Exploring The Affordances of Game-Aware Streaming to Support Blind and Low Vision Viewers: A Design Probe Study

NOOR HAMMAD, Carnegie Mellon University, USA
FRANK ELAVSKY, Carnegie Mellon University, USA
SANIKA MOHARANA, Carnegie Mellon University, USA
JESSIE CHEN, Carnegie Mellon University, USA
SEYOUNG LEE, Carnegie Mellon University, USA
PATRICK CARRINGTON, Carnegie Mellon University, USA
DOMINIK MORITZ, Carnegie Mellon University, USA
JESSICA HAMMER, Carnegie Mellon University, USA
ERIK HARPSTEAD, Carnegie Mellon University, USA

This paper explores new ways to support blind and low vision (BLV) game stream participants. Prior work on game-aware streaming systems has focused on the potential for viewer interaction and personalization for sighted viewers, but how such systems impact interaction and personalization for BLV viewers remains largely unexplored. Most streaming experiences have significant visual information but no non-visual or sensemaking alternatives, which can exclude BLV viewers from understanding and interacting with the streaming experience. Building on the pre-existing system MARS, we developed a design probe that makes game data available to stream viewers in personalizable visual and non-visual formats. We use this probe to investigate the needs of BLV game stream viewers through qualitative interviews and live prototype testing sessions on Twitch. In addition to the technical contributions of our probe, our work addresses how game-aware streaming technologies can align with the needs and motivations of BLV viewers, and paves the way for novel designs in future iterations of game-aware streaming technologies.

CCS Concepts: • **Human-centered computing → Participatory design**; **Accessibility design and evaluation methods**.

Additional Key Words and Phrases: Live Streaming, Twitch, Accessibility, Blind and Low Vision, Games Accessibility

ACM Reference Format:
Noor Hammad, Frank Elavsky, Sanika Moharana, Jessie Chen, Seyoung Lee, Patrick Carrington, Dominik Moritz, Jessica Hammer, and Erik Harpstead. 2024. Exploring The Affordances of Game-Aware Streaming to Support Blind and Low Vision Viewers: A Design Probe Study. 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, 21 pages. https://doi.org/10.1145/3663548.3675665

Authors' Contact Information: Noor Hammad, nhammad@cs.cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Frank Elavsky, fje@cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Sanika Moharana, smoharan@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Jessie Chen, jessiech@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Seyoung Lee, snlee@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Patrick Carrington, pcarrington@cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Dominik Moritz, domoritz@cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Jessica Hammer, hammerj@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; Erik Harpstead, harpstead@cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

1 INTRODUCTION

Live streaming via platforms such as Twitch and YouTube has become popular form of content. These platforms have grown exponentially in viewership numbers in the past decade, attracting millions of viewers monthly. While these platforms are used to stream different live content [12, 14], game streams remain the most popular content on Twitch. In a game stream, the streamer shares both their live gameplay and a video feed of themselves. Typically the streamer will use the video to discuss what is happening in the game, talk to their audience, or both [33]. Meanwhile, viewers and moderators participate in a chat that accompanies the stream. Chat can be used to communicate with other viewers, connect with the streamer, or for input commands into the game system [13, 18, 35].

While viewers in game streams value interaction [27], the baseline streaming interface makes participation difficult. First, by virtue of being the one playing the game, the streamer can interact with in-game objects, leaving viewers dependent on the streamer to access game-related information. Second, game streams can become very fast-paced, either due to the genre of the game being played or due to a large volume of chat messages from viewers [13]. Third, game streams are videos and therefore only offer visual information. To date, these challenges have been studied in the context of sighted participants. However, blind and low vision (BLV) viewers are impacted by the same interface limitations while also dealing with additional barriers due to inaccessibility [24]. For instance, the volume of chat can overwhelm assistive tools like screen readers, and the lack of customizability of the viewing interface can prevent users from adapting the experience to their needs.

Systems that offer novel interactions and personalizations for viewers offer an opportunity to address these challenges. These *game-aware* streaming systems aggregate multiple forms of data, such as game data or viewer input data, to create personalized, interactive viewer interfaces that are responsive to the live game being played [20]. Audience Participation Games (APGs), for instance, use viewer input to affect the game mechanics of a streamer's game [15, 35]. Lessel et al's *HelpStone* uses game log data to create viewer overlays for the game *Hearthstone* that support viewer comprehension and improved streamer-viewer communication [28]. The Metadata Augmented Real-time Streaming (MARS) system builds on these ideas to support a broader range of games, as well as allowing viewers to manipulate the viewing interface [20]. Twitch has also identified the need for these novel interactions for viewers and have introduced their own tools which create "Twitch Extensions" that can be installed by a streamer onto their channel [4].

In other words, game-aware streaming systems provide rich metadata streams that can be viewed by the viewer in different forms (e.g., visually or non-visually) that best fit their needs. We know from other work on accessibility that metadata is critical for access, whether it is automatically generated [40] or created by human participants [24]. By using the unique affordances of game-aware metadata, however, we can explore the potential to move beyond merely translating visual content for the consumption of BLV users toward an exploration of a design space that is responsive to their own needs and desires to engage with live streams. We therefore seek to explore how game-aware systems can support access to game streams through three mechanisms: exposing game data in screen reader-friendly format, enabling manipulation of data display properties, and providing tools for exploring real-time and historical game data.

In this paper, we present a design probe study of a game-aware streaming system to better understand how game-aware systems can support BLV viewers. Serving as the design probe is a visual novel called *Bloomwood Stories* [38], which utilizes the MARS system to enable game-aware features. We conducted semi-structured interviews with ten participants who identified as BLV to understand their current use of technologies and existing streaming affordances while viewing game streams. We followed with a live test of the design probe, where a streamer played *Bloomwood Stories* live on Twitch. We then conducted interviews with participants to understand how the prototype impacted their experience and to explore future iterations of game-aware stream systems. Using inductive coding and affinity diagramming, we surfaced themes on how novel features from game-aware streaming might already meet the needs of BLV viewers, as well as identify areas for improvement and innovation. Throughout, our work is centered around two research questions:

(1) How do existing game-aware streaming technologies support, or fail to support, game stream viewing for BLV viewers?
(2) What qualities, challenges, and design opportunities do BLV viewers envision for game-aware streaming technologies?

2 RELATED WORK

Our work examining the affordances of game-aware streaming to support the needs of BLV viewers draws on multiple bodies of literature. In the following, we review literature on game streaming platforms, streaming and games accessibility, and existing game-aware systems.

2.1 Game Streaming

Live streaming's increasing popularity has led to the emergence of many different platforms of live streaming, such as Instagram Live, YouTube Live, TikTok Live, and Discord [3, 6, 29, 43]. While all of these platforms share similarity in the synchronous interaction afforded to their users, we specifically focus on game streaming on Twitch, which has the largest share of game stream viewership [2]. Twitch allows streamers to broadcast their gameplay to a live audience of viewers who are able to synchronously interact using text chat [33]. It is a free to use platform for both streamers and viewers, with streamers able to use open broadcasting software [1] to capture their gameplay content. As Twitch has grown in popularity, it has become an ecosystem that is home to many genres of streams such as software development or creative streams [12, 14]. Despite this increasing diversity of content, Twitch's most popular content remains game-based content [2]. There are also different types of game streams as well, with eSports being popular on the platform [7].

The traditional game streaming interface often includes a video feed of gameplay, a camera view of the streamer overlaid on the gameplay feed, and a text chat panel for viewers to interact with one another and the streamer playing the game [33]. A key aspect of the Twitch streaming experience is "liveness", in which communities can form around shared experiences, creating a "virtual third place" or "digital hearth" in the process [15, 18]. During a stream, viewers are mutually aware of themselves and the streamer, allowing them to create a feeling of digital co-presence [10]. These communities can also extend their interactions offline to other platforms such as Discord [36].

A substantial body of work explores the differing needs and motivations of viewers on Twitch. Cheung et al's study of spectators of the game *StarCraft 2* defined nine different viewer personas, which includes "socializers" who are interested in interacting with other spectators, and "assistants" who enjoy offering feedback to the streamer to help them progress in the game. [8] Further work has shown that viewers are motivated by social interaction, information-seeking, community building, or entertainment [17, 22, 37]. Viewers' motivation to engage in streams for social interaction and community building are key aspects of the viewer experience, which is not the case with most mass media [18, 37]. Additionally, Twitch streams can often be hours in length, and "drop-in spectatorship", where viewers join a stream already in progress, is a common phenomenon [41]. This drop-in spectatorship means viewers can have various levels of context when joining a stream. Despite having differing viewer motivations, Lessel et al's survey of viewers revealed how viewers rank interactive components of a streaming interface to be their most valued features, which illustrates the need for novel game streaming interfaces that afford interactive features for viewers [27]. Additionally, how these varying motivations apply to BLV viewers remains largely unknown.

2.2 Streaming and Game Accessibility

While game and stream audio are often part of a streaming experience, information about the game's contents is still predominantly presented visually through video. This essential functionality is a barrier for BLV viewers of game streams [24] as well as for BLV streamers themselves [23].

Our work takes inspiration from and builds on Killough and Pavel's work, which engages the particular accessibility challenges posed by game streaming [24]. Overall they found that text, audio, and pre-recorded text methods of description helped increase the accessibility of the stream content for viewers with visual impairments, but they identify that video game description presents new challenges for describers. Describers must prioritize fast-paced and complex content as it unfolds as well as articulate and communicate domain-specific information related to the game. Experienced audio describers may not possess the domain-knowledge required for a particular game stream while community-members with domain knowledge may not possess the expertise and skills of an experienced audio describer. Crowd-sourcing description from community members often resulted in finer details about contents on screen being deprioritized in favor of higher-level, major details.

We recognize the tensions that exist in producing descriptions in game streams, however we instead wish to investigate how existing game data could provide viewers with interactive control over their experience of the game and stream information, which may also help address these issues of lost information due to pacing and prioritization.

There has been a recent focus on accessible game development [26] that we believe provides a rich body of potential metadata that could be provided to live stream viewers interactively. Brown and Anderson's evaluation of the state of accessibility in video games found that modern titles developed with accessibility in mind provide players with customizability and control over font size and family, contrast, colorblind state toggles, UI scaling, screen-reading, and timing for the display and progression of on-screen text [5]. Many of these capabilities fall outside of the typical scope of content that would be prioritized in game streams or live description, since they are functional controls of the player's game experience. However, if game aware interfaces can be made available to stream viewers [20], we see an opportunity to include accessibility metadata and functionalities as well.

2.3 Game-Aware Streaming Systems

While game streaming has evolved into an entertainment medium with millions of users, the limitations of the existing interface have not seen a similar evolution. There has since been a growing interest in designing new interfaces to improve the overall streaming experience. These new systems leverage various streams of data to support viewer-side features.

One approach to evolving the game streaming experience relies on using chat data. While the text chat on Twitch gives viewers and streamers an opportunity to connect in real-time, it can be hard to follow in large-scale streams with hundreds of thousands of viewers [13]. Miller et al's Conversation Chat Circles is a system that uses "chat neighborhoods" to limit the messages a viewer sees, helping them communicate more effectively in a large-scale stream [30]. Pan et al's TwitchViz system uses chat data to render visualizations of key chat behavior and moments of gameplay [32]. Similarly, Chung et al's VisPoll system uses chat data for sensemaking, this time for streamers to help make sense

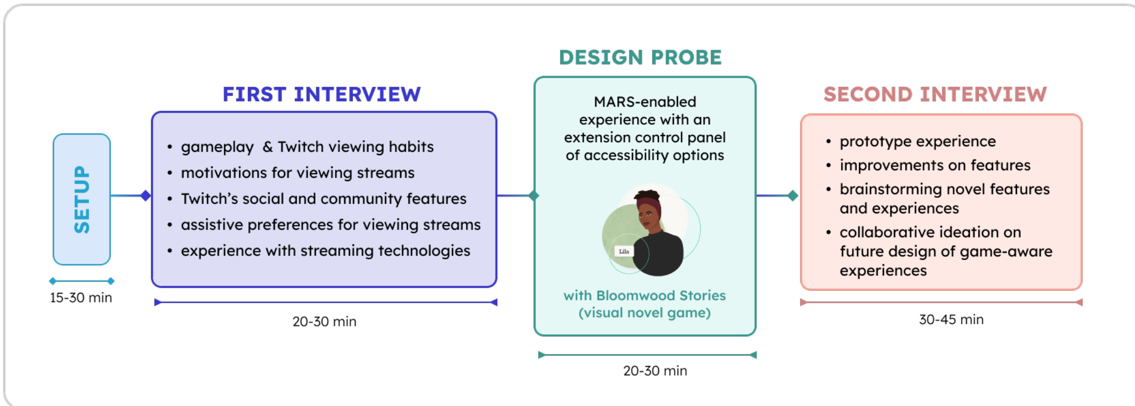


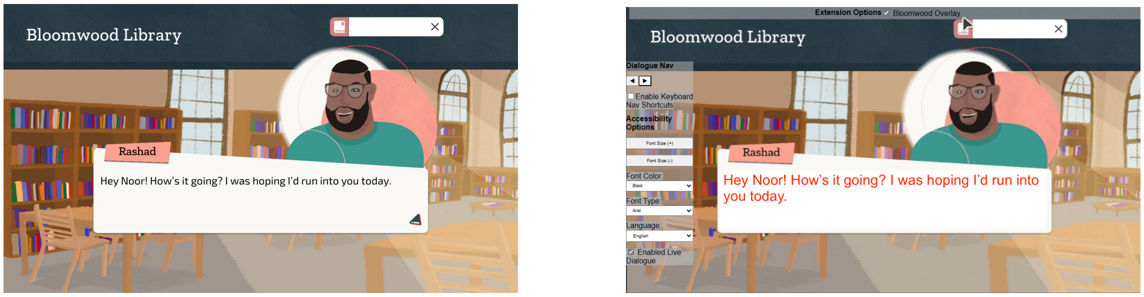
Fig. 1. A summary of each phase of the design probe study

of chat data at scale [9]. Real-time, large-scale chat data can also be used to mitigate toxicity in Twitch streams, which Yen et al's system StoryChat addresses by using viewer messages to create graphical narratives [42].

Beyond managing large-scale viewer chat limitations, game-aware systems can also leverage chat data to enable gameplay interactions for viewers. Audience Participation Games (APGs) allow viewers to directly impact gameplay [15, 35]. Early versions of APGs such as *Twitch Plays Pokemon*, *Choice Chamber*, and Glickman et al's APG toolkit used Twitch's text chat to aggregate viewer input, which is then transmitted to the streamer's game and used to impact game progress [15, 34]. Recent APG system design has sought to move away from the aggregation-based approaches that are commonly used to feed viewer input into the game, particularly Danmaku Participation Games on the BiliBili game streaming platform that enables independent viewer control of gameplay [39].

Beyond chat data, game-aware interfaces have also begun to incorporate broader metadata from games to create novel experiences [20]. These systems scrape data in real-time from the streamers' game, using that data to allow viewers to interact with game objects from the streamed game. Hammad et al formalized the concept of game-aware interfaces with their MARS system (Metadata Augmented Real-time Streaming). MARS is a generalizable architecture that enables the game data from any Unity game to be rendered in real-time onto a Twitch viewer's interface. MARS makes use of the Twitch Extension API, which is a recent addition to the Twitch platform that allows front-end developers to customize a viewer's interface via web development approaches [4]. Additional game-aware systems include *HelpStone*, which uses game log data from the deck building game *Hearthstone* to generate interactive overlays for viewers interested in communicating with the streamer [28]. These game-aware systems enable personalized viewer interfaces and allow viewers to independently interact with the game.

The design and development of novel viewer-side experiences will continue to evolve through the efforts discussed above. However, while these systems' touted capabilities are meant to enable new means of interaction for Twitch viewers, the existing needs of BLV viewers have not been considered in their development to date. Moreover, it is unclear whether new features such as APGs, chat augmentations, or personalized game overlays require new mechanisms for BLV viewers to interface with them.



(a) *Bloomwood Stories* game view that is seen by the streamer

(b) *Bloomwood Stories* extension is enabled, with the viewer adjusting the font type, size, and color.

Fig. 2. *Bloomwood Stories* game view as seen by the streamer, and overlay features enabled in design probe

3 DESIGN PROBE STUDY

To approach understanding the game-aware streaming needs of BLV game stream viewers, we conducted a design probe study. We invited individuals who were interested in game streaming and identified as BLV to take part. Each study session consisted of a set-up phase to give participants access to the live prototype, followed by the three main parts of the session. In the first interview, participants were asked about their current gameplay and viewing habits, along with questions about their assistive technology preferences while viewing game streams. This was followed by the prototype test itself, which served predominantly as a probe to facilitate ideation in the second interview. This second interview focused on the participants' experience with the prototype to generate ideas and improvements towards the accessible game-aware streaming design space. The resulting data was then transcribed and grouped using affinity diagramming by multiple research team members. A summary of our study methodology can be found in Figure 1.

Below, we break down the study details, starting with a detailed description of the game-aware prototype we used.

3.1 The Design Probe

A crucial component of our study design is the use of a live prototype. By allowing the participants to try a game-aware prototype while viewing a stream, we are able to surface shortcomings in existing features and interaction mechanisms and facilitate ideation with participants to inform the design of future game-aware streaming systems.

To create our design probe we made use of the MARS toolkit [20]. The specific game we chose to build a MARS-enabled experience around is an award-winning visual novel game titled *Bloomwood Stories* [11, 38]. In this single-player text-heavy game, players interact with community members in the fictional neighborhood of *Bloomwood* to address their community needs. Progression in the game hinges on dialogue interactions with the non-playable characters (NPCs), with choice-based dialogue selection serving as a core mechanic in gameplay. *Bloomwood Stories* was presented in the original MARS paper as a notional example of how a MARS-enabled extension could support accessibility needs, however, their implementation was superficial and did not engage with actual needs of BLV users.

Bloomwood Stories was developed using the Unity game engine. Using MARS' Unity-side tools, the game was outfitted with MARS' metadata tracking system that sends game data to an extension running on a Twitch viewers' web browser. While MARS allows any game data to be sent to Twitch, our prototype focuses on the game's dialogue scenes. Specifically the system exposes metadata about each dialogue line characters say as well as the response options

Participant	Vision Impairment	Assistive Technology	Twitch Usage
P1	Totally blind	Screen reader	More than three times a week
P2	Totally blind	Screen reader	More than three times a week
P3	Legally blind (low vision in one eye and totally blind in the other)	Screen reader	Once a week
P4	Low vision	Text magnifiers	Once a month
P5	Low vision	Text magnifiers	More than three times a week
P6	Low vision	Text magnifiers/contrast	More than three times a week
P7	Totally blind	Screen reader	More than three times a week
P8	Low vision	Text magnifiers	More than three times a week
P9	Severe low vision	Text magnifiers/contrast	More than three times a week
P10	Low vision	Screen reader	More than three times a week

Table 1. Summary of participants and their visual impairments, assistive technologies, and Twitch usage

available to the player to continue the conversation. We chose to focus on these scenes specifically because this kind of dialogue interaction is common in many games and its text-centric nature presents potential challenges to BLV users.

The MARS-enabled extension for *Bloomwood* displays as an overlay on top of the Twitch video player and is designed to explore three key features: presenting dialogue text in screen reader-compatible format, enabling manipulation of dialogue display properties, and providing tools for viewers to understand the game at their own pace with real-time and historical dialogue. Figure 2 demonstrates the streamer's view of *Bloomwood Stories* side by side with the viewer's extension overlay with the accessibility options panel.

The dialogue data from the game is overlaid on top of where it would be displayed within the streamed video. This dialog is rendered using SVG tags and can be accessed by screen reader users as well. The interactive elements of the extension include ARIA tags to provide additional semantics to help with screen reader navigation.

An extension control panel is displayed to the left of the interface and provides a collection of accessibility options that can be used to modify the dialog text display. This panel gives viewers the ability to select alternate fonts, increase or decrease the font size, and change the font color of the dialogue text in real-time. These display manipulation controls are also responsive to keyboard commands to avoid the need for click based user interactions.

To provide sensemaking support to viewers who may want or need to read at their own pace, the control panel also includes "next" and "previous" arrows that let viewers browse previous dialogue from the game that is now off-screen due to the streamer having moved on in the in-game conversation. Dialogue browsing can also be modified using the "Live Dialogue" toggle, which will automatically update the dialogue to the streamer's current position in the dialogue tree. While these dialogue browsing features were included to help support viewers who wish to read at their own pace, it is also important due to the nature of "drop-in" spectatorship in Twitch, and because Twitch does not allow viewers to rewind a stream to see what they missed.

By design, the feature set of the *Bloomwood Stories* MARS probe is intentionally limited. We do not include all of the game's text, such as the text in the game's menus, as part of the metadata that is sent to Twitch. Furthermore, we did not implement any major features outside of text manipulation and text browsing. The design space of accessible game-aware prototypes requires further exploration, and we do not yet know what features BLV viewers want. As such, we built our prototype to act as a probe that illustrates the potential of game-ware interfaces, thus allowing our participants to start thinking about features they want in future iterations of game-ware interfaces

3.2 Participants

For this study, we targeted two types of participants; individuals who identified as blind and those who identified as low vision. We were particularly interested in users of assistive technologies while viewing game streams, such as users of screen readers, magnification software, and font manipulation tools. Our inclusion criteria specified that participants be United States residents, be aged 18 or older, identify as BLV, have an interest in playing video games, and have an interest in viewing streams on Twitch.

Our recruitment process utilized multiple channels to reach potential participants. We first reached out to games accessibility organization *AbleGamers*, who helped spread our recruitment flier via their Player Panels. We also used social media platforms such as X, and Reddit to reach prospective participants. Specialized online communities, including discord servers for games accessibility practitioners, and members of games accessibility and live stream communities. We also posted our flier on the AudioGames forum, which is dedicated to using audio, as opposed to graphics, as the main output from video games. Prospective participants first filled out an online screener survey that is linked in our flier, and selected participants were emailed consent and scheduling information.

In total, we recruited 10 participants (table 1), five of which were predominantly users of screen readers, and five who were using other assistive technologies such as magnification software and text manipulation tools. While all of our participants were avid gamers that had an interest in viewing streams on Twitch, four of our participants were active streamers or had been streaming on the platform in the past.

3.3 Session Set-Up

Before commencing the interviews and prototype test, the research team used the start of each study session as an opportunity to set-up the prototype test. Prototype set-up was done at the start of the study session in case technical issues emerged, which allowed the team to identify issues early and adapt accordingly. All study sessions were conducted remotely via Zoom, and each session was recorded for later transcription. For each session, a minimum of four research team members were present on Zoom at all times; one team member served as the interviewer for all sessions, one team member served as the streamer playing *Bloomwood Stories* for all sessions, and two team members supported the sessions by taking notes and by acting as confederate viewers who interacted via the Twitch chat during the prototype test.

Using Zoom allowed participants to share their screens and provide observational data to the research team. Zoom also allowed the research team to help participants set up their personal device for the prototype test. The research team helped participants log into the provided account, navigate to the Twitch channel that *Bloomwood Stories* was being streamed to, and enabled the extension on the participants' video player. Because the prototype used in the study is a development build that had not passed the Twitch Extension publication process, participants needed to use a Twitch account created by the research team that had permission to use our pre-release version of an extension. Unfortunately, the button used to enable an extension for viewers on the Twitch platform is not accessible by screen reader. To circumvent this issue, the research team used Zoom's Remote Control feature to temporarily take control of the participants' device to enable the extension.

Depending on the participant's personal device, assistive technology settings, and browser settings, session set-up lasted approximately 15-30 minutes. After ensuring the prototype was working for the participant and that the streamer's audio and video were being streamed properly, the first interview, prototype test, and second interview were able to commence.

3.4 First Interview

The goal of the first interview in the study session was to develop an understanding of the participant's current Twitch viewing experience. Twitch is a multi-faceted platform where gameplay and social interaction are central to the viewing experience, yet viewers have different approaches and motivations for engaging with the platform [18]. As such, while the interview was semi-structured in nature, we grouped the interview around the participant's gameplay habits, Twitch viewing habits, motivations for viewing streams, and outlook on Twitch's social and community features. This was followed by their assistive preferences while viewing streams, and their experiences with streaming technologies such as Twitch Extensions or *Twitch Plays Pokemon*. Prior to the study session, we reviewed the participant's screener survey responses to help tailor the interview questions to their accessibility needs and Twitch experience. The first interview lasted 30-45 minutes, and was followed directly by the prototype test.

3.5 Prototype Test

Upon concluding the first interview, the next step in the study session involved the participant getting hands-on experience with the *Bloomwood Stories* MARS-enabled extension prototype. The research team member serving as the streamer launched the game and played the game while streaming on Twitch. The streamer played the game, providing commentary on the game while interacting with chat, however they were deliberately instructed not to read out in-game dialogue word for word in order to allow room to test the affordances of the prototype extension. Two other research team members joined the stream as viewers to facilitate a more active text chat for the participant. These research team members were not given a script, but freely interacted with the streamer and the participant using the chat as is common in Twitch streams. Participants joined the stream of the game while having the prototype enabled.

Prior to interacting with the prototype, participants were given an overview of the game and were instructed to use either their screen reader or the extension's accessibility options to manipulate the game's text. Though the prototype did not include any features that impacted TWitch chat, participants were told that they were free to interact with the chat as much or as little as they wanted.

Participants were then free to interact with the prototype while viewing the game stream. They were able to interact with the streamer and other viewers using the chat, and were given the opportunity to openly explore the prototype features while viewing. The lead interviewer in the study remained on the Zoom call to observe the participant's shared screen, taking notes on their interaction with the prototype to revisit during the second interview. Should the participant have any new technical difficulties while viewing the stream, they were able to directly interact with the lead interviewer to resolve the issues. The prototype test lasted 20-30 minutes, depending on how long the streamer took to play the first section of the game.

3.6 Second Interview

The second interview that followed the prototype test was grounded in the participant's experience with the prototype. The questions in this interview focused on the participant's prototype experience, improvements on the features in the prototype, and on brainstorming novel features and experiences they wanted to see beyond the *Bloomwood* game-aware experience. The observational data collected in the prototype test was used to guide the prototype discussion. Discussions about future design of game-aware experiences took the form of a collaborative ideation with the interviewer. Similar to the first interview, the second interview was 30-45 minutes long, and each session was approximately two and a half hours in length. Upon concluding their participation in the study, participants were compensated with a 50 dollar gift card.

To begin analyzing the data from the two interviews, three members of the research team undertook the task of transcribing the interviews. This was followed by an inductive coding approach by two research team members. Both independently coded the same two interviews, to help establish common understanding across an initial set of codes and determine guidelines for developing a codebook. After reaching a consensus on the robust quality of the codes from two interviews, each researcher then independently coded four sessions. While coding independently, the two members collaborated closely by iteratively refining them as they proceeded with their analysis. Throughout the coding process, the team members meticulously familiarized themselves with the interview transcripts and memos taken during the sessions, reading through the material to ensure a comprehensive understanding of the data. This initial coding pass generated 105 codes, which were thematically categorized using affinity diagramming. At this stage, two additional research team members joined the analysis process to collectively discuss the generated codes, working collaboratively to reflect on interpretations of the codes that materialized into 15 higher-level categories. This iterative process involved thorough deliberation and discussion until consensus was reached about the emergent categories

4 RESULTS

The collected data from the study includes the semi-structured interviews as well as the observational notes taken during the prototypes. We classified the data into three broad categories: Current challenges and practices, existing features, and envisioning an ideal future for game-aware streaming. Below, we elaborate on the findings in each category.

4.1 Current Challenges and Practices

Since the first interview in the study was designed to understand participants' prior experience engaging with game streams, a category of codes emerged to describe the various aspects of how participants currently interact with Twitch. Below we elaborate on our findings, some of which corroborate prior work's findings with regards to the current barriers and practices [24], while others reveal new insights into participant motivations, views on communities on Twitch, and use of assistive technology.

4.1.1 Game Streaming Background. All of our participants were avid gamers who played games on a weekly or daily basis. Participants played a variety of game genres, ranging from deck building games such as *Hearthstone* (P1, P2, P3, P7), role-playing games such as *Final Fantasy* (P6, P9), and farming simulation games such as *Stardew Valley* (P4, P5, P10).

In terms of game stream viewing, participants ranged from daily viewers (P1, P2, P7, P8, P9), weekly or monthly viewers (P3, P5, P6, P10) to occasional viewers who generally preferred recorded content on YouTube (P4). Types of game streams viewed by participants include Let's Plays (P2,P6, P9), speed runs (P5, P6), lore breakdown streams (P4), and indie game streams (P9). While participants had differing preferences when it came to viewing on the Twitch mobile app (P1, P2, P6) or the browser version (P3, P5, P7, P8, P9, P10), they valued both platform versions to support multi-tasking behaviors.

When asked about familiarity with game-aware streaming features such as Twitch Extensions, participants had either never heard of Twitch Extensions (P2, P3, P5) or had some experience but had encountered accessibility barriers (P1, P9, P10). When asked about APG systems, most had heard of chat-based systems such as *Twitch Plays Pokemon* (P2, P3, P4, P9, P10). While participants were curious about interacting with APGs mechanics, the use of text commands discouraged them from actively contributing to APG gameplay (P5, P9), with one participant referring to APGs as being "too big (in terms of chat)".

4.1.2 Current Assistive Technology Use. In our study, all participants relied on some type of established tool like screen readers (NVDA, VoiceOver), screen magnifiers or contrast settings. Others had created workarounds such as keyboard shortcuts, customizing mouse pointers, or creating custom monitor display stations based on type of task, gameplay and stream (P5, P6, P7, P9). A third strategy was to use multiple platforms simultaneously, such as viewing the stream on the computer while magnifying the chat comments on a smartphone (P4, P5), or listening to stream audio on one device while using a screen reader on another (P2, P3, P7). However, participants reported that managing these personalizations can be extremely taxing (P3, P4, P7).

Another approach participants used was strategic platform selection. Participants reported that the Twitch interface was unusually noisy (P2, P6, P9, P10) and had a complex UI that made it difficult to navigate. YouTube is more accessible and has a simpler interface (P1, P2, P3, P5, P6, P7), but focuses on pre-recorded video. Participants reported waiting for recording of streams to appear on YouTube, rather than participating live on Twitch, as a way of securing access.

For some participants, the stream itself served as an assistive technology. Watching a stream allowed them to *vicariously play* a game that would not otherwise be accessible to them (P2, P5, P7,P8). "Oftentimes I get to experience games that I can't play if they (streamers) are doing a Let's Play. I still know the story of a lot of games I can't actually play on my own. Because they (streamers) are just reading the dialogue that I can't see. And doing the difficult gameplay that I could never do myself." (P2). Streamers' audio also served to increase access, particularly for streamers who focus on describing the game (P2, P3, P5, P7, P10). This behavior is typically intended to support engagement for sighted viewers, but the live audio narration serves as "an accidental benefit that makes their stream more accessible" (P2). However, participants reported that streamers would sometimes stop describing their actions during fast-based moments (P10), and that it could sometimes be frustrating to view a game they could not themselves play (P3, P4).

4.1.3 Stream-Level Access Barriers. Participants identified streamer's audio description as an access feature that could fluctuate between streams. Similarly, they identified access barriers that might be more or less problematic across different streams.

First, the streamer's audio balance "really makes or breaks an experience" (P2). BLV viewers often have to calibrate across sound quality from microphone and speaker, streamer's volume and screen reader, chat interruptions, and music accompanying the game (P2, P3, P6, P7). Tradeoffs negotiating distinctive audio sources and levels become even more difficult when attempting to balance while multitasking (P3) or if the streamer chooses to adjust audio during the game, changing "settings between music, sound effects, and speaking, things like that (P7).

Second, some games presented much higher barriers to accessibility than others, even beyond the availability of accessibility mods (P2, P3). Participants specifically reported challenges with fast-paced games such as fighting games, where even a well-intentioned streamer cannot keep up with describing the action on the screen (P10), as well as with open world games and others that require freeform exploration (P3, P9). One participant compared this challenge in gameplay viewing to consuming social media, where the video-first and image-rich content with no informative audio required detailed added verbal explanation to truly understand its significance (P7). Interestingly, this led several participants to seek out first-person shooter games (P3, P9). Although they are fast-paced, they rely on a small set of in-game actions (e.g. move, shoot, reload), most of which can be tracked through the default audio cues. Additionally, the fast-paced moments typically require little explanation to understand why they are significant, and can be reduced to a short description such as "hit," "miss," or "headshot."

Finally, the stream chat experience was not only inaccessible itself, but also as made the rest of the stream less accessible by burdening them “with another source of sound” (P2). While participants shared they would like to interact with chat more, its current state of navigation, audio feedback, translation of elements like graphical components (GIFs and memes), chat badges, icons, emojis, or even language characters convolute screen reader interaction (P2, P3, P4, P5, P7). This was particularly problematic in larger streams, with more than a hundred or so viewers, as well as in streams where the cultural norms emphasized the use of symbols, GIFs, and emojis [13]. Given these interruptions, many participants chose to migrate their conversation to separate forums or channels like Discord, where they can chat in a more controlled environment.

4.2 Existing Features

A second category of codes referenced features that are already implemented in the probe: screen reader access to in-game dialogue, text manipulation, and dialogue browsing functionality. As a design probe, our goal was not to provide a “finished” experience, but rather to use these features to elicit participants’ challenges and suggestions. In this section, we report what we learned.

4.2.1 First Impressions of The Experience. Overall, participants appreciated the features offered by the probe, particularly real-time access to dialogue using a screen reader without relying on the streamer to describe the scene (P1,P3,P5, P10), on-demand information to understand game narrative (P6), and font size manipulation (P4, P9). These personalized aspects offered by the extension created a more personable streaming experience for them. "I felt a lot more attached to it (the game). I mean, obviously, this is going to be different because it’s a much smaller stream, but that’s what I like. Even just beyond the chat, having that extra attachment to the game... It was exactly what I needed" (P3).

While all participants expressed excitement about the potential of game-aware streaming technology, most experienced a challenging learning curve. "I’d say at first (understanding the extension) was a little difficult, just kind of understanding the way that the extension was laid out. But I think overall it did enhance my experience. And once I started to understand it more, I was able to go through the story with (the streamer), and I thought it was a pretty great experience" (P1). In particular, screen reader participants needed extra support to distinguish between our probe and alt-text (P3), an alternative chat display (P2), or a participatory interface (P1, P9). Participants suggested both social and technical responses, including creating an “Extension Help” button that viewers can click on at any time (P10), having the streamer describe what the system does (P2), or adding bots in chat to help support the system (P2). Participants also were concerned about how the system might be discoverable outside the research context, and suggested approaches such as adding an #accessibility tag to the stream title (P9).

4.2.2 Text Manipulation Features. The ability to manipulate the look of in-game dialogue was only relevant to visually impaired participants, who valued this ability (P4, P5, P6, P8, P9). Their feedback focused on how to iterate this feature, such as by adding more fonts (P4), introducing a color picker (P9), or allowing users to change the background color against which text appears (P5, P9). Participants also wanted more fine-grained control over font size. Some viewers wanted magnification support, which would temporarily increase the text size on hover (P4, P5, P8). Others wanted to replace the font size buttons with a slider, which would both provide finer control and prevent Twitch’s video player from inadvertently entering or exiting full screen mode (P4, P6).

While our system only supported customization of in-game dialogue, visually impaired participants wanted their styling changes to apply to all other in-game text (P9), as well as to the interface itself (P4, P5, P6). We see this as an indication of the feature’s value. Additionally, participants suggested extending the visual customization to chat, and including a filtering feature to remove emojis and badges from chat usernames (P4, P9). This chat filtering feature was also desired by screen reader participants to help them make sense of many chat messages at once (P2, P3, P7).

One concern with offering customization options was that they can be overwhelming (P6). To overcome this concern, participants suggested that future systems could provide presets of settings. These presets could be shared between viewers via a code (P6), added to the interface as macros (P9), or be directly modified by users, for example by remembering when a user hides an interface element (P6, P8).

4.2.3 Improving System Navigation. Most participants identified ways to improve the navigation of our probe. Screen reader participants suggested we could add more headings or ARIA landmarks (P1. P2, P3, P10), implement more shortcuts to reach different features (P3, P7), or collapse the accessibility panel entirely (P1, P3). While these changes can be made relatively easily, participants also identified problems that are rooted in Twitch's underlying technical architecture.

From a development perspective, we were not able to control where Twitch embeds the probe into the DOM, and as a result this made navigating to and from the probe and text chat difficult for screen reader users. "If I want to chat, for me to go back and forth, I have to hit the headings each time to go back. And there's not really a shortcut for that, so it's a little bit hard to do that on the fly. It does take a few seconds (to navigate)" (P1). To mitigate this issue, participants suggested shortcuts dedicated specifically to task switching (P1, P2, P3), or having a "focus mode" that switches to live reading of incoming dialogue if the viewer is in chat (P1, P7, P10). Similarly, Twitch Extensions that overlay video content can only appear within the bounds of the video. For low-vision participants, this meant that the panel would sometimes obscure their view of the dialogue. This led to suggestions for a button to collapse and open the accessibility panel as needed (P5, P6, P9).

4.2.4 Establishing Context. Our probe allowed users to consume live in-game dialogue, or to browse it at their own pace. However, our probe's focus on dialogue illuminated areas where description of *game context* was also needed, in order for the dialogue to make sense. First, participants identified gaps in *dialogue-level context*, such as giving screen readers access to the name of the character that was speaking (P2, P3, P7). *Game-mechanical context* included requests for clearer identification of the options available to the streamer (P1, P3, P10) and of what choices the streamer had previously selected (P10). "I think it would be cool if it would auto speak what option the streamer clicks. It could say 'help this person find their jobs' and then a second later, it just automatically reads me the new dialogue" (P10). *Game state context* was needed for when the streamer switched to the in-game map, as there is no dialogue available in that part of the game. This resulted in confusion among participants who were left wondering why the dialogue disappeared (P2, P3, P7, P9, P10). Further attention is needed to provide enough context when transitioning from one game state to another, such as announcing text that says "in the city map now" when the streamer navigates to the map (P2). Finally, *stream-level context* was important for addressing the drop-in nature of streaming (P2, P5, P6, P9). Suggested solutions revolved around adding stream summaries, such as by adding a panel outside of the video that updated at regular intervals with what has happened thus far (P2, P5) or by allowing viewers to view past data to "rewind the action they missed" (P6).

4.3 Envisioning an Ideal Future for Game-Aware Streaming

A series of codes emerged that described features or ideas that were not directly addressed in the prototype participants experienced, but rather discussed ideas that could be implemented in a broader ecosystem around game-aware streaming.

These codes thus account for desired features beyond the probe's capabilities and insights into how participants envision a desired state of the world.

4.3.1 Reducing Overload. As described in Section 4.1.3, game audio was a major accessibility issue. Depending on the stream, the game's audio may compete with the streamer's microphone or background music, which can interfere with game sensemaking efforts (P1, P2, P3). Alternatively, participants also expressed that there are different types of game audio (e.g. sound effects, character audio), and that the importance of the game's audio depends on the type of game being played (P6, P7); for a game like *Super Mario Maker*, where the "jump" sound effect is prevalent in gameplay, the streamer's audio should take precedence over the game's sound effects (P6).

While our probe did not include audio manipulation features, participants suggested a range of approaches to address audio overload. This included changing the streamer's volume separately from the game's volume (P2, P6, P7), lowering the game music's volume compared to spoken dialogue (P3, P4), or adding a speech recognition system which dynamically adjusts the various sources of game sound (P9).

Participants also suggested visual strategies for reducing overload. For example, some games render visuals at a faster rate than *Bloomwood Stories*, such as first-person shooters or open world games. Participants felt it would be useful to slow down the gameplay to make sense of visual information (P5). Descriptions of visual data could be dynamically provided if a streamer is silent during a game scene (P1), or in an "alt text" format that labels scenes for viewers to access as needed (P3, P7).

4.3.2 Sensemaking with Game Data. The probe participants tested was built around a visual novel game, which is a 2D game genre that uses dialogue choice as a key gameplay mechanic. Other genres of games differ in their types and complexity of mechanics, and participants discussed how game-aware streaming features might apply to these other genres. Participants were interested in how different forms of data could be used beyond in-game dialogue, such as using lighting data or graphical data (P2, P10). Participants valued the ability to access real-time game information on-demand (P1, P2, P5, P6, P7, P10), and wanted to extend this to other games. This could be used to gain deeper insight into a game's lore (P3), access the statistics of enemies (P5, P7), get insight into strategies for competitive games (P5, P6), or to learn more about new games (P10). Similarly, participants wished to see the dialogue browsing from the probe generalized to other forms of "off-screen interaction" to allow them to explore game environments or different choice outcomes that the streamer did not choose in role-playing games (P5, P6, P10).

Moving beyond textual representations of data, there was also an interest in leveraging sound cues in the game for sensemaking; existing games have sound cues that can indicate when a player navigates to the menu screen, unsheathes a weapon to attack an enemy, or dodges an enemy attack in a game like *Mortal Kombat* (P3). Prior work has demonstrated that these sound cues are used by BLV gamers to navigate the game, and participants expressed using these sound cues in their own gameplay [16] (P3, P4, P7). Allowing viewers to manipulate the volume of these existing sound cues, however participants suggested that future interfaces can create "3D spatial audio" sound cues that are not present in the game to help viewers interact with a game's environment more effectively (P7, P10).

Finally, many participants wanted access to the *aesthetic* and *social* experiences of play (P1,P2, P3, P7). For example, one participant wanted descriptions to capture the richness of the world in games like *World of Warcraft* or *Skyrim* (P3, P10). Participants also were interested in access to visual "easter eggs," which are in-jokes that reference shared knowledge among the gamer community (P3). Participants suggested that these visual descriptions could be generated using the game's actual metadata (P3), AI-driven techniques that automate descriptions (P1, P3, P7) or community-driven approaches (P1, P3, P10). Descriptions could be entirely generated by the community, as is the case in Killough et al's

[24] generation of audio descriptions (P1, P3), or through the use of a "sound database" that game developers provide which the community can edit and improve (P10).

4.3.3 Participation and Collaboration. A heavily requested feature type from participants was APG features for viewers to impact game mechanics (P1, P2, P3, P5, P6, P10). One such example was for the horror game *Phasmophobia*, where players are tasked with identifying ghosts at a haunted site [25] (P2). A game-aware interface for *Phasmophobia* could leverage asymmetric game design and give viewers access to the "compendium" containing ghost-hunting details that the streamer needs [21]. This format of asymmetric play requires collaboration between the streamer and the viewers to beat the game (P2, P6); however, participants also expressed an interest in enabling antagonistic APG features which allow viewers to hinder the streamer's gameplay (P10). While the initial design space for APGs has already been identified in previous work by Seering et al, the question of how to best design APG mechanics and interfaces for BLV viewers remains an open question that requires further research [35].

While this is not an accessibility feature, an additional collaborative mechanic that was mentioned was region tagging in the game space which allows a viewer to highlight a part of the game space and have that region annotated for the streamer and other viewers (P5).

4.3.4 Alternative Modalities and Outputs. Our probe was browser-based, but we saw participants ideate different interaction modalities and output mechanisms that would be useful for them in the future. A popular suggestion by participants was to create a conversational interface that allows viewers to input queries about the game's status. This included querying about the streamer's current location in the game (P2), using *Dungeons and Dragons* dungeon master-style narrations for context setting (P1), providing rich visual descriptions of the scenery in a game (P1, P2), and querying about the status of NPCs in games like *Stardew Valley* which center around interactions with NPCs (P10). One participant suggested tagging locations in the game such that viewers are able to input the tag as a query and the system will respond to the nearest objects in the game that possess that tag (P10).

While these participants were inspired by the terminal-style interface of LLM systems such as ChatGPT to execute their queries [31], another participant suggested utilizing voice commands instead to input queries (P3). Screen reader participants also brainstormed solutions to prevent the over-reliance on audio output, specifically through haptic-based controller interfaces (P3) and braille readers (P1, P3, P7). In particular, participants noted that any braille reader solution should output the game-aware interface content and not the chat, since chat can be too fast and chaotic (P3, P7). "NVDA readings try to read stream chat. And (I'm) trying to keep up with dialogue and listen at the same time. Especially in a huge stream that has 'machine gun' chat... So this can be difficult where Braille might be a good option (for the extension)." (P7). While alternative output formats such as haptics and braille readers are promising ideas, a design challenge such an interface would need to manage is how to avoid viewer confusion and cognitive overload due to multitasking (P1).

5 DISCUSSION

Our findings on our participants' current practices extend previous work in two ways. First, we identified a motivation for stream viewing that is not currently present in the literature: *vicarious play*, or access to a game through streaming that is not otherwise accessible. This finding is a contribution to the larger body of work on streaming motivation, but it also showcases the value of working closely with specific communities of stakeholders such as BLV participants.

Second, we extend previous work on audio description for streaming. Killough and Pavel [24] suggest that describers must prioritize fast-paced and complex content. Our participants' relationship to FPS games suggests another alternative:

that in some games, fast-paced content can be meaningfully “chunked” without losing the overall narrative of the stream. It is much less important whether a streamer’s targeting reticle sways side to side or up and down, than whether they hit or missed their target. Guidelines for effective audio description - and the ease of learning to describe the game effectively - can therefore vary based on the genre of the game.

Finally, we note that a common theme in our interviews was a strong desire to engage with game streaming, despite all the barriers that participants describe. This is important because many barriers to accessibility are rooted in ableist assumptions about the BLV community and their relationship to games and streaming (P1, P4, P7, P9). Given that most video games and streaming experiences are visually focused, participants reported encountering misconceptions that people with visual impairments simply do not want to play (P1, P7). Our data clearly shows this is wrong. BLV gamers go to great lengths to participate in their hobby, and even more so to participate in game streams. We therefore call the attention of sighted gamers to correct their misconceptions.

Participants’ reactions to our probe showed a need for personalization not just for the in-game dialogue, but for the probe’s interface itself. For example, low-vision users needed their font changes applied to the interface, while blind users did not need to access the font customization system at all.

Unfortunately, many of the challenges identified by participants were caused by platform-level features of Twitch. The MARS system already provides workarounds to a number of technical challenges related to the Twitch platform [20], but issues such as the inability to control where code is inserted into the DOM is unlikely to be resolved without institutional cooperation from the platform itself.

Pressuring Twitch to make platform-level changes will likely require concerted effort from the BLV community and its allies. There is already a community of blind developers, gamers, and viewers who actively build relationships with other stakeholders in the streaming world, such as sighted game developers and content creators. For example, participants pressed for better game developer education about the needs of the BLV community, and sought to engage sighted developers directly with BLV gamers (P2, P3, P5, P6, P8, P9). Corresponding with content creators who hold deeper connections with the streaming platform could be an alternate form of influence (P3). Building solidarity between sighted and BLV streaming stakeholders is therefore a key part of addressing these apparently technical challenges.

We believe that this ties to the issue of discoverability. Right now, accessible game streams are the exception rather than the norm. While BLV streamers are in control of their own accessibility features, BLV viewers and other stakeholders must find streams that meet their needs, often without much guidance. How might we make accessible streaming the norm? A game-aware approach provides one possible direction, by exposing metadata that can be used for many different purposes. For example, in design-based focus groups, streaming stakeholders identified ten ways that real-time metadata could improve the streaming experience, unrelated to accessibility [19]. This approach may therefore support building the coalitions needed to make accessible streaming the norm.

Our probe, as a single-player experience, did not directly target the community aspects of streams. However, we found that the theme of community engagement provided a lens through which to understand our participants’ vision of the future.

First, we found that there was a tension between wanting to stay synchronized with the live game stream, and wanting the freedom to make sense of the game on their own time. For example, one participant expressed that the idea of “off-screen interaction” (for example, exploring the game environment independently of the streamer) would pull them away from engagement with the streamer (P2). Similarly, the idea of slowing down gameplay during fast-paced games (P5) was in tension with experiencing surprising moments at the same time as the rest of the stream.

Second, we saw that participants wanted increased opportunities to participate in the game stream, for example as evidenced by their desire for access to APG mechanics. Interestingly, the mechanics they envision primarily operate through influencing the streamer's gameplay, in contrast to social mechanisms of influence that rely on relating to other viewers [35]. This may be because Twitch chat is so inaccessible that they cannot yet imagine a future in which they have the full ability to participate; future probes could fruitfully explore this direction.

Finally, participants conceptualized themselves both as beneficiaries of improved accessibility systems, and as contributors to them. For example, participants felt they could improve audio descriptions, particularly with AI support. This desire may be partly motivated by wanting a better experience for themselves and other members of the BLV community. Furthermore, participating in community-level efforts may allow them to experience a sense of being peers with sighted viewers who may have more access to these features. This equalizing power amongst gamers can strengthen the bonds within the distributed digital hearth that makes up a streaming community [15].

Beyond extensions, accessibility features, and assistive preferences, participants expressed various perspectives on other factors that greatly influence accessible live streaming and gaming experiences. Additionally participants elaborated upon the misinterpretation of accessibility automatically meaning narrow assistive technology features that only benefit a small group.

In the past, building accessible games has taken a more utilitarian perspective – as a translation incorporating functionally assistive aspects that strive for more productive outcomes of a game experience. However, our participants distinctly commented how the nature of game streaming has made it an essential part of popular culture, a social phenomena, and a way of participating in everyday society that ought to be accessible for people with disabilities. Gaming is now an experiential form of engagement that is part of digital spaces and these moments within streaming are not parasocial but rather have universal reach. The benefits of access make for more equitable online communities for all. Motivated by this, participants recognized the potential for game-aware extensions to go beyond simple translations of content between modalities to make information available toward an approach that could serve to broaden access to gaming more generally. Participants reiterated how dynamic text data provides a baseline that opens doors to selective parsing, filtering, and manipulation that can form the basis of richer interactions beyond mere text (P7). For example, extensions could also prove valuable to non-English and multilingual speakers (P8) and provide opportunities to modify game data through language translation. Additionally, these tools could be used to help bridge the access gap, from streamers not utilizing screen readers towards data for viewers wanting audio feedback fostering further social connections.

While there have been efforts towards making gaming more inclusive, there remains a long way to go in involving more people with disabilities. In addition to playtesting and getting feedback from disabled gamers, participants advocated that it's crucial to collaborate early in the conception and creation process of games to spearhead avenues of access into major genres that continue to be inaccessible to BLV people, such as first person shooter games or multi-player online battlefield arena games (P1). As accessibility of games largely depends on those who make them, participants noted how there remain cultural barriers across gaming that pose limitations and boundaries for accessibility-oriented change, which create ripple effects for inaccessibility across game streaming. For example, game studios and developers of games with substantial video game market share in certain regions, hold more power in dictating these practices and trends compared to others. They determine a bottleneck of who is allowed to be involved in such enterprises and what types of player experience they are designing for (P1, P7). Many times this is motivated by the belief that supporting disabled gaming is not a profitable undertaking, whereas people with disabilities are still willing to spend on such endeavors.

As participants envisioned ideal futures for game streaming experiences, some reflected on the boundaries of creativity and its long association with visual mediums that has inherently excluded BLV creators. Understanding there are different kinds of play and variances in what a game experience is will generate more opportunities and acceptance for alternate pathways of creating (P4). Other participants envisioned an audio focused social media platform, allowing them to interact with media via audio comments. Parallel to the social and interactive chat component of Twitch, mediums that support listening to audio-based on comments and responding in the same manner can create greater non-visual access to video-driven content like game streaming.

6 LIMITATIONS & FUTURE WORK

While our design probe study explored a multitude of technical dimensions and desired features for BLV viewers, to participate in more accessible live streaming experiences, a number of limitations arose that could be explored further in future work.

Our focus in this project was on the accessibility needs specific to BLV viewers. The potential exists to use game-aware streaming techniques to explore ways to support additional accessibility needs within game streaming (e.g., providing more contract controls, toggling a visual dot to help with motion sickness, drawing a static overlay on top of fast moving content, etc.). Exploring how game-aware interfaces can help address a broader set of needs would be a fruitful avenue for future work.

Additionally, the probe we used in our study was designed for a 2D visual novel game whose primary mechanic was dialogue-driven. Given that the primary metadata was character dialogue, the game-aware interactions for viewing, browsing, and manipulating dialogue was relatively straightforward to design. However, the design space of games can offer much more complex interactions and thus more complex forms of metadata that can drive more complex game-aware interactions. Participants noted that the current approach for text interaction may not work for more fast-paced games (P3), where dialog or other data may appear dynamically around the scene and pop in and out of existence. How to design accessible supports for real-time sensemaking in these more complex environments remains an open question for future work.

Our probe implementation faced a number of technical limitations. For example, the MARS system is currently limited to games built within the Unity game engine and the Twitch streaming platform. Additionally, overlay Twitch extensions are currently only possible on desktop browsers. A number of our participants mentioned how they often access Twitch on mobile devices or on a combination of desktop and mobile as the split in platforms provides better affordances for dealing with the volume of chat interactions. Along with these platform limitations, the prototype used in this study did not offer any chat-specific features, and though we were not able to simulate a high volume chat in our streaming, managine chat remains a major concern for BLV viewers. Finally, our study protocol required us to make use of Zoom's remote control feature to initially enable the overlay extension because of the inaccessibility of Twitch's built in extension controls. Each of these technical challenges is surmountable in principle but they would require coordination across technical components of Twitch and other live streaming ecosystems to provide more reliable accessibility tools to users.

Our work focused on game streaming accessibility, however future work must explore developing viewer-side support tools in non-game streaming contexts as well. While the challenges in non-gaming streaming contexts overlap with game streaming in some ways, such as the inaccessibility of the streamed activity to BLV viewers, they also present new challenges with regards to how data can be collected from the stream to render information to viewers. Thus, while we cannot overgeneralize our findings beyond game streams, we treat our work here as a call to action to create accessible streaming experiences for the broader streaming ecosystem.

7 CONCLUSION

Game live streams present rich and multifaceted experiences to their viewers who engage with them for a number of unique and varied reasons, from entertainment to learning and social connection. Despite this, the viewing experience of BLV viewers is hindered by the limitations of the real-time and video-centric nature of common live streaming platforms. While the highly dynamic nature of video games contributes to this challenge it also presents an opportunity in providing access to the rich metadata created by games as they are played to drive viewer interfaces that are aware of the games' content and expose affordances to BLV viewers to understand and engage with streams.

In this work we developed a design probe of a game-aware streaming interface to explore the potentials for the use of live game metadata to support the needs and desires of BLV viewers of live streams. Over the course of ten interactive sessions with BLV viewers we surfaced a number of themes related to existing challenges engaging with live streams, impressions of potential design concepts, and desires for an ideal future of interactive live streaming. Through the process we have demonstrated the potential of a game-aware approach to allow us to move beyond an approach that merely translates continent from one medium to another toward a design space that leverages the unique affordances of games and streaming to open up a new frontier of accessible interactive designs.

REFERENCES

[1] 2021. Open Broadcast Software. https://obsproject.com/

[2] 2021. TwitchTracker Statistics and Charts. https://twitchtracker.com/statistics

[3] 2024. Discord. https://discord.com/. Accessed: 2024-4-25.

[4] 2024. Extensions Reference. https://dev.twitch.tv/docs/extensions/reference/. Accessed: 2024-4-25.

[5] Mark Brown and Sky LaRell Anderson. 2021. Designing for Disability: Evaluating the State of Accessibility Design in Video Games. *Games and Culture* 16, 6 (2021), 702–718. https://doi.org/10.1177/1555412020971500 arXiv:https://doi.org/10.1177/1555412020971500

[6] Bytedance. 2024. TikTok. https://www.tiktok.com/en/. Accessed: 2024-4-25.

[7] Sven Charleer, Kathrin Gerling, Francisco Gutiérrez, Hans Cauwenbergh, Bram Luycx, and Katrien Verbert. 2018. Real-Time Dashboards to Support eSports Spectating. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 59–71.

[8] Gifford Cheung and Jeff Huang. 2011. Starcraft from the stands: understanding the game spectator. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 763–772.

[9] John Joon Young Chung, Hijung Valentina Shin, Haijun Xia, Li-Yi Wei, and Rubaiat Habib Kazi. 2021. Beyond Show of Hands: Engaging Viewers via Expressive and Scalable Visual Communication in Live Streaming. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21, Article 109)*. Association for Computing Machinery, New York, NY, USA, 1–14.

[10] Vaibhav Diwanji, Abigail Reed, Arienne Ferchaud, Jonmichael Seibert, Victoria Weinbrecht, and Nicholas Sellers. 2020. Don't just watch, join in: Exploring information behavior and copresence on Twitch. *Comput. Human Behav.* 105 (April 2020), 106221.

[11] Morgan C Evans, Adela Kapuscinska, Maya Greenholt, Junchao Lin, Xuanyuan Liu, Tianyi Zhang, Jessica Hammer, and Geoff Kaufman. 2021. Designing a Self-Efficacy Game for Health Literacy in Marginalized Communities. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Yokohama</city>, <country>Japan</country>, </conf-loc>) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 406, 6 pages. https://doi.org/10.1145/3411763.3451609

[12] Travis Faas, Lynn Dombrowski, Alyson Young, and Andrew D Miller. 2018. Watch me code: Programming mentorship communities on twitch. tv. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–18.

[13] Colin Ford, Dan Gardner, Leah Elaine Horgan, Calvin Liu, A m Tsaasan, Bonnie Nardi, and Jordan Rickman. 2017. Chat Speed OP PogChamp: Practices of Coherence in Massive Twitch Chat. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 858–871.

[14] C Ailie Fraser, Joy O Kim, Alison Thornsberry, Scott Klemmer, and Mira Dontcheva. 2019. Sharing the studio: How creative livestreaming can inspire, educate, and engage. In *Proceedings of the 2019 on Creativity and Cognition*. 144–155.

[15] Seth Glickman, Nathan McKenzie, Joseph Seering, Rachel Moeller, and Jessica Hammer. 2018. Design Challenges for Livestreamed Audience Participation Games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY*

’18). Association for Computing Machinery, New York, NY, USA, 187–199.

[16] David Gonçalves, Manuel Piçarra, Pedro Pais, João Guerreiro, and André Rodrigues. 2023. "My Zelda Cane": Strategies Used by Blind Players to Play Visual-Centric Digital Games. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Hamburg</city>, <country>Germany</country>, </conf-loc>) *(CHI ’23).* Association for Computing Machinery, New York, NY, USA, Article 289, 15 pages. https://doi.org/10.1145/3544548.3580702

[17] Daniel Gros, Brigitta Wanner, Anna Hackenholt, Piotr Zawadzki, and Kathrin Knautz. 2017. World of Streaming. Motivation and Gratification on Twitch. In *Social Computing and Social Media. Human Behavior.* Springer International Publishing, 44–57.

[18] William A Hamilton, Oliver Garretson, and Andruid Kerne. 2014. Streaming on twitch: fostering participatory communities of play within live mixed media. In *Proceedings of the SIGCHI conference on human factors in computing systems.* 1315–1324.

[19] Noor Hammad, Erik Harpstead, and Jessica Hammer. [n. d.]. Towards a Design Framework for Data-Driven Game Streaming: A Multi-Stakeholder Model. *Proc. ACM Hum.-Comput. Interact.* UNDER REVIEW.

[20] Noor Hammad, Erik Harpstead, and Jessica Hammer. 2023. The View from MARS: Empowering Game Stream Viewers with Metadata Augmented Real-time Streaming. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (<conf-loc>, <city>San Francisco</city>, <state>CA</state>, <country>USA</country>, </conf-loc>) *(UIST ’23, Article 62).* Association for Computing Machinery, New York, NY, USA, 1–13.

[21] John Harris and Mark Hancock. 2019. To Asymmetry and Beyond! Improving Social Connectedness by Increasing Designed Interdependence in Cooperative Play. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI ’19, Paper 9).* Association for Computing Machinery, New York, NY, USA, 1–12.

[22] Zorah Hilvert-Bruce, James T. Neill, Max Sjöblom, and Juho Hamari. 2018. Social motivations of live-streaming viewer engagement on Twitch. *Computers in Human Behavior* 84 (2018), 58–67. https://doi.org/10.1016/j.chb.2018.02.013

[23] Joonyoung Jun, Woosuk Seo, Jihyeon Park, Subin Park, and Hyunggu Jung. 2021. Exploring the Experiences of Streamers with Visual Impairments. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2 (Oct. 2021), 1–23.

[24] Daniel Killough and Amy Pavel. 2023. Exploring Community-Driven Descriptions for Making Livestreams Accessible. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (<conf-loc>, <city>New York</city>, <state>NY</state>, <country>USA</country>, </conf-loc>) *(ASSETS ’23).* Association for Computing Machinery, New York, NY, USA, Article 42, 13 pages. https://doi.org/10.1145/3597638.3608425

[25] Kinetic Games. 2020. Phasmophobia. Video game. https://store.steampowered.com/app/739630/Phasmophobia/

[26] Jozef Kulik, Jen Beeston, and Paul Cairns. 2021. Grounded Theory of Accessible Game Development. In *Proceedings of the 16th International Conference on the Foundations of Digital Games* (Montreal, QC, Canada) *(FDG ’21, Article 28).* Association for Computing Machinery, New York, NY, USA, 1–9.

[27] Pascal Lessel, Maximilian Altmeyer, and Antonio Krüger. 2018. Viewers’ Perception of Elements Used in Game Live-Streams. In *Proceedings of the 22nd International Academic Mindtrek Conference* (Tampere, Finland) *(Mindtrek ’18).* Association for Computing Machinery, New York, NY, USA, 59–68.

[28] Pascal Lessel, Alexander Vielhauer, and Antonio Krüger. 2017. Expanding Video Game Live-Streams with Enhanced Communication Channels: A Case Study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI ’17).* Association for Computing Machinery, New York, NY, USA, 1571–1576.

[29] Meta. 2024. Instagram. https://www.instagram.com/live/.

[30] Matthew K Miller, John C Tang, Gina Venolia, Gerard Wilkinson, and Kori Inkpen. 2017. Conversational Chat Circles: Being All Here Without Having to Hear It All. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI ’17).* Association for Computing Machinery, New York, NY, USA, 2394–2404.

[31] OpenAI. 2021. ChatGPT: A Large Language Model for Conversational Agents. OpenAI Blog. https://openai.com/blog/chatgpt

[32] Rui Pan, Lyn Bartram, and Carman Neustaedter. 2016. TwitchViz: A Visualization Tool for Twitch Chatrooms. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA ’16).* Association for Computing Machinery, New York, NY, USA, 1959–1965.

[33] Karine Pires and Gwendal Simon. 2015. YouTube live and Twitch: a tour of user-generated live streaming systems. In *Proceedings of the 6th ACM multimedia systems conference.* 225–230.

[34] Dennis Ramirez, Jenny Saucerman, and Jeremy Dietmeier. 2014. Twitch plays pokemon: a case study in big g games. In *Proceedings of DiGRA.* 3–6.

[35] Joseph Seering, Saiph Savage, Michael Eagle, Joshua Churchin, Rachel Moeller, Jeffrey P Bigham, and Jessica Hammer. 2017. Audience participation games: Blurring the line between player and spectator. In *Proceedings of the 2017 Conference on Designing Interactive Systems.* 429–440.

[36] J T Sheng and S R Kairam. 2020. From virtual strangers to irl friends: relationship developmentc in livestreaming communities on twitch. *Proceedings of the ACM on Human-Computer* (2020).

[37] Max Sjöblom and Juho Hamari. 2017. Why do people watch others play video games? An empirical study on the motivations of Twitch users. *Computers in human behavior* 75 (2017), 985–996.

[38] The Oh Lab. 2021. Bloomwood Stories. http://theohlab.org/bloomwood-stories. Accessed: April 5, 2023.

[39] Piaohong Wang and Zhicong Lu. 2023. Let’s Play Together through Channels: Understanding the Practices and Experience of Danmaku Participation Game Players in China. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY (Oct. 2023), 1025–1043.

[40] Yujia Wang, Wei Liang, Haikun Huang, Yongqi Zhang, Dingzeyu Li, and Lap-Fai Yu. 2021. Toward Automatic Audio Description Generation for Accessible Videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Yokohama</city>, <country>Japan</country>, </conf-loc>) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 277, 12 pages. https://doi.org/10.1145/3411764.3445347

[41] Saelyne Yang, Jisu Yim, Juho Kim, and Hijung Valentina Shin. 2022. CatchLive: Real-time Summarization of Live Streams with Stream Content and Interaction Data. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22, Article 500)*. Association for Computing Machinery, New York, NY, USA, 1–20.

[42] Ryan Yen, Li Feng, Brinda Mehra, Ching Christie Pang, Siying Hu, and Zhicong Lu. 2023. StoryChat: Designing a Narrative-Based Viewer Participation Tool for Live Streaming Chatrooms. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Hamburg</city>, <country>Germany</country>, </conf-loc>) *(CHI '23, Article 795)*. Association for Computing Machinery, New York, NY, USA, 1–18.

[43] Youtube. 2024. YouTube. https://www.youtube.com/live/.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009