Exploring Visual Scanning in Augmented Reality: Perspectives From Deaf and Hard of Hearing Users

Sanzida Mojib Luna
Niantic x RIT Geo Games and Media Research Lab
Rochester Institute of Technology
Rochester, New York, USA
sl8472@g.rit.edu

Garreth W. Tigwell
School of Information
Rochester Institute of Technology
Rochester, New York, USA
garreth.w.tigwell@rit.edu

Konstantinos Papangelis
Niantic x RIT Geo Games and Media Research Lab
Rochester Institute of Technology
Rochester, New York, USA
kxpigm@rit.edu

Jiangnan Xu
Niantic x RIT Geo Games and Media Research Lab
Rochester Institute of Technology
Rochester, New York, USA
jx3896@rit.edu

Abstract

Sensory-intensive and attention-demanding tasks like visual scanning, interacting with 3D objects, comprehending and following instructions, etc. are becoming more common in Augmented Reality (AR) environments as the technology expands through diverse fields. It is important to understand how these types of tasks are experienced by Deaf and Hard of Hearing (DHH) people, especially if those tasks involve any sound or compete with attention shifts (e.g., observing someone signing) in both real and virtual environments. Our current research specifically aims to identify the challenges that DHH users encounter when engaging in visual scanning in an AR environment. Using *Angry Birds AR* as a probe in our research, 11 DHH participants, with varying hearing abilities played seven rounds of the game, followed by a short structured interview and a long semi-structured interview. Our findings revealed that subtle audio cues and excessive visual indicators impacted participants' performances negatively. Additionally, they positioned themselves strategically for maximum spatial awareness but faced challenges with AR visual cues due to the lighting conditions in the real environment. We further suggested design implications such as customizable, user-friendly haptic and textual feedback, and intelligent spatially aware mechanisms for AR.

CCS Concepts

• **Human-centered computing** → **Accesibility**; **Mixed / Augmented reality**.

Keywords

Human-computer Interaction; Accessibility; Augmented Reality

ASSETS '24, October 27–30, 2024, St. John's, NL, Canada

ACM ISBN 979-8-4007-0677-6/24/10
https://doi.org/10.1145/3663548.3688535

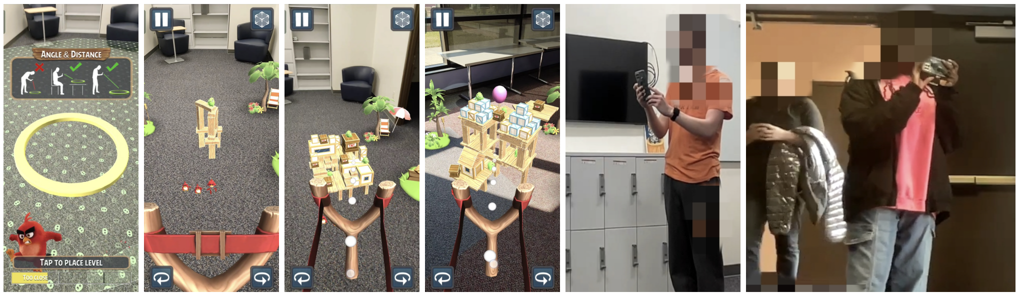
ACM Reference Format:
Sanzida Mojib Luna, Garreth W. Tigwell, Konstantinos Papangelis, and Jiangnan Xu. 2024. Exploring Visual Scanning in Augmented Reality: Perspectives From Deaf and Hard of Hearing Users. In *The 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24), October 27–30, 2024, St. John's, NL, Canada.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3663548.3688535

1 Introduction

Augmented Reality (AR) as a technology is spanning across various domains, including education [30], healthcare [5, 33], training [13, 14], entertainment [9], etc. With the advancements of AR, the technology is not limited to overlaying digital objects in the real environment only; rather, the scope of tasks users can perform within these environments has expanded significantly. Today, most AR environments require users to perform several common tasks, such as visually scanning the surrounding environment [14], interacting with and manipulating 3D elements [23], and following instructions [12]. In AR, these tasks are uniquely challenging, as they require users to process information from the real world and digital overlays, often while navigating or interacting with the environment.

For instance, visual scanning tasks in AR involve systematically observing and interpreting both the real-world environment and virtual overlays through devices like smartphones or AR glasses. Visual scanning tasks involve i) searching for and identifying both physical and virtual elements in the augmented space (Figure 1a), ii) distinguishing between real and virtual objects (Figure 1a and 1b), iii) interpreting visual cues and indicators overlaid on the real environment (Figure 1c and 1d), iv) adapting to changes in both the physical surroundings and the digital overlays (Figure 1c, 1d, 1e, and 1f), and v) maintaining spatial awareness of the real environment while engaging with virtual elements (Figure 1e and 1f).

Deaf and Hard of Hearing (DHH) people encounter challenges in comprehending specific audio cues [32] in various real and virtual environments while doing these tasks. They rely more on visual and haptic cues [21] and frequently shift attention to maintain spatial awareness [11]. Despite these needs, AR environments lack sufficient scholarly attention to the challenges DHH people might face while performing these tasks. Identifying the challenges can



(a) (b) (c) (d) (e) (f)

Figure 1: Instances of visual scanning tasks in *Angry Birds AR* , (a) an interface of a device screen of a player looking for a suitable place for placing the AR objects in the augmented space, (b) virtual objects are defined by vibrant colors, (c) in a private artificially-lit environment, virtual cues (white dots) indicating projectile path, which the shot bird will follow, (d) in a public naturally-lit environment, the same virtual cues (white dots) indicating projectile path, (e) a player playing in a private environment where he has to be mindful of the surrounding objects, (f) a player playing in a public environment where she has to be mindful of the surrounding objects and moving non-player(s).

help us create design guidelines for more inclusive and immersive AR environments accessible to users regardless of their hearing abilities.

To gain this understanding, we used a mobile AR single-player game (*Angry Birds AR*) as a probe[1] to investigate how DHH users with varying hearing abilities perform tasks such as visual scanning, interacting with AR objects, and following instructions. This study focuses on the challenges that DHH users face during visual scanning tasks in AR when using handheld devices such as smartphones. Our work addresses the gap in understanding the challenges DHH users encounter while engaging in visual scanning in AR.

In our study, 11 DHH participants played seven rounds of the game, followed by a short structured interview and a long semi-structured interview. Our research revealed that participants' performance was negatively impacted by subtle audio cues and excessive visual indicators in AR. Additionally, participants strategically positioned themselves to enhance spatial awareness but faced challenges perceiving visual cues in AR due to their positioning and real-world lighting conditions. Based on these findings, we recommended several design implications for AR, including customizable and user-friendly textual and haptic feedback, dynamic environmental adjustments, and intelligent, spatially aware mechanisms to address these challenges.

Our findings revealed unique challenges for DHH users in the AR environment, different from non-AR contexts. Unlike traditional digital interfaces where DHH users interact with a single-screen space, AR demands engagement with both physical and virtual elements, creating a complex visual landscape. This dual interaction increases the impact of subtle audio cues and excessive visual indicators, as users must concurrently process information from both realms. Additionally, balancing spatial awareness in the physical world with visibility in the AR overlay presents a challenge specific to AR. These findings highlight the unique cognitive load AR imposes on DHH users.

[1]Here we refer probe as a tool for collecting data [3].

2 Related Work

2.1 AR For DHH Users

AR has been used effectively to create accessible and assistive environments that prioritize the requirements of DHH users. For example, AR combined with automatic speech recognition (ASR) and text-to-speech synthesis (TTS) has been used to provide live captions in real time for DHH users [19]. Additionally, significant research has focused on the position [10], sound identity [7], and customization [22] of live AR captions within the gaze of DHH users. In addition, various virtual sign language interpreters have been designed and implemented for different contexts, such as home entertainment [28] and classrooms [17], enhancing the daily lives of DHH users. Moreover, existing research has investigated how AR can create and enhance accessible communication of DHH users with various devices, such as smart speakers in home settings [18], as well as among DHH people with diverse hearing abilities and communication methods within collaborative AR environments [16].

2.2 DHH Users and Visual Scanning

Several studies have been conducted regarding visual scanning, in both real and virtual environments, in particular, eye-tracking involving DHH users. In visual languages, such as sign language, eye gaze plays crucial roles (i.e., gaze patterns accompanying types of verbs [25], locative pronouns [26], gaze fixation on the face and upper-body [20], etc.), which has been demonstrated by research using eye-tracking. In addition, the perception of emotions by DHH users using facial features under different conditions [1] and facial and body postures [2] has been investigated using visual scanning.

The work aforementioned in 2.1 has focused mainly on using AR to improve the awareness of DHH users about their surroundings. In contrast, the work in 2.2 has utilized visual scanning to understand various perspectives of DHH communication and how they perceive their surroundings. However, these studies overlook the barriers DHH users might face while performing tasks related to visual scanning in AR environments.

3 Method

Our study was divided into two main segments: (i) gameplay using the AR game *Angry Birds AR* [24] followed by a short interview and (ii) one-on-one longer interviews. *Angry Birds AR* is a single-player AR game, with a first-person view. The game is ideal for our needs, being user-friendly for those with little AR experience and not relying on audio cues for essential information. This makes it suitable for offering DHH people with firsthand experience of visual scanning in AR. The game tasks involve setting up an AR environment, following instructions, interacting with AR objects, etc., to advance the levels. Players start by using their smartphone camera to scan the environment and locate pigs within the AR structure. They must find a flat, non-reflective, and textured surface to place the AR island and the wooden tower, guided by visuals on the screen and a bottom bar indicating the suitability of the surface (Figure 2). Players scan an area of about 5 to 8 square feet to set up the AR structure and begin the gameplay. The goal is to use a slingshot to launch birds at wooden towers occupied by pigs, aiming to knock them all down.

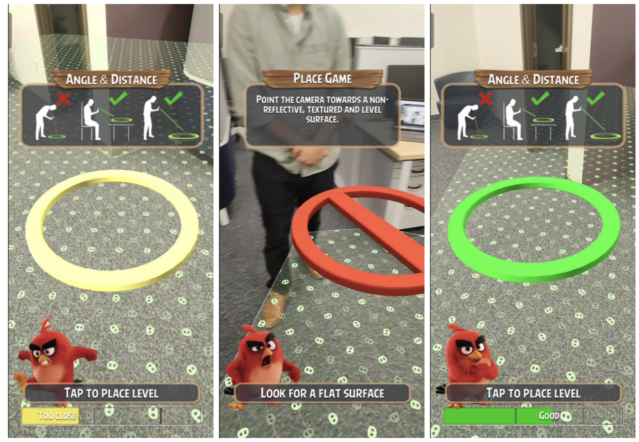
3.1 Participants

We promoted our study by distributing flyers at a local college and sending emails to students and staff. Interested individuals were directed to complete an online "Participant Registration" form. This form gathered details, including preferred pronouns, DHH identification, weekly gaming hours, and familiarity with AR. We recruited 11 participants, aged 19 to 28 years. Among them, eight were hard of hearing, and three were deaf. Seven preferred spoken English, while four preferred American Sign Language (ASL). Through the interviews that we conducted after the gameplay, we found out that, all the participants were familiar with the 2D version of our selected game, however, none of them had any experience with the AR version of it. Similarly, participants' AR experience spanned from novice to expert levels based on their registration responses. Please refer to Appendix A for more information about the participants' demographics.

3.2 Procedure

Participants attended a gameplay session in our lab where they received a briefing about the game and session procedures. Each participant played the game individually on a provided smartphone, Realme 8 Pro (RMX3081) with a 6.4-inch screen, using the default settings of the game, such as - music, sound, and vibration enabled. For ASL users, we provided interpreters to facilitate communication between participants and researchers during gameplay sessions and both interviews. Explicit consent was obtained for external video and audio recordings of the gameplay, short and long interviews. Participants used the think-aloud [6] method, verbally or through signing, to express their thoughts during gameplay, and were instructed to record their device screens.

Participants completed seven game levels in two settings: the first three in an artificially lit research lab (private setting) and the remaining four in a naturally lit hallway (public setting). It is important to note that, participants voluntarily offered their thoughts while playing at the same time, without interrupting their gameplay or impacting their engagement with the game. Their actions were video recorded for later observational analysis, and game progress was reset after each session. Participants then returned to the lab for a brief interview. The interview covered their initial reactions, preferred game space, favorite features and reasons behind it, suggestions for improvements, and reasons for suggesting said



(a) (b) (c)

Figure 2: Interfaces of *Angry Birds AR* before the game starts. The game is instructing the player (a) to look for a space in a further position, (b) to look for a flat space, and (c) the space is good enough for AR placement.

(a)

(b)

Figure 3: Player is holding the slingshot with a bird to shoot in (a)private setting, and (b) public setting. The white dots show projectile paths where the bird might land.

improvements. Participants were invited to longer, semi-structured online interviews via Zoom [34] at a later date, where they were asked about their gameplay experiences, challenges, and recommendations for improvements. An ASL interpreter assisted ASL users and the researchers in communication. Each participant received a $50 e-gift card for their time.

4 Findings

For data analysis, we used data triangulation [27], incorporating gameplay sessions, screen recordings, interview recordings, and observational notes. All data were imported into NVivo [15] and coded using thematic analysis [4], employing both semantic and latent approaches to identify recurring themes. We identified two key themes regarding visual scanning: the impact of subtle audio cues and excessive visual cues, and the challenge of balancing real-world spatial awareness with AR visibility amidst varying environmental conditions.

4.1 Impact of Subtle Audio Cues and Excessive Visual Indicators

Five out of 11 participants reported confusion during visual scanning in the AR interface. They faced challenges grasping the meaning of specific visual indicators while placing the AR structure, despite multiple indicators being present simultaneously on UI (Figure 2), all aimed at guiding them towards a particular task. For Instance, P04 said,

> "I think you [researcher] told me that you're supposed to tap it. Because when I was trying to find a perfect surface, there's a bar in the bottom and I couldn't understand what it exactly meant." [P04, HoH]

Although there were no significant audio instructions, the background music and sound contributed to task engagement by providing a sense of immersion and continuity. This was further confirmed by P06 who was hard of hearing and could process the background sound and music,

> "If it had no sound, I think it would negatively impact my gameplay. Because then I would feel like my actions had no impact ." [P06, HoH]

Participants' recommendations were more clear visual feedback, such as "written instructions only" [P03], or "one indicator at a time" [P11]. Additionally, all participants (11/11) had to replay at least one of the seven levels. When asked why, a few (3/11) said the game directed them to replay even though they thought they had defeated all the targets (knocked all the pigs off the AR tower). P11 explained,

> "I think in one round I didn't see any pigs, so I tried to break the tower with the last bird, and then I had to play the round again" [P11, HoH]

We observed the screen recordings and found that one of the targets (pigs) was still on the other side of the tower. Participants could either move around the AR structure physically or use on-screen rotate buttons to view the other side of the structure. Moreover, the targets made subtle sounds when they were on top of the AR structure. However, participants could not find them, assumed all targets were defeated, and focused elsewhere (destroying the AR tower) instead. In both scenarios—placing AR structures into the real world and searching for targets—participants missed the subtle audio cues, which, while not crucial, affected their task performance negatively.

4.2 Balancing Spatial Awareness In Real World And Visibility in AR

Participants played last four rounds of game in an L-shaped hallway which had few pieces of furniture, public access, and was naturally lit. When we asked them about which space they preferred for the gameplay, 7 out of 11 expressed their preference for a larger space, which was the hallway. They explained that the additional space allowed for greater movement and offered more flexibility for visual exploration. We observed most of the participants (10/11) positioned themselves at the cross-section of the hallway. We also observed at least one non-player pass by when they were playing. Few of the participants (3/11) also mentioned about visual challenges due to sunlight (Figure 3).

> "The white dots [projectile path] were harder to see in the other [second] location because of the sunlight. So, I actually preferred the first location for that reason." [P03, Deaf]

We observed that participants were reluctant to change their positions despite visual challenges that could be mitigated by doing so. This reluctance suggested that altering their position might reduce their awareness of non-players moving around them, even though it would enhance the visibility of visual cues in AR. Additionally, over half of the participants (6/11) expressed concerns about being distracted by non-players during gameplay.

> "There were maybe one or two people passing by. But if there's a lot of people around me, I can see that being a little bit awkward and would definitely be distracting." [P10, HoH]

5 Discussion and Limitations

Participants required more time to position the AR structure, a key visual scanning task in the game, and often had to replay levels. This was largely due to challenges in processing subtle audio cues, including background music and noise from the targets, which were intended to guide the player toward the targets. These audio cues, crucial for immersion and performance in AR [31], were challenging to grasp for DHH users with varying levels of hearing abilities, especially those with minimal to no residual hearing. To address this, participants suggested alternative methods like textual cues ("written instructions only," [P03]) and dynamic environments ("one indicator at a time," [P11]). Furthermore, haptic feedback can be an efficient way to indicate off-screen targets that can assist in maintaining task continuity and enhance immersion and performance in AR for DHH users. Additionally, excessive visual cues led to confusion, negatively affecting performance as DHH users rely heavily on visual information [8]. This aligns with web design guidelines [29], indicating the need for clear, user-friendly visual feedback systems and customizable indicators in AR for DHH users.

The strategic positioning of the participants in public settings suggested that they chose to enhance spatial awareness while interacting with AR elements. They wanted to be able to see the moving non-players to avoid disturbing and being distracted by them. However, this positioning caused glare from direct sunlight, making it difficult to see visual cues in AR like white dots. They were reluctant to change their position for better visibility of AR cues, fearing it might compromise their awareness of moving non-players in the surroundings. They also expressed concerns about playing in crowded areas, where increased attention to moving non-players would further distract them from the task. The challenge requires integrating spatially aware features like visual and haptic cues for moving people and objects, and mini-maps of the environment into AR. These features would help DHH users enhance spatial awareness without compromising AR task performance. Designers must balance these visual aids to avoid overwhelming users while ensuring effectiveness.

Our study provides valuable insights into DHH users' experiences with AR, yet it is important to acknowledge its limitations. Although effective for engagement, the game-based AR environment may limit our findings' generalizability to other AR applications. Games typically feature more structured interactions and clearer goals than utilitarian AR applications in education, healthcare, or workplace training. The gamified nature of the task might have influenced participants' behavior and attention patterns in ways not fully reflective of real-world AR use. In addition, including hearing participants in the study would have strengthened our findings by allowing us to compare visual scanning patterns. This comparison would help confirm that the challenges identified are unique to DHH participants in AR and highlight the need to address them. Moreover, the simplicity of the game's visual scanning tasks may not capture the complexity required in more sophisticated AR applications. Furthermore, while eye-tracking could provide precise gaze data, it might have altered participants' behavior and reduced ecological validity. Our approach enabled the natural engagement of participants with AR, providing insights into authentic visual scanning strategies and challenges.

6 Conclusion and Future Work

Our research is the primary step to bridge the gap in understanding the challenges of DHH users with varying hearing abilities while performing visual scanning in AR. We recruited 11 DHH participants who engaged in seven rounds of the AR game *Angry Birds AR*. Following gameplay, each participant joined a short structured and a longer semi-structured interview. Our findings revealed that subtle audio cues and excessive visual indicators negatively impacted participants' performance. This can be mitigated with alternative approaches, such as dynamic environments and user-friendly haptic and textual cues. Additionally, participants had to compromise AR visibility to maintain real-world spatial awareness. Incorporating intelligent spatially aware mechanisms in AR could enhance performance without shifting attention. Despite our limitations, these challenges are likely relevant in various AR contexts. Future research should validate and expand these findings in diverse AR applications to better understand the needs of DHH users in AR environments.

References

[1] Maria Bianca Amadeo, Andrea Escelsior, Mario Amore, Gianluca Serafini, Beatriz Pereira da Silva, and Monica Gori. 2022. Face masks affect perception of happy faces in deaf people. *Scientific Reports* 12, 1 (2022), 12424. https://doi.org/10.1038/s41598-022-16138-x

[2] Brittany A Blose and Lindsay S Schenkel. 2024. Facial and Body Posture Emotion Identification in Deaf and Hard-of-Hearing Young Adults. *Journal of Nonverbal Behavior* (2024), 1–17. https://doi.org/10.1007/s10919-024-00458-9

[3] Kirsten Boehner, Janet Vertesi, Phoebe Sengers, and Paul Dourish. 2007. How HCI interprets the probes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>San Jose</city>, <state>California</state>, <country>USA</country>, </conf-loc>) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 1077–1086. https://doi.org/10.1145/1240624.1240789

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[5] Mathilde R. Desselle, Ross A. Brown, Allan R. James, Mark J. Midwinter, Sean K. Powell, and Maria A. Woodruff. 2020. Augmented and Virtual Reality in Surgery. *Computing in Science & Engineering* 22, 3 (2020), 18–26. https://doi.org/10.1109/MCSE.2020.2972822

[6] David W Eccles and Güler Arsal. 2017. The think aloud method: what is it and how do I use it? *Qualitative Research in Sport, Exercise and Health* 9, 4 (2017), 514–531. https://doi.org/10.1080/2159676X.2017.1331501

[7] Ru Guo, Yiru Yang, Johnson Kuang, Xue Bin, Dhruv Jain, Steven Goodman, Leah Findlater, and Jon Froehlich. 2020. HoloSound: Combining Speech and Sound Identification for Deaf or Hard of Hearing Users on a Head-Mounted Display. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) *(ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 71, 4 pages. https://doi.org/10.1145/3373625.3418031

[8] Ramy Hammady, Minhua Ma, Carl Strathern, and Mostafa Mohamad. 2020. Design and development of a spatial mixed reality touring guide to the Egyptian museum. *Multimedia Tools and Applications* 79, 5 (2020), 3465–3494. https://doi.org/10.1007/s11042-019-08026-w

[9] Shiu-Wan Hung, Che-Wei Chang, and Yu-Chen Ma. 2021. A new reality: Exploring continuance intention to use mobile augmented reality for entertainment purposes. *Technology in Society* 67 (2021), 101757. https://doi.org/10.1016/j.techsoc.2021.101757

[10] Dhruv Jain, Bonnie Chinh, Leah Findlater, Raja Kushalnagar, and Jon Froehlich. 2018. Exploring Augmented Reality Approaches to Real-Time Captioning: A Preliminary Autoethnographic Study. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems* (Hong Kong, China) *(DIS '18 Companion)*. Association for Computing Machinery, New York, NY, USA, 7–11. https://doi.org/10.1145/3197391.3205404

[11] Raja S. Kushalnagar, Gary W. Behm, Aaron W. Kelstone, and Shareef Ali. 2015. Tracked Speech-To-Text Display: Enhancing Accessibility and Readability of Real-Time Speech-To-Text. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 223–230. https://doi.org/10.1145/2700648.2809843

[12] Gun A. Lee, Seungjun Ahn, William Hoff, and Mark Billinghurst. 2020. Enhancing First-Person View Task Instruction Videos with Augmented Reality Cues. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 498–508. https://doi.org/10.1109/ISMAR50242.2020.00078

[13] Kangdon Lee. 2012. Augmented reality in education and training. *TechTrends* 56 (2012), 13–21. https://doi.org/10.1007/s11528-012-0559-3

[14] Mark A Livingston, Lawrence J Rosenblum, Simon J Julier, Dennis Brown, Yohan Baillot, J Edward Swan, Joseph L Gabbard, Deborah Hix, et al. 2002. An augmented reality system for military operations in urban terrain. In *Interservice/Industry Training, Simulation, and Education Conference*, Vol. 89.

[15] LumiVero. n.d.. NVivo - Lumivero. https://shorturl.at/h8mn9. [Accessed 05-06-2024].

[16] Sanzida Mojib Luna, Jiangnan Xu, Konstantinos Papangelis, Garreth W. Tigwell, Nicolas Lalone, Michael Saker, Alan Chamberlain, Samuli Laato, John Dunham, and Yihong Wang. 2024. Communication, Collaboration, and Coordination in a Co-located Shared Augmented Reality Game: Perspectives From Deaf and Hard of Hearing People. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 36, 14 pages. https://doi.org/10.1145/3613904.3642953

[17] Le Luo, Dongdong Weng, Guo Songrui, Jie Hao, and Ziqi Tu. 2022. Avatar Interpreter: Improving Classroom Experiences for Deaf and Hard-of-Hearing People Based on Augmented Reality. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 318, 5 pages. https://doi.org/10.1145/3491101.3519799

[18] Roshan Mathew, Garreth W Tigwell, and Roshan L Peiris. 2024. Deaf and Hard of Hearing People's Perspectives on Augmented Reality Interfaces for Improving the Accessibility of Smart Speakers. In *International Conference on Human-Computer Interaction*. Springer, 334–357. https://doi.org/10.1007/978-3-031-60881-0_21

[19] Mohammad Reza Mirzaei, Seyed Ghorshi, and Mohammad Mortazavi. 2012. Combining Augmented Reality and Speech Technologies to Help Deaf and Hard of Hearing People. In *2012 14th Symposium on Virtual and Augmented Reality*. 174–181. https://doi.org/10.1109/SVR.2012.10

[20] Laura J Muir and Iain EG Richardson. 2005. Perception of sign language and its application to visual communications for deaf people. *Journal of Deaf studies and Deaf education* 10, 4 (2005), 390–401. https://doi.org/10.1093/deafed/eni037

[21] Keita Ohshiro and Mark Cartwright. 2022. How people who are deaf, Deaf, and hard of hearing use technology in creative sound activities. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 66, 4 pages. https://doi.org/10.1145/3517428.3550396

[22] Yi-Hao Peng, Ming-Wei Hsi, Paul Taele, Ting-Yu Lin, Po-En Lai, Leon Hsu, Tzu-chuan Chen, Te-Yen Wu, Yu-An Chen, Hsien-Hui Tang, and Mike Y. Chen. 2018. SpeechBubbles: Enhancing Captioning Experiences for Deaf and Hard-of-Hearing People in Group Conversations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3173574.3173867

[23] Patrick Reipschläger and Raimund Dachselt. 2019. DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 29–41. https://doi.org/10.1145/3343055.3359718

[24] Resolution Games. n.d.. Angry Birds AR: Isle of Pigs — Resolution Games — resolutiongames.com. https://www.resolutiongames.com/angry-birds-ar-isle-of-pigs. [Accessed 05-06-2024].

[25] Robin Thompson, Karen Emmorey, and Robert Kluender. 2006. The relationship between eye gaze and verb agreement in American Sign Language: An eye-tracking study. *Natural Language & Linguistic Theory* 24, 2 (2006), 571–604. https://doi.org/10.1007/s11049-005-1829-y

[26] Robin L Thompson, Karen Emmorey, Robert Kluender, and Clifton Langdon. 2013. The eyes don't point: Understanding language universals through person marking in American Signed Language. *Lingua* 137 (2013), 219–229. https://doi.org/10.1016/j.lingua.2013.10.002

[27] Veronica A Thurmond. 2001. The point of triangulation. *Journal of nursing scholarship* 33, 3 (2001), 253–258. https://doi.org/10.1111/j.1547-5069.2001.00253.x

[28] Vinoba Vinayagamoorthy, Maxine Glancy, Christoph Ziegler, and Richard Schäffer. 2019. Personalising the TV Experience Using Augmented Reality: An Exploratory Study on Delivering Synchronised Sign Language Interpretation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300762

[29] World Wide Web Consortium (W3C). 2018. Web Content Accessibility Guidelines (WCAG) 2.1 - Guideline 1.3 Adaptable. https://www.w3.org/TR/WCAG21/#adaptable. Accessed: 2024-06-26.

[30] Hsin-Kai Wu, Silvia Wen-Yu Lee, Hsin-Yi Chang, and Jyh-Chong Liang. 2013. Current status, opportunities and challenges of augmented reality in education. *Computers & education* 62 (2013), 41–49. https://doi.org/10.1016/j.compedu.2012.10.024

[31] Jing Yang, Prasanth Sasikumar, Huidong Bai, Amit Barde, Gábor Sörös, and Mark Billinghurst. 2020. The effects of spatial auditory and visual cues on mixed reality remote collaboration. *Journal on Multimodal User Interfaces* 14, 4 (2020), 337–352. https://doi.org/10.1007/s12193-020-00331-1

[32] Kyrie Zhixuan Zhou, Weirui Peng, Yuhan Liu, and Rachel F Adler. 2024. Exploring the Diversity of Music Experiences for Deaf and Hard of Hearing People. *arXiv preprint arXiv:2401.09025* (2024). https://doi.org/10.48550/arXiv.2401.09025

[33] Egui Zhu, Arash Hadadgar, Italo Masiello, and Nabil Zary. 2014. Augmented reality in healthcare education: an integrative review. *PeerJ* 2 (2014), e469. https://doi.org/10.7717/peerj.469

[34] Zoom. n.d.. Zoom: One Platform to Connect. https://zoom.us/ [Accessed 05-06-2024].

A Demographic Information of the Participants

Table 1: Participants' Demographic Information.

Players' Pseudo Name	Gender	Age	Experience Using AR	Identity	Preferred Mode of Communication
P01	Male	21	Limited experience	HoH	Spoken English, Written English
P02	Female	22	Heard of it but never had any experience with it	HoH	Spoken English, Written English, American Sign Language (ASL)
P03	Male	19	Professional in AR	Deaf	Written English, American Sign Language (ASL)
P04	Male	21	Limited experience	HoH	Spoken English, Written English
P05	Male	20	Heard of it but never had any experience with it	HoH	Spoken English, Written English
P06	Male	28	Professional in AR	HoH	Written English, American Sign Language (ASL)
P07	Male	22	Limited experience	HoH	Spoken English, Written English
P08	Female	23	Limited experience	Deaf	American Sign Language (ASL)
P09	Female	19	Limited experience	Deaf	American Sign Language (ASL)
P10	Male	20	Limited experience	HoH	Spoken English, Written English, American Sign Language (ASL)
P11	Male	20	Limited experience	HoH	Spoken English, Written English, American Sign Language (ASL)