# Immersive Echoes: Impact of Spatial Audio on Virtual Social Engagement and Realism

Jamie Kim
Georgia Institute of Technology
jkim3678@gatech.edu

## ABSTRACT

This study evaluated the impact of spatial audio cues on enhancing user engagement and realism in virtual social gatherings. Conducted with three college students aged 20-24 from Georgia Tech, the research focused on the potential of spatial audio to bridge the gap between virtual and physical experiences. Due to logistical convenience, the study pivoted away from a larger, more diverse participant group than proposed prior. Key findings from this study suggest that spatial audio significantly improves users' understanding of virtual environments, making interactions feel more dynamic and lifelike. Participants reported enhanced clarity in discerning the direction and proximity of sounds, which contributed to a more immersive and realistic experience. This underscores the effectiveness of audio cues in enriching virtual social interactions.

## Keywords

Spatial Audio, Virtual Reality, User Engagement, Audio Cues, Social Virtual Gatherings, User Experience (UX), Immersive Technology, Auditory Spatialization.

## 1. INTRODUCTION

Since the Covid-19 pandemic, the transition from physical to virtual environments for socialization, education, and professional interactions has accelerated. This shift has uncovered many shortcomings in virtual platforms, particularly in the aspects of engagement, emotional satisfaction, and authenticity. Current virtual environments often fail to capture the sensory richness and emotional depth of in-person gatherings, leading to a lesser sense of connection amongst users.

Several studies have shown that auditory enhancements might be the next stop in bridging the gap between real life interactions and virtual interactions. A study by Angelika C. Kern, “Audio in VR: Effects of a Soundscape and Movement-Triggered Step Sounds on Presence” demonstrates how soundscapes and movement-triggered sounds can enhance the sense of presence in VR. This suggests that sophisticated audio systems can significantly improve realism and user engagement (Kern). Additionally, research done by Elizabeth T. Davis indicates that audio cues enhance visual perception and spatial awareness, helping users feel more 'present' in virtual settings (Davis).

Despite these findings, most virtual gathering platforms these days lack effective spatial audio cues, resulting in conversations that feel disjointed and impersonal. Real-world social interactions benefit from the intuitive cues provided by the direction and volume of sound, aspects that are often poorly replicated in virtual environments. This study’s goal is to address these issues by dynamically adjusting audio to reflect the intensity and direction of voices, thereby enhancing the perception of proximity and interaction in virtual spaces. Research by Martin Gibbs on proximity-based chat systems in video games illustrates the potential of audio to create more lifelike and engaging virtual social interactions (Gibbs). This was my main inspiration for creating a virtual space with dynamic audio cues based on proximity.

### 1.1 Research Questions

This study was guided by the following research questions:

- How do spatial audio cues impact the user’s sense of engagement and realism while interacting in a virtual space?
- Does the implementation of directional and spatial audio improve the user’s sense of presence and emotional connection with others in a virtual space?
- How intuitive do users find the interface and audio controls for navigating and interacting within the virtual space?
- How much does the availability of interaction options influence the user’s perceived versatility of the platform?
- How well does each room encapsulate a different atmosphere?

## 2. METHOD

The study protocol was designed to evaluate the impact of spatial audio cues on user engagement and realism within a virtual environment, focusing on three specific scenarios: a virtual birthday party, a professional networking event, and a casual game night. Each participant was given a structured walkthrough of the virtual platform, beginning with an overview of the platform's goals and a tutorial on navigating and using its features. The introduction was standardized to ensure all participants received the same information and no participant had an advantage over others.

### 2.1 Evaluation Scenarios

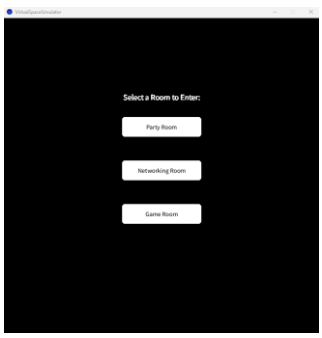


**Figure 1. Main Menu Screen**

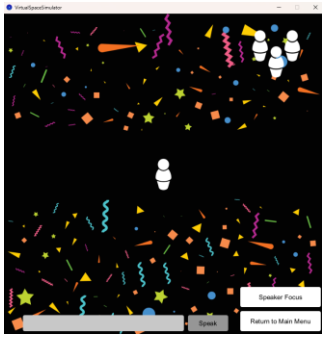


**Figure 2. Party Room**

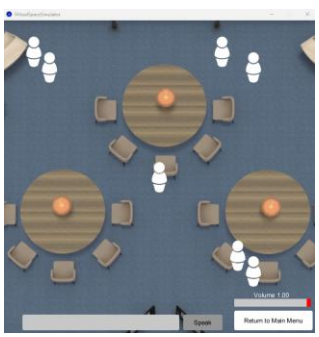

**Figure 3. Networking Room**

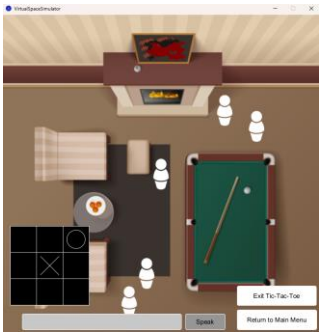

**Figure 4. Game Room**

Participants were guided through the three scenarios in a controlled setting. For each scenario specific tasks were assigned:

i. **Virtual Birthday Party:** Navigate to the party room, engage in a conversation, and use the Speaker Focus Toggle to make a toast.

ii. **Professional Networking Event:** Move within the networking room, engage in conversations, and adjust the Individual Volume Slider to manage audio levels of the conversations.

iii. **Casual Game Night:** Join a game room and play Tic-Tac-Toe, focusing on the ease of interaction and audio cues in a relaxed setting.

These tasks were chosen to evaluate the platform's capability to host engaging social interactions and to test the intuitiveness and effectiveness of spatial audio cues in various social contexts.

## 2.2 Data Collection

Data Collection was conducted through the following qualitative and quantitative methods:

i. **User Interaction Logs:** For more accurate quantitative data, a user interaction log feature was coded into the simulator. These logs are .txt files that record real-time interactions when a user clicks on the screen within the simulator. The logs contain essential data such as the timestamp of each click, the x and y coordinates of the mouse cursor at the moment of click, and the current scenario or context in which the interaction occurs.

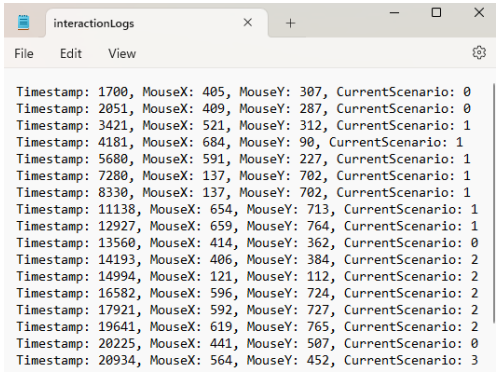


**Figure 5. Example Raw Interaction Log .txt File**

ii. **Task Completion Time:** Measured to assess the intuitiveness of navigation and interaction within the platform.

iii. **Behavioral Observations:** Notes were taken on user behavior and emotional responses during the tasks to gather qualitative insights.

iv. **Spatial Proximity Estimation:** After each scenario, participants were asked to estimate the proximity of voices they interacted with and describe their sense of space and directionality.

Post-experience, each participant was interviewed to collect qualitative data on their overall experience and perceptions of realism and engagement. An anonymous survey was also conducted to gather additional feedback.

## 2.3 Data Control

To ensure consistency across sessions, all user positioning and audio cues were predetermined using JSON data, setting fixed points for interactions and conversations within the virtual environment. This controlled setup allowed us to isolate the impact of spatial audio on the user experience without other variables influencing the outcomes.

## 2.4 Shortcomings

The study's shortcomings include a small and homogeneous sample size, consisting only of three college students. This prevented the collection of data from reflecting broader user experiences. Additionally, the simulator's inability to facilitate real-time interactions with real users meant that all scenarios were predefined and controlled, limiting the dynamic and spontaneous nature of real-world social interactions. This setup potentially impacted the realism and engagement levels perceived by participants, constraining the assessment of spatial audio's effectiveness in truly immersive virtual environments.

# 3. RESULTS

This section will present the findings from three participants who were guided through a series of scenarios on a virtual platform designed to test the impact of spatial audio on user engagement and realism. Given the small sample size, the analysis includes both quantitative metrics and qualitative observations to provide a comprehensive view of user experience.

## 3.1 Quantitative Data

### *3.1.1 User Interaction Logs*

The following graphs plotted using matplotlib show a comparison of the sample user positions to the click positions of the participants. The positional click data was retrieved from the user interaction logs.

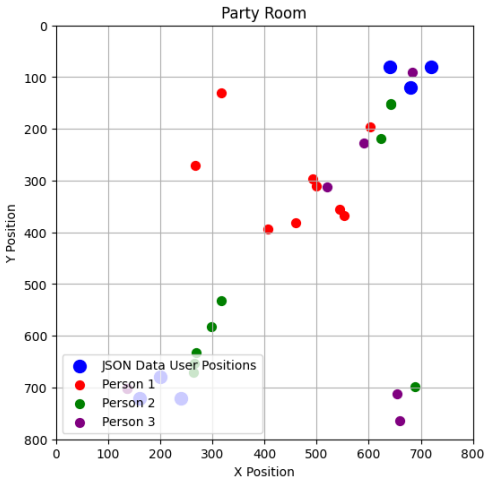


**Figure 5. Party Room Participant Clicks**

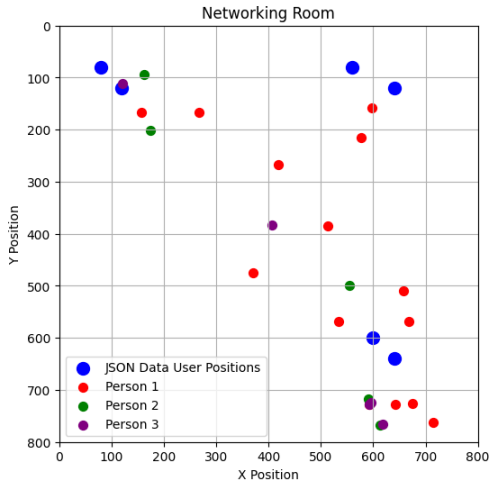


**Figure 6. Networking Room Participant Clicks**

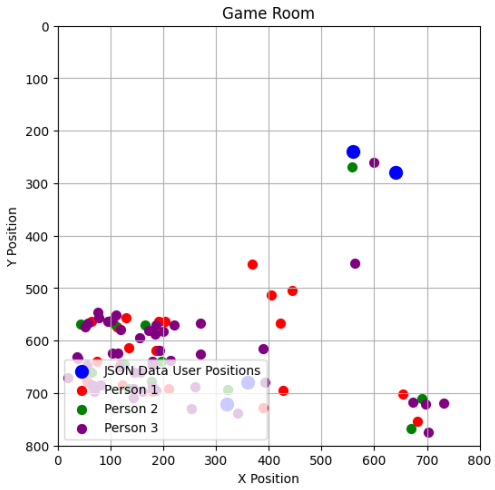


**Figure 7. Game Room Participant Clicks**

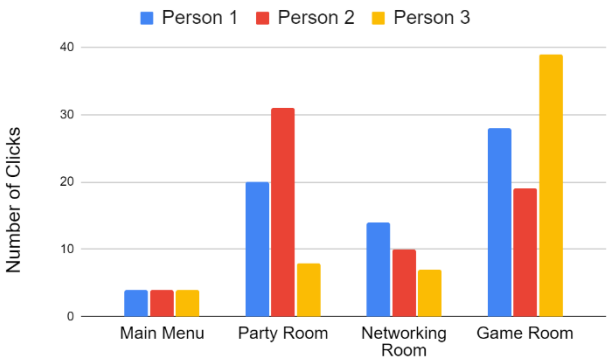


**Figure 8. Overall Number of Clicks Per Participants**

The plots created from the log data show varied engagement across scenarios. The Game Room saw the highest activity, especially with Person 3. The bottom left corner in the Game Room also contains the area where the interactive Tic Tac Toe board pops up, so the concentration of clicks in that area signifies that the participants were able to engage with the UI elements. We can also see how participants' click positions are close by the sample users, showing that they were able to listen and engage with their conversations. Overall, these findings are encouraging, demonstrating that participants could adeptly move towards and engage with other users, thereby affirmatively assessing the virtual environment's navigational and interactive design.

### *3.1.2 Task Completion Time*

The task completion times for each room were calculated from the user interaction logs.

**Table 1. Task Completion Time Raw Data (In Seconds)**

| Scenarios | Person 1 | Person 2 | Person 3 |
|---|---|---|---|
| Main Menu | 26.53 | 23.45 | 18.53 |
| Party Room | 23.23 | 13.05 | 10.14 |
| Networking Room | 11.02 | 6.61 | 6.03 |
| Game Room | 29.87 | 9.36 | 14.4 |

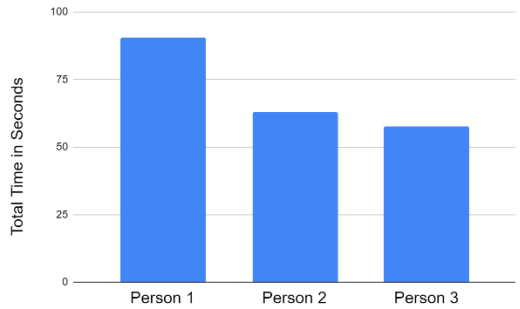


**Figure 9. Total Task Completion Time Per Participant**

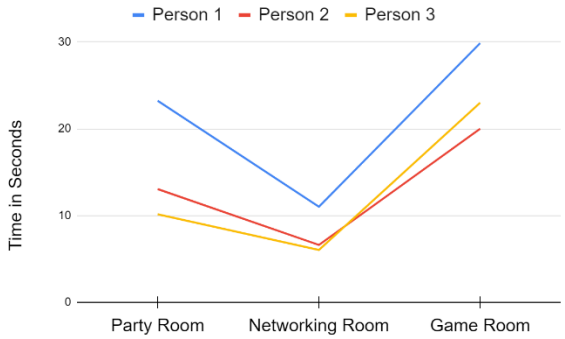


**Figure 10. Time Spent in Each Scenario Per Participant**

The task completion time across various scenarios reveals some interesting insights into user behavior and platform interaction. The significant decrease in completion times from the Party Room to the Networking Room indicates that participants quickly adapted to the virtual environment's flow, improving efficiency in similar settings.

### *3.1.3 Perceived Proximity Accuracy*

To assess the accuracy of perceived proximity within the virtual environment, participants estimated the distance to sound sources relative to the positioning of other sample users. The following table presents the accuracy of participants' estimates as a percentage of the actual distances.

**Table 2. Participant Accuracies**

| Scenarios | Person 1 | Person 2 | Person 3 |
|---|---|---|---|
| Party Room | 98% | 87% | 83% |
| Networking Room | 93% | 94% | 89% |
| Game Room | 95% | 88% | 90% |

The perceived proximity accuracy among participants was generally high across all scenarios, indicating their ability to accurately sense the location of other users within the virtual environment. These results suggest that participants relied on sound cues to gauge the proximity of other users, highlighting the effectiveness of spatial audio in creating immersive virtual environments. Further refinement of audio spatialization techniques could potentially enhance users' spatial awareness and interaction experiences in virtual spaces.

## 3.2 Qualitative Data

### *3.2.1 Recurring Themes in Feedback*

- Realism: Participants noted that spatial audio cues made virtual environments feel more lifelike. One mentioned, "It felt like people were actually around me, not just sounds coming from my headphones."
- Engagement: Users reported higher engagement levels during scenarios with more interactive UI elements, particularly the Tic Tac Toe board in the Game Room.
- Emotional Connection: Several comments included feelings of being 'part of the space' especially when audio cues were used to track the positioning of other users.

### *3.2.2 Behavioral Observations*

Throughout the evaluation sessions several behavioral observations were recorded as the participants navigated through the simulator. All participants seemed to exhibit a natural inclination to explore different areas within the virtual space. For example, all three participants clicked around more in positions without other sample users just to test out the feature of the dynamic audio and the visually updating avatar. Also, the interactive UI elements seemed to help the participants engage with the virtual space more as it allowed them to feel like they were also adding on their presence to the space in real time. For example, the speaker focus button plays a Happy Birthday tune when pressed, and participants seemed excited to wish the sample users on the screen a Happy Birthday. There was a notable increase in activity in the Game Room, where participants seemed particularly drawn to the interactive Tic Tac Toe board. This suggests that interactive elements within the environment captured the participants' interest and encouraged interaction. Additionally, participants demonstrated social behaviors such as approaching groups of users to engage in conversations, indicating a desire for social interaction and connection within the virtual environment. Overall, the behavioral observations underscored the participants' immersion in the virtual space and their willingness to explore and interact with the platform's features, highlighting the potential for virtual environments to facilitate engaging and interactive experiences.

## 3.3 Anecdotal Evidence

- Participant 1: "Figuring out where to join conversations using sound was pretty straightforward. It made it easier for me to decide where I wanted to go next."
- Participant 2: "The controls were easy to understand, and I had no trouble moving around and interacting with the environment. It felt natural, like I've done it before."
- Participant 3: "I liked the interactive elements; they added a nice touch to the overall experience. It made the virtual environment feel more engaging and lifelike."

# 4. DISCUSSION

The study results provided insightful observations into how spatial audio cues enhance engagement and realism within a virtual space. The significant correlation between the presence of spatial audio cues and user engagement was evident, as participants demonstrated a higher level of interaction and immersion, particularly highlighted by their navigation and interaction patterns. This suggests that spatial audio significantly impacts a user's sense of presence and emotional connection within a virtual space, thereby affirming the hypothesis that directional and spatial audio can enhance the feeling of being part of the virtual environment. Furthermore, the high task completion efficiency in different scenarios shows that users found the interface and audio controls intuitive, supporting smooth navigation and interaction. The variation in engagement levels across different rooms, particularly with increased interactions where more dynamic audio cues were present, indicated the effectiveness of these audio elements in creating a diverse atmospheric experience within each virtual space.

## 4.1 Strengths and Weaknesses

One of the key strengths of this study was the integration of both quantitative and qualitative data, which provided a multi-dimensional view of user experiences. Quantitative data from user interaction logs and task completion times offered concrete evidence of user behavior, while qualitative feedback and behavioral observations provided deeper insights into user perceptions and interactions. However, the study also presented several weaknesses. The limited participant number (only three) could affect the generalizability of the findings. Moreover, while the study captured immediate user reactions and interactions, it did not assess long-term engagement or the effectiveness of spatial audio over extended periods, which could be crucial for understanding user adaptation and sustained interest.

## 4.2 Study Design and Constraints

The study design, focusing on immediate interaction within a controlled virtual environment, effectively captured the impact of spatial audio on user engagement and realism. However, this is notably different from interactions with actual users in real time within a virtual space. Additionally, the study only focus on a few predefined scenarios and settings, which may not have fully explored the potential of spatial audio in a wider range of virtual environments.

# 5. CONCLUSION

Significant insights into the effects of spatial audio on user engagement and realism in virtual environments have been revealed through this research. Through a combination of quantitative metrics and qualitative observations, the study shows evidently that spatial audio cues are critical in enhancing the virtual experience. The audio elements significantly improved participants' sense of engagement and realism, fostering a deeper sense of presence and emotional connection within the virtual environment. Participants were also able to accurately localize sounds and other users, which facilitated more natural and intuitive interactions.

Each room's atmosphere was distinctly enhanced by different genres of sample conversations and interactive elements, which helped encapsulate different thematic experiences effectively. The variety of interaction options available influenced users' perceptions of the platform's versatility. This leaves room for the virtual space being used to adapt to different user preferences.

Feedback on the intuitiveness of the interface and audio controls was overwhelmingly positive, indicating that users found the navigation straightforward and user-friendly. This ease of interaction highlights the importance of intuitive design in virtual environments, enabling users to interact without the need for complex instructions.

To address the limitations of this study and to enhance the robustness of future research, more evaluations should be conducted with a larger and more diverse participant base. Expanding the scenarios to include a wider variety of environments and longer interaction periods could also provide more comprehensive insights into the impacts of spatial audio. The addition of more interactive features beyond the given UI elements and the chat box could significantly enrich the user experience.

For example, incorporating tools that allow users to personalize their avatars or manipulate environmental elements could help them express themselves more uniquely and vividly in the virtual space. These steps would move us closer to creating a virtual environment that not only mimics but feels as nuanced and responsive as real life.

# 6. REFERENCES


[1] Davis, E. T., Scott, K., Pair, J., Hodges, L. F., & Oliverio, J. (1999). Can Audio Enhance Visual Perception and Performance in a Virtual Environment? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 43(22), 1197-1201. https://doi.org/10.1177/154193129904302206

[2] Gibbs, Martin, et al. Proximity-Based Chat in a First Person Shooter: Using a Novel Voice Communication System for Online Play. 12 2006.

[3] Kern, Angelika C., and Wolfgang Ellermeier. “Audio in VR: Effects of a Soundscape and Movement-Triggered Step Sounds on Presence.” *Frontiers*, Frontiers, 4 Feb. 2020, www.frontiersin.org/articles/10.3389/frobt.2020.00020/full