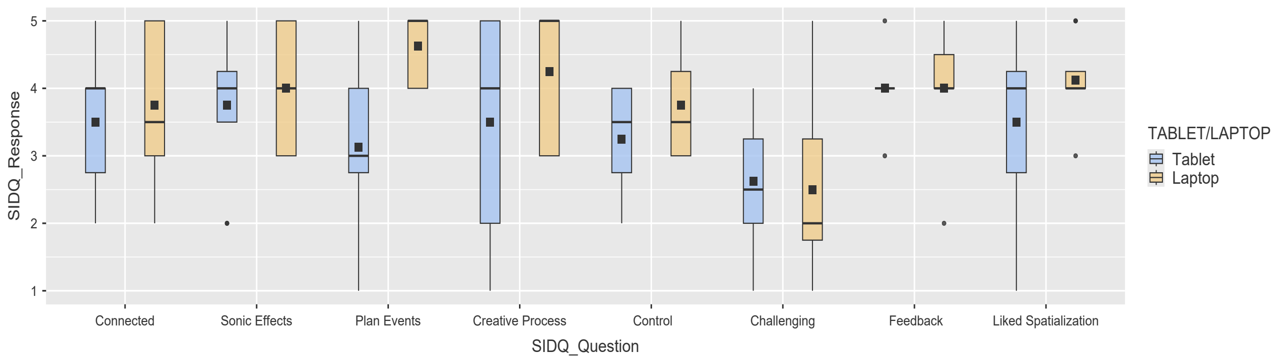



**Figure 5: SIDQ results for tablet and laptop users. Squares indicate means, lines indicate medians; outliers are marked as dots.**

dense interaction scenarios. This indicates a potential imbalance between visual interaction and auditory feedback, and suggests increased cognitive load when multiple users act simultaneously. In this context, participants appeared to prioritize visual information for orientation and coordination, which may further reduce the perceptual salience of individual sound contributions. This can affect the ability to track one's own actions within the shared sonic result, particularly as the number of active sound sources increases.

These observations point to a tension between visual coordination and auditory differentiation, where support for group awareness may come at the expense of individual perceptual clarity.

## 6.2 Collaborative Spatial Interaction and Shared Agency

The qualitative findings provide insight into how participants coordinated within the shared spatial environment. A central aspect was the reliance on the shared visual representation of trajectories and sound objects. This visual layer allowed participants to observe each other's actions and adjust their own behavior accordingly, often without explicit role assignment.

In addition to visual awareness, participants described verbal communication as a primary means of coordination. Speech was used to establish shared intent, negotiate actions, and organize group behavior. At the same time, non-verbal strategies emerged, including synchronization of movements, imitation of others, and temporary withdrawal from active control in order to observe the evolving scene.

These behaviors resemble established practices in collaborative music-making, such as synchronization and role differentiation. In this context, interaction was not structured through predefined roles, but emerged through continuous adjustment and negotiation among participants.

Together, these observations describe a form of interaction in which control is distributed and dynamically reconfigured across the group. This interaction can be understood as the emergence of *shared agency*, in which participants continuously adapt their actions in relation to others, the shared visual representation, and the evolving spatial sound scene.

Shared control over trajectories, combined with the visibility of actions, enabled participants to influence the overall sonic outcome collectively. Rather than assigning fixed responsibilities, interaction unfolded as a distributed process in which agency was negotiated in real time. In this sense, agency is not located in a single user, but emerges from the interaction between participants and the system.

## 6.3 Usability Opportunities

The findings point to several design implications for collaborative immersive audio systems. First, the shared visual representation plays a central role in coordinating group interaction. This highlights the importance of tightly integrated multimodal interaction, in which visual and auditory elements are aligned to support both coordination and perceptual clarity.

Second, the frequent use of verbal communication highlights the importance of explicit coordination mechanisms in collaborative spatial interaction. In co-present studio settings, verbal interaction functions as an effective and flexible modality for establishing shared intent and coordinating group behavior. This suggests that collaborative spatial systems can benefit from explicitly supporting such multimodal interaction, in which speech, gesture, and visual feedback are combined. At the same time, in remote or performance contexts where verbal communication may be limited, this points to the need for alternative coordination mechanisms embedded within the interface.

Third, the observed variability in tasks related to spatial planning reflects the complexity of balancing precision and accessibility in interaction design. This points to opportunities for developing interaction techniques that support both detailed control and intuitive use across different input modalities.

Overall, these observations indicate that interaction characteristics in collaborative spatial systems can serve as indicators of design potential, pointing toward new directions for supporting distributed and shared forms of interaction.

# 7 Conclusion and Future Work

This paper presented WeSp, a browser-based interface for collaborative trajectory-based spatialization. A group user study with music experts examined its potential for shared interaction, combining quantitative measures with qualitative analysis.

The results indicate that WeSp supports collaborative engagement through a low barrier to entry, shared visual representations, and the emergence of interaction strategies grounded in musical practice. At the same time, challenges related to multimodal cognitive load and coordination highlight areas for further investigation. The audio-visual tension is a well-known problem in electronic music contexts, where visualization may distract from relevant audio content.

The findings further highlight the role of communication and shared representations in shaping collaborative spatial interaction, extending beyond purely interface-based considerations.

Future work will address the balance between visual and auditory feedback, with particular attention to supporting clearer auditory perception in complex group scenarios. Another direction concerns the development of alternative coordination mechanisms that reduce reliance on verbal communication, especially in performance contexts.

While the present study focuses on co-present interaction, where verbal communication plays a central role in coordination, future work may explore alternative forms of *talking* within collaborative spatial systems. This includes interface-based communication mechanisms that extend or replace speech in remote or performance contexts, enabling coordination through gesture, visual signaling, or shared interaction states.

## 8 Ethical Standards

The user study was conducted at Technical University Berlin as part of an ongoing research project on interactive spatial audio interfaces. Participation was voluntary, and all participants provided informed consent. All participants were adults. Questionnaire responses and demographic data were collected anonymously and were not linkable to individuals. No personally identifiable information was recorded. The study involved minimal risk and complied with applicable data protection regulations.

## References

[1] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa Place: United Kingdom Publisher: Hodder Arnold.

[2] Nick Bryan-Kinns. 2013. Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies* 71, 1 (2013), 76–90. https://doi.org/10.1016/j.ijhcs.2012.02.004 Place: Netherlands Publisher: Elsevier Science.

[3] Nick Bryan-Kinns and Fraser Hamilton. 2012. Identifying mutual engagement. *Behaviour & Information Technology* 31, 2 (Feb. 2012), 101–125. https://doi.org/10.1080/01449290903377103 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01449290903377103.

[4] Nick Bryan-Kinns, Wei Wang, and Yongmeng Wu. 2018. Thematic Analysis for Sonic Interaction Design. In *From Data to Design Workshop at British Computer Society Human-Computer Interaction Conference 2018.* https://doi.org/10.14236/ewic/HCI2018.214

[5] Ivica Ico Bukvic, Disha Sardana, and Woohun Joo. 2020. New Interfaces for Spatial Musical Expression. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME-20).* Zenodo, 249–254. https://doi.org/10.5281/zenodo.4813342

[6] John Chowning. 1977. The Simulation of Moving Sound Sources. *Computer Music Journal* 1 (June 1977). https://doi.org/10.2307/3679609

[7] Thierry Coduys and G. Ferry. 2004. Iannix. Aesthetical/symbolic visualisations for hypermedia composition. In *Proceedings of the Sound and Music Computing Conference.* Paris,France. https://doi.org/10.5281/zenodo.849179

[8] Thomas Deacon and Mathieu Barthet. 2023. Invoke: A Collaborative Virtual Reality Tool for Spatial Audio Production Using Voice-Based Trajectory Sketching. In *Proceedings of the 18th International Audio Mostly Conference (AM '23).* Association for Computing Machinery, New York, NY, USA, 161–168. https://doi.org/10.1145/3616195.3616211

[9] Thomas Deacon, Tony Stockman, and Mathieu Barthet. 2017. User Experience in an Interactive Music Virtual Reality System: An Exploratory Study. In *Bridging People and Sound,* Mitsuko Aramaki, Richard Kronland-Martinet, and Sølvi Ystad (Eds.). Springer International Publishing, Cham, 192–216. https://doi.org/10.1007/978-3-319-67738-5_12

[10] Jeremie Garcia, Jean Bresson, and Thibaut Carpentier. 2015. Towards interactive authoring tools for composing spatialization. In *2015 IEEE Symposium on 3D User Interfaces (3DUI).* IEEE, Arles, France, 151–152. https://doi.org/10.1109/3DUI.2015.7131745

[11] Andrew Johnston. 2011. Beyond Evaluation : Linking Practice And Theory In New Musical Interface Design. In *Proceedings of the 2011 International Conference on New Interfaces for Musical Expression (NIME11).* https://doi.org/10.5281/ZENODO.1178053

[12] Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. 2007. The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international conference on Tangible and embedded interaction (TEI '07).* Association for Computing Machinery, New York, NY, USA, 139–146. https://doi.org/10.1145/1226969.1226998

[13] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. *USAB 2008* 5298, 63–76. https://doi.org/10.1007/978-3-540-89350-9_6

[14] G. Marentakis and S. Mcadams. 2013. Perceptual impact of gesture control of spatialization. *ACM Trans. Appl. Percept.* 10, 4, Article 22 (oct 2013), 21 pages. https://doi.org/10.1145/2536764.2536769

[15] Justin Mathew, Stéphane Huot, and Brian Katz. 2017. Survey and implications for the design of new 3D audio production and authoring tools. *Journal on Multimodal User Interfaces* 11 (April 2017). https://doi.org/10.1007/s12193-017-0245-z

[16] Liang Men and Nick Bryan-Kinns. 2019. LeMo: Exploring Virtual Space for Collaborative Creativity. In *Proceedings of the 2019 Conference on Creativity and Cognition (C&C '19).* Association for Computing Machinery, New York, NY, USA, 71–82. https://doi.org/10.1145/3325480.3325495

[17] Nils Peters, Georgios Marentakis, and Stephen Mcadams. 2011. Current Technologies and Compositional Practices for Spatialization: A Qualitative and Quantitative Analysis. *Computer Music Journal* 35 (March 2011), 10–27. https://doi.org/10.1162/COMJ_a_00037

[18] Federico Schumacher, Vicente Espinoza, Francisca Mardones, Rodrigo Vergara, Alberto Aránguiz, and Valentina Aguilera. 2021. Perceptual Recognition of Sound Trajectories in Space. *Computer Music Journal* 45, 1 (March 2021), 39–54. https://doi.org/10.1162/comj_a_00593

[19] Hyunkyung Shin and Henrik von Coler. 2024. Arcube: Hybrid spatial interaction for immersive audio. In *Proceedings of the 2024 ACM Symposium on Spatial User Interaction.* 1–3.

[20] Hyunkyung Shin and Henrik von Coler. 2025. AR Matchmaking: The Compatibility of Musical Instruments with an AR Interface. *space* 20 (2025), 29.

[21] D. Stowell, A. Robertson, N. Bryan-Kinns, and M. D. Plumbley. 2009. Evaluation of live human–computer music-making: Quantitative and qualitative approaches. *International Journal of Human-Computer Studies* 67, 11 (Nov. 2009), 960–975. https://doi.org/10.1016/j.ijhcs.2009.05.007

[22] Anna Xambó. 2015. *Tabletop tangible interfaces for music performance : design and evaluation.* Ph. D. Dissertation. The Open University, Milton Keynes.

[23] Anna Xambó, Gerard Roma, Robin Laney, Chris Dobbyn, and Sergi Jordà. 2014. SoundXY4: Supporting Tabletop Collaboration and Awareness with Ambisonics Spatialisation. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME-14).* https://doi.org/10.5281/zenodo.1178985

[24] Cem Çakmak, Anil Camci, and Angus Forbes. 2016. Networked Virtual Environments As Collaborative Music Spaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME-16).* 106–111. https://doi.org/10.5281/zenodo.1176002

[25] Anil Çamcı, Kristine Lee, Cody J. Roberts, and Angus G. Forbes. 2017. INVISO: A Cross-platform User Interface for Creating Virtual Sonic Environments. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17).* Association for Computing Machinery, New York, NY, USA, 507–518. https://doi.org/10.1145/3126594.3126644

[26] Anıl Çamcı, Aaron Willette, Nachiketa Gargi, Eugene Kim, Julia Xu, and Tanya Lai. 2020. Cross-platform and Cross-reality Design of Immersive Sonic Environments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Romain Michon and Franziska Schroeder (Eds.). Birmingham City University, Birmingham, UK, 127–130. https://doi.org/10.5281/zenodo.4813270