‘Did you see that!?’ Enhancing the Experience of Sports Media Broadcast for Blind People

Cagatay Goncu[1,2]
Daniel J. Finnegan[3]

[1] Monash University, Australia
[2] Tennis Australia
cagatay.goncu@monash.edu
[3] Cardiff University, UK
finnegand@cardiff.ac.uk

Abstract. Accessibility in sports media broadcast (SMB) remains a problem for blind spectators who wish to socialize and watch sports with friends and family. Although popular, radio’s reliance on low bandwidth speech results in an overwhelming experience for blind spectators. In this paper we focused on two core issues: (i) how SMB can be augmented to convey diegetic information more effectively, and (ii) the social context in which SMB are consumed. We chose tennis broadcasts for our investigations. Addressing issue (i), we developed a system design and prototype to enhance the experience of watching tennis matches, focusing on blind spectators using audio descriptions and 3D audio, and evaluated our system with (n=12) in a controlled user evaluation. Our results indicate how audio descriptions gave clear information for the tennis ball placements, 3D audio provided subtle cues for the ball direction, and radio provided desired human commentary. For issue (ii), we conducted an online questionnaire (n=15) investigating the social context in which blind spectators consume SMB. Participant feedback indicated there is a demand for more accessible SMB content such that people can consume SMB by themselves and with their friends. Participants were enthusiastic for a revised system design mixing elements from 3D audio and audio description. We discuss our results in the context of social SMB spectatorship, concluding with insights into accessible SMB technologies.

Keywords: accessibility · blind and low vision people · sports broadcasting · tennis · spectatorship

1 Introduction

Sports Media Broadcast is a hugely popular pastime enjoyed by millions of people, and for many is a viable alternative to attending events in person. SMB lets viewers experience what attending the event is like by simulating the experience of co-presence: microphones and cameras capture events happening on and around the sports pitch, delivering content regarding actions in the sport event and the reactions of fans in the arena. The most dominant form of SMB is on-demand televised (TV) experience, where events are broadcast with audio narration and commentary from professional commentators, celebrities, and journalists who dictate all the action as it happens in the arena. While providing a state of the art experience for sighted spectators, blind spectators must tune in to radio broadcasts to receive information which is normally captured in the video stream, for example team formation provided via info graphics.

Both radio and TV SMB suffer from problems with respect to accessibility: for example blind people may miss out on essential action happening within the SMB if the actions are not described. Broadcast media for blind people typically requires extra information in the audio channel: for example, audio descriptions which provide context and detail around what is happening in the TV broadcast. Television shows typically insert brief snippets of spoken audio in between dialogue to convey information that is captured in the video, for example the emotions of the characters like disgust and fear on their faces [36]. However, this is not systematically described in a fast paced environment like SMB: instead, blind spectators are encouraged to consume a separate channel such as radio, removing them from the shared social experience. In a social setting, blind people may therefore depend on social interventions: friends dictating parts of the action or describing the meta data around a game event, for example, who scored the recent goal which won the soccer game. Interestingly, in the early 1920s, the genesis of sports broadcasting, the British Broadcasting Commision (BBC) radio commentators used an experimental structured commentary technique for soccer[4]. Sports commentators using this system coined the phrase 'back to square one' as the soccer field was divided into 8 cells, and commentators used each cell's associated number to relay the position of the ball and the players. For example, during a match the commentator would provide real time audio commentary while their colleague would declare 'Square 2' ... 'Square 6' as the ball moved around the pitch[5]. Though innovative at the time and providing essential information for spectators, this technique of structured commentary providing clear descriptions of where the football is has been forgotten in time and is no longer used in main stream sports broadcast commentary [18].

Our goal is to innovate and provide a better blind spectatorship experience than de facto radio broadcasting. Potential avenues lie in creating augmented broadcasts which add extra information such as audio descriptions and sonifications to the original SMB: here we focus on augmentations conveyed using 3D audio as it provides spatial information without the need of description or commentary. These augmentations could provide blind spectators with a better experience, with information regarding events and actions that go unreported in the SMB commentary. This will provide a complete picture of what is happening, without impacting the experience of sighted users in social consumption environments, and help the blind spectators form a mental image as similar to graphical information [20].

[4] An image of the football pitch showing the demarcation of zones can be found at `http://news.bbc.co.uk/sport2/hi/football/1760579.stm`

[5] `https://www.theguardian.com/notesandqueries/query/0,5753,-1811,00.html`

Therefore, we seek to address the following research questions:

RQ 1: How do blind spectators consume SMB and what, if any, preferences do they have regarding radio or TV?
RQ 2: Is augmented SMB more effective in conveying SMB information compared to radio?

To facilitate social SMB consumption, we present an iteratively designed prototype system for spatializing tennis SMB events using an augmented audio channel to aid blind spectators interpret the event. The system captures key information during the SMB event and renders it in real time so that blind people can consume broadcasts with their sighted friends and family without negatively affecting one another's overall experience. We also present the results of a questionnaire which investigates the current methods of consuming SMB by blind spectators and their overall experience.

We chose tennis because: (i) it is a very structured game, (ii) during game play it has long quiet time slots which creates opportunities for SMB augmentation, and (iii) we can access detailed match data from Tennis Australia. Our system incorporates elements from different modern day SMB–commentary from radio, multi channel information from TV–, to enhance the experience for blind spectators. Our main contribution is an analysis of the shortcomings in SMB: particularly how it falls short of presenting the full content of a sport event to blind spectators, and a potential prototype system that can remove some of these shortcomings. We conclude that SMB for blind spectators may, and arguably should, raise a new design challenge for the HCI community.

2 Background

The dominant forms of technology mediated broadcast for sports spectatorship are radio and television, with both providing real time commentary describing what is happening throughout the broadcast. For blind spectators, radio remains the most effective mode of consumption as greater emphasis is placed on describing what is happening via commentary due to the lack of a visual broadcast channel. Technology has been applied over the past few decades to enhance the experience of sports spectatorship [37]. Most work around augmenting media broadcast revolves around enhancing the experience through the use of predominantly visual displays [34, 12], via augmented reality [48], automated processing of statistical data [2, 35], and enabling consumption of ever growing repositories of visual information. Existing methods for making pre-recorded videos accessible to the visually impaired community use audio descriptions (AD)–spoken dialogue detailing what is displayed on the screen. However, ADs may not be suitable for live broadcast: they require identifying appropriate gaps in the video, so that descriptions do not overlap with other spoken dialogue content. In a live broadcast, one does not know when appropriate gaps for inserting ADs may appear. Recent work has investigated automatic generation of ADs applied to live sports, showing subjective improvement in understanding by blind SMB consumers [19, 24, 43]. However, these systems focus on very simple event based commentary, for example a point was scored, and are limited in their ability to capture more nuanced play which spectators may wish to follow, for example, the trajectory of the ball. Recent work concerning a system designed for general audience to sonify soccer games used computer vision techniques to capture movement of the ball and the players, generating sonifications for key features such as possession, passes, steals, and goals [39]. In this study pitch mapping was used to map the ball's distance to goal to the pitch of the sound, and this was found more enjoyable. It was also reported that key moments sonification which provided events that focused on passes, steals, and goals were found more useful. The study also reported the need for a more deeper investigation in improved spectator experience.

Why focus on the spectator experience? SMB is a popular pastime for blind people [32] with many people spending a significant amount of time spectating sports events on television. Sports are seen as extremely difficult to follow when watching television [32]: factors contributing to this difficulty involved the clarity of dictation, long scenes of quiet exposition where much attention is focused on visual camera panning with little to no audio description, and poor sound quality. While work from the HCI community has sought to develop systems to aid blind participation in sport [22, 10, 17], it has curiously neglected sports spectatorship and the problem of creating accessible environments for spectating live sports events. We are interested in the role of HCI in sports spectatorship for blind people due to several benefits it may bring, for example, to health and wellbeing [21] and social opportunities [9]. For example, the feeling of catharsis that can come from "vicarious participation in sports" achieved through spectatorship [47], and the sense of recreation accompanying SMB consumption. Other work has developed systems to enable blind participation in sport for exercise and recreation, for example using sonification techniques to render player and ball movement [38], and using spatial audio for locating the ball positions in virtual reality games [46, 42].

We begin by observing the motivations: what drives sport spectatorship? Several studies refer to the appeal of vicarious achievement or achievement gained through observing others [33], but also the social factors surrounding sports spectatorship [44, 27]. One may infer the motivations driving sports spectatorship for blind people are the same, however research is scarce in this regard. Classic work in HCI has created taxonomies of interaction and how to provide meaningful engagement for the spectator [37], however these taxonomies are not discussed in the context of accessible design, for example being able to observe one perform gestures in public without necessary audio description or tactile information pertaining to the gesture or other event. Designing for and facilitating the spectator experience for blind people is what we focus our attention on in this paper.

Augmenting SMB via Sports Commentary: Modern approaches to augmentation generally operate by broadcasting meta information in another representation, for example audio descriptions, to a client's companion device. One core issue persists with these approaches: they rely on the meta information that is broadcast with the main channel. This meta-information takes the form of a descriptive narrative over what is happening on the screen. Users are required to concentrate on this descriptive content while simultaneously paying attention to the main content of the program e.g. dialogue and action sounds. Thus, content consumers are subjected to divide their attention between diegetic events–events happening within the broadcast content that they can perceive, such as crowds roaring–and non-diegetic descriptive information. In e-sports, researchers have explored how to enhance the spectator experience of live game streams through enhanced communication channels [28], focusing not on what to present but how and when to present it [5]. As e-sports events are inherently digitally mediated, all the relevant information is immediately available and thus the risk of missing any information which the commentator fails to report may be mitigated by displaying the raw information directly to the observer in whatever format they choose. However these options are still limiting for blind consumers, creating barriers and compounding the issue with information overload and confusion by requiring the spectator to attend to even more channels. Attending to multiple sources of information is a difficult task, causing higher cognitive load and even stress on the consumer, and attention is known to be poor unless people are trained in divided attention tasks [41].

Since the early days of radio, sports has been broadcast with a commentator whose role is to provide a real time account of what is happening in the game. Sports commentary itself is typically provided in two concurrent flavors: the first providing descriptive content regarding the event while the other provides more dramatic commentary, known as 'color commentary'. One aspect of SMB color commentary is the added value brought by the commentator who creates a narrative discourse [26]. Their role focuses on conveying the emotional aspect of sports, and their commentary is thus fuelled with passion and exuberance. As they seek to augment the broadcast experience for listeners by storytelling based on similar events in previous games, this value proposition hinges on the listener's capability to map current events in the game to the story being told.

Finally, color commentators provide contextualized information for listeners: for example, if a player wins a point by hitting a between-the-legs shot, the color commentator may explain this as extraordinary as the player is normally right footed. If listeners can't fully comprehend the current game events then this will negatively impact their ability to enjoy the story. Furthermore, they may begin to confound the story with current events and misinterpret historical from current events in the broadcast. With respect to tennis, commentary typically comes between points where color commentators discuss the players themselves and their historical games. Consequences for blind spectators can be severe and can lead to further barriers: for example, missing commentary impacts a blind spectator's ability to engage with particular events happening in the game. As social peers may begin a discussion regarding an event with missing commentary, blind spectators cannot engage fully with the shared social discourse. Although follow up discussions (i.e. post broadcast analyses) provide opportunities for in-depth and discussion which blind spectators may engage in, this would fall short of the same problems in programming/broadcasting where programme directors and media personalities may not share the same events and/or prioritize some events over others.

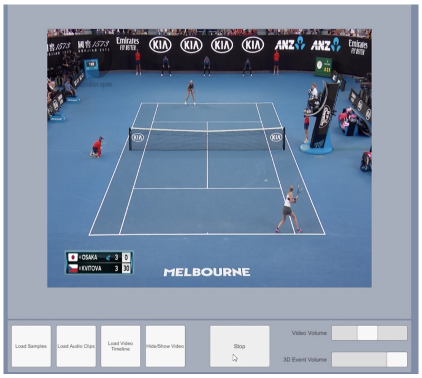
3 Pilot Studies and Initial Prototype Design

Thus, there are opportunities for the role of technology to enhance SMB and provide a better blind spectatorship experience than de facto radio broadcasting. Before embarking on our system design, we conducted pilot studies with two blind spectators to identify the level of access to tennis SMB. In the pilot studies, we did unstructured interviews with one male and one female participant both of whom follow sports, in particular tennis matches, on both radio and TV. We conducted these interviews iteratively while we were doing improvements in the prototype system. In each iteration (in total 3) we asked the questions to both pilots participants.

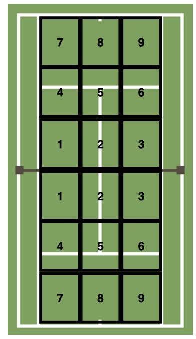
The first study involved informal discussions around the context in which they consume SMB. They reported that they use radio to follow tennis matches, while using their mobile phone to browse various applications and web sites. They also commented how they use radio receivers with headphones in public spaces so not to ruin the experience for others. From their comments, it was clear that they could get a general description of the game, but they could not get a lot of information about how the game is actually played, for example where the ball bounced and/or other game specific actions such as whether the players hit the ball hard or with a heavy top spin. For instance, if a player hits a down-the-line forehand after having a long sprint and wins the point, radio commentators have just enough time to describe this action: in this case there is no time to describe the ball's movement and/or the player locations. Our pilots also reported watching the matches on TV is not a good option either, as the commentary does not describe much about the game play.

Based on the first pilot study, we considered augmenting tennis SMB in a way that blind spectators can access game play information. Our system is designed for blind spectators to consume SMB using an augmented broadcast channel delivering audio description and 3D audio content which a) does not overwhelm the spectator and b) does not occlude other elements of the original broadcast. This facilitated the investigation of our **RQ 1**. Our system is designed to be used in a setting similar to watching a live-TV broadcast at home or in a public space, instead of going to a match in a stadium. As the radio broadcasts are already fully occupied, we decided to augment the TV broadcasts. In tennis, crowd noise is kept to a minimum during game play and broadcast commentators also keep quiet during this period. This gives us commentator free time slots that can be used to add additional information about the game play including audio descriptions and 3D audio content covering events and actions that go unreported in the SMB commentary. This facilitated the investigation of our **RQ 2**.

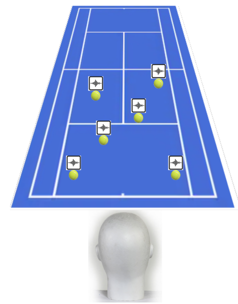
A screenshot of our system prototype's user interface is shown in Figure 1a. The prototype uses pre-recorded video footage and 3D positional data to create



(a) System prototype user interface



(b) Tennis court grid system used for audio descriptions.



(c) Sample 3D audio placements used for 3D audio

Fig. 1: System components

augmented SMB content consumable using off the shelf headphones. The 3D positional data are provided from the ball tracking system, an industry standard system for tracking ball movements in major sporting games. The ball tracking data (JSON files containing timestamped Cartesian coordinate system positional information about the ball) are used to select synthesized audio descriptions, and as input to a 3D audio rendering subsystem. Our prototype takes video recordings of tennis matches and the ball tracking data as its inputs and generates the augmented videos with audio description and 3D audio as its output. Pre-recorded video are used for conducting the user evaluation in an easier fashion, however the prototype is suitable for live consumption as it uses the ball tracking data format that is identical to the real time tracking data which can be received within 100ms. Our prototype also features simple slider bar elements for adjusting the volume of the original broadcast and audio description / 3D sounds separately.

In the prototype, left and right of the tennis court are defined from the spectator's perspective sat behind the nearest tennis player as in Figure 1a. The server is not fixed to a particular side of the screen, so the server can be either the bottom player or the top one. Each side of the court is divided into a 3 by 3 grid providing finer detail as shown in Figure 1b. The audio descriptions provide information about the bounce location of each ball during the play. These are vocalized using synthesized speech from descriptions provided by an expert tactile graphics transcriber.

We use binaural audio to implement our 3D audio augmentation. Binaural is a form of 3D audio based on how we hear sounds naturally. It requires only stereo channels which are filtered using anthropomorphic models of the human head and pinnae called head related transfer functions (HRTFs). Though it is possible to create bespoke HRTF models for individual listeners, this is logistically challenging and was not feasible for our prototype. Instead we used a generic (HRTF) which is good enough for sound localization when users are given the opportunity to train using them [1, 7]. During piloting, the comments from participants regarding their use of headphones to not disturb others in public spaces was the main factor driving our design decisions for spatialized audio. As our system leverages headphones for 3D, the participants will use our system as they do their radios. We decided on using a binaural HRTF design for 3D audio due to it not requiring specialized equipment, for example multi channel headphones.

3.1 Issues with Pilot System Design

The second pilot study involved gathering feedback on the initial prototype design. Discussions with our pilot participants focused on issues around audio descriptions and the augmented SMB channel. For example, we first used 'short' and 'long' to describe the ball bounce location as an indication of its distance to the net. Although these labels were accurate, they caused problems for our pilot participants because of their game specific meanings. In tennis, the term 'long' is used to represent an out-of-bounds hit due to the ball landing beyond the opposing baseline. Similarly the term 'short' is used when the ball bounces near the service line.

For 3D audio, initially we had conceived an egocentric perspective where the listener is placed in the middle of the court. In this setup, sounds are rendered in 3D based on the actual ball bounce location obtained from the ball tracking data. For example, when the ball hits the upper left corner of the tennis court, our 3D sound augmentation would be heard to the front left of the listener while a ball hitting the lower right of the court would be heard from behind and to the right of the listener. This initial design caused two problems: first, the 3D audio was difficult for users to interpret as they found it disconcerting to imagine themselves standing in the center of the court during game play. Secondly, the 3D sounds did not correlate with the typical coordinate system used in tennis commentary. This caused further confusion for the pilot study participants who found it cognitively demanding to interpret. To summarize, the issues identified in our pilot studies are: **I1)** *Terminology used in audio descriptions regarding the location of the ball must consider game specific factors.* and **I2)** *3D audio must consider listener perspective and the coordinate system used in SMB commentary.*

4 Final System Design

To solve **I1**, we changed the terminology for 'short' and 'long' to 'near' and 'far' respectively. Thus, the grid squares are labelled left far (cell 7), centre far (cell 8), right far (cell 9) for the row furthest from the net; left near (cell 1), centre near (cell 2), right near (cell 3) for the row closest to the net; and left middle (cell 4), centre middle (cell 5), right middle (cell 6) for the middle of the court. The labels are converted to synthesised speech recordings by using the

Speech Synthesis manager[6] on macOS. These recordings are then used as audio descriptions which can be reused in every match without any modification.

As an example, consider the following play: (i) the server starts the point with a serve to the centre middle service line, (ii) the receiver returns a short ball to the left side, and (iii) the server hits the ball down-the-line on the right side on the court, but misses the shot. The system gets the exact ball bounce positions from the ball tracking data, and matches it to one of the cells in the grid. Using this match it uses the relevant audio description. Thus, for this particular 3 shot point, the system provides 'centre middle', 'left near', and 'right far' audio descriptions at the time the ball bounces on the court.

The audio descriptions only describe the locations of ball bounces. They will not let you know whether the ball is in or out. This is based on the design consideration for not occluding other elements of the original broadcast. Original TV broadcast still has the line umpire calls in the video, so the out-calls will still be audible. Therefore, for this particular example, the final output starts with the original TV broadcast sounds, i.e. crowd noise, at the beginning. Then, 'centre middle', 'left near', and 'right far' audio descriptions will be heard at the specific times of the ball bounces. As before, the out-call from the umpire that comes from the original TV broadcast will be heard.

To solve **I2**, we placed the viewer behind the court and rendered 3D audio from the perspective of the video camera capturing the TV footage. This provided a 1:1 mapping between the visual scene and our 3D audio augmentation. This was done to mimic the same conditions of a blind spectator watching a tennis match in front of a TV. Though TV broadcasts use different view angles during a match, for example focusing on individual players in between serves for dramatic effects, they use a fixed view angle when the ball is in play as seen in Figure 1a. This view does not change based on the server position. As the players change over after every two games, the server will be seen either at the top or the bottom of the view. We use this view for the audio augmentation to avoid any disorientation. A short click sound was used for our 3D audio augmentation. Therefore, for the same example described in Section 3, 3D positioned click sounds are rendered in time when the ball bounces off the tennis court. As with audio descriptions, out-calls from the original broadcast umpire are heard.

5 Questionnaire and User Evaluation

We conducted a mixed methods approach involving an online questionnaire and a controlled user evaluation, both with blind participants. The online questionnaire probed participants on their spectating experience. Their comments give insight into how blind spectators consume SMB, and what their preferences and experiences are (**RQ 1**). Their comments give insight into how blind spectators consume SMB. In the controlled user evaluation, we evaluated how our system may enhance the experience for blind spectators using augmented SMB

[6] `https://developer.apple.com/documentation/applicationservices/speech_synthesis_manager`

delivered through audio descriptions and 3D audio. We tasked participants with answering questions based on what they viewed during the evaluation, and their preferences in three different forms of SMB. Task performance in the evaluation determined the effectiveness of SMB augmentations compared to radio (**RQ 2**).

We recruited 14 blind participants (8 female), all legally blind (i.e., a visual acuity of 20/200 or less). using opportunity sampling advertised using social media channels and email lists. We disseminated the online questionnaire to 14 participants, 12 of which were included in the user evaluations: 2 participants could not do the user evaluations due to poor internet connections. Participants ranged from 26 to 74 years old ($\overline{x}$ = 49, sd=17), and granted their consent to take part. A $30 e-gift card was given to participants for their time.

Online Questionnaire Participants completed the questionnaire before the user evaluation, so that our system design would not effect their comments. We used an online system (Google Forms) to disseminate the questionnaire. To ensure the questionnaire was accessible for our participants, we used simple radio buttons and text field controls instead of drop down menus and linear scales. The questionnaire contained open ended questions on sports media broadcast consumption, asking participants which sports they follow and/or play, how they follow these sports, what they like/dislike about SMB, what type of SMB they want to have, and the social environment of consumption. Two questions specifically focused on tennis were: (i) "If tennis is one of the sports that you have been following, how do you follow the tournaments and matches?", and (ii) "Do you play tennis? Which other sports do you play or would like to play?".

User Evaluation COVID-19 forced us to conduct evaluations remotely over video conferencing. Thus we presented our prototype system output as pre-recorded videos on a single platform (Google Forms) that both serves the stimuli and records participants' responses. Doing so minimizes context switching between apps, avoiding additional stress on participants. As we aware of disadvantages and problems around conducting remote evaluations with participants with disabilities [31], we gathered feedback on the online form from a member of the blind community working at Monash University before recruiting our participants. Based on their feedback we updated the form to ensure screen readers can read it without any issues on desktop and mobile platforms, for example partitioning the form into sections so that participants could navigate it easier. We disseminated the controlled user evaluation links to them by email at the beginning of the sessions to mitigate difficulty issues accessing the chat messages on the video conferencing platforms

Our user evaluation consisted of 3 presentation conditions (audio description, 3D audio, and radio). We did not consider the combination of presentation methods to keep the study duration under 90 minutes, avoiding participant fatigue. In creating videos for our user evaluation, we sampled points from three different tennis matches: Osaka vs. Kvitova, 2019; Cilic vs. Federer, 2018; and Djokovic vs. Nadal, 2019. We created 4 MP4 videos (1 for training, and 3 for testing) for each condition, resulting in 12 videos in total for our evaluation. These videos were counterbalanced to avoid learning effects. Each video captures a different point in one of these matches: for example, video 2 in the 3D audio condition was the 3rd point of the 7th game in the 2nd set of the Osaka vs. Kvitova, 2019 match, and video 1 of the radio condition was the 1st point of the 3rd game in the 1st set of the Djokovic vs. Nadal, 2019 match. Using video recordings of our prototype system output enabled us to conduct the user evaluations remotely, as participants would not need to install our software locally. Each of these videos represent a point of a tennis match with increasing complexity having 2, 3 and 5 shots each, including the training material. The videos were then uploaded to a private online repository and embedded into the online form. For each video we gave participants three tasks, designed by an expert tennis player, to represent the game and player's actions: one of the fundamental needs of spectators [14]:

T1: Which player (server or receiver) won the point?
T2: Can you describe the movement of the ball?
T3: How competitive/skilful was the point in a scale of 1 (low skilful) to 5 (high skilful)?

T1 is simply a binary choice between the server or receiver, T2 is a multiple choice question (each choice is a textual description of a point play), and T3 is a numeric value from a 5 point scale, with the correct answer pre-determined by one of the researchers involved in this study who is also an expert reviewer.

Each participant performs the tasks for each of the 3 presentation conditions in a repeated measures design: audio description (AD), 3D audio, and radio (R). The order of the presentation conditions was balanced using a Latin square. Participants were presented the training video first, followed by 3 test videos. They answered T1, T2 and T3 for each of these test videos. After completing all the trials, participants were asked to rank their preferred presentation conditions. We used preference rankings as they are an important factor in evaluating spatial audio reproduction systems [11]. They were then invited to provide comments explaining what they like and dislike for each of the presentation conditions. Additionally, they were asked about the main considerations that affected their preferences. We used video conferencing software to monitor participants during the evaluation. Participants used screen sharing to let us observe their progress. They also shared their computer sound with us, so that we could hear the screen reader prompts. Each evaluation took on the order of 90 minutes to complete.

6 Results

Online Questionnaire *What blind people like and dislike about SMB content:* Radio was the preferred medium for following sports due to several distractions from the TV broadcast. Comments from the online questionnaire raised issues with TV advertisements interrupting the game and causing confusion, and TV providing a worse *'picture' (P4)* compared to the one perceived from listening to radio commentary. However some preferred the color commentary from TV commentators as they add a *'sense of humour' (P6)* while others found this to be *'excessive' (P13)*. Though radio was perceived as providing more detail about what is happening compared to TV broadcasting, one participant commented on the declining quality of radio broadcasting. They described how commentators spend more time *'having a chat with one another' (P12)* and growing interest in *'chemistry between broadcaster and expert summariser [sic] rather than a focus on the game itself' (P12).*

> "I like that radio is designed to not be able to see the picture so they often describe visual aspects of the game as well as commenting eg. weather, crowd numbers and they add a lot more statistics and description to their commentary. It makes it very difficult if radio commentary or TV commentary gets distracted and they go off track and are discussing things that are not relevant to the game. . . . the TV tends to do this more often because the assumption is made that you can keep up by watching they [sic] vision while they discuss something else. Also TV commentary often uses less descriptive language, eg. did you see that, have a look at that kick." (P3, emphasis ours)

The social context(s) in which blind people consume SMB content: Several participants noted how they typically watch SMB content with others as a social activity. The context in which social interaction takes place differed across participants: for some, SMB content was enjoyed at home with family and friends while others would go to pubs and bars. Finally, a select few would attend live events and bring with personal radios to listen to live commentary while watching the game. *"I tend to listen on my own via radio, or with family/friends. I sometimes go to the pub to watch football with friends." (P12).* Almost all participants said they would also watch tennis broadcasts by themselves.

Summary: Results from our online questionnaire reveal the diversity in participants' relationship with modern SMB and the issues they face. Most participants declared radio as their first preference, and had strong opinions on the commentary provided in traditional broadcasting and its impact on their consumption experience and ability to comprehend the action in and around the game. They were also concerned about others as they mostly consume SMB in social contexts: key issues revolve around not wanting to negatively impact the experience of others while at the same time relying on friends and family for bespoke descriptions about what is happening.

User Evaluation To analyse the results from the user evaluation tasks (T1, T2, T3), we conducted a Cochran's Q test, a significance test suitable for repeated measures block designs where the response variable is binary. For determining the victor of the match point (T1), a Cochran's Q test determined that there was no difference in the proportion of participants who were correct over the different presentation conditions, $\chi^2(8) = 10.0, p = .26$. There was also no difference in the proportion of participants who could correctly determine the movement of the ball (T2) ($\chi^2(8) = 12.5, p = .13$) nor who could determine the skillfulness of the players (T3) in the broadcast event ($\chi^2(8) = 7.7, p = .46$). Figure 2a shows the proportion of correct answers for each task in each presentation condition.

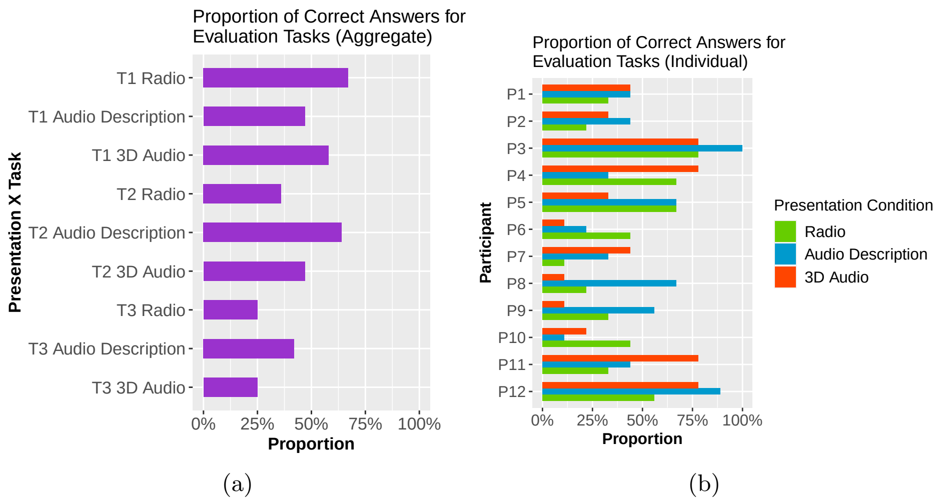


Fig. 2: User evaluation results: Proportions of correct answers for all participants, and individually.

Figure 2b shows the proportion of correct answers for each participant[7]. Some participants (P3, P12) are 'power users' and score highly in each presentation. Most of the participants performed better with audio description than radio. Half of the participants (P1, P2, P4, P7, P11, P12) did better with 3D audio than radio. With the exceptions of P6 and P10, all participants performance with radio is at best on par with audio descriptions and 3D audio.

To test for a difference in preference rankings between all three presentation conditions, we conducted a Friedman Test on the preference votes. There was no difference between preferences across all participants for Radio, 3D audio, or Audio Description, $\chi^2\,(2) = 0.5, p = .78$. Figure 3, Panel A shows the preference rankings (1st, 2nd, and 3rd) for each presentation condition, and Panel B shows the proportion of correct answers across all tasks in each presentation condition. Audio description has 51%, 3D audio has 44%, and Radio has 43% correct answer proportion. This is striking: participants can't correctly interpret *more than half of the game play information for radio–the de facto medium–yet it is the most preferred presentation condition.*

We used thematic coding to analyse participant comments, explaining what they like or dislike about the presentation conditions, and the main considerations for their comments. We used two independent coders to look through the comments, and identified keywords and phrases to construct codes. Finally, we then categorized these codes into the following themes: the context where they can use each condition, the cognitive load introduced by the augmentations, the ability of tracking the ball location, the amount of information each condition provide, the quality of information they get, the distractions they experience, and the factors effecting their overall experience. For each theme, we report the

[7] P3 and P11 indicated that they passionately follow sports.

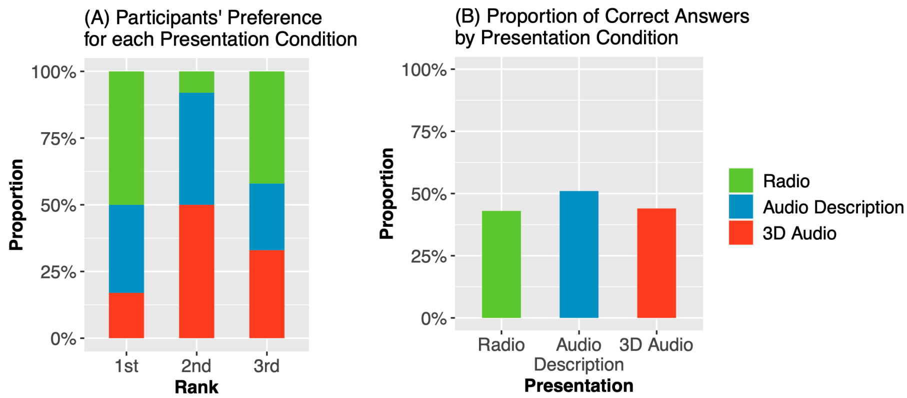


Fig. 3: Participant preference towards each SMB medium used in our user study.

Table 1: The mappings between codes and themes used in the thematic analysis, and the total number of times the codes are mentioned by the participants. The themes that have more relevance to tennis are marked with a star (*).

Themes	Codes	No. of Mentions
Tracking the ball (*)	position of the ball	10
	precision	5
	raw play information	10
Pros and Cons of Commentary	capturing player emotions	2
	natural feeling of radio	9
	dislike of synthetic speech	9
	randomness of commentary	8
Information Load	overwhelming	7
	high concentration for AD and 3D audio	9
Spectator Experience	radio with headphones	11
	social desire to follow the action with others	3
	distractions from others	3
Accuracy and Consistency of the Augmentation (*)	density of information	6
	expanding bandwidth of information	7
Usefulness of the Augmentation (*)	3D audio providing subtle feedback	3
	3D audio difficult at first	2

number of participants in parenthesis. See Table 1 for our map between codes and themes generated.

Tracking the Ball: Participants liked how our system gave consistent and precise ball locations in contrast to radio that depends on the human commentators (10). Many participants stated they want to get raw play information about the status of activity (including ball and player positions) on the field (10): Some participants liked the clear audio description for the placement of the ball, however they were not very happy with the way they took away the crowd noise (5). They found the grid structure helpful for ball placement locations, and the fact that they could picture this information on the court.

Pros and Cons of Commentary: Emotions were considered as a part of the experience (2). Radio was considered as the most natural way, and audio description was considered a mechanical feedback (9). Radio was the most familiar presentation condition for the participants. Participants liked the human commentators, however they pointed out that sometimes they missed the play, and they had to guess the position of the ball and the players (8). They also said that radio provided the vibe of the game better than the others (5). Another issue regarded the random information given by commentators about the players and their shot selections (3). Some participants found this distracting (5), while others said that this information helped them understand the player strategies (2). Some participants complained about radio's lack of timeliness and consistency, being prone to human commentator error, bias, and irrelevant chat during points (8). Some participants said that they had their own favorite commentators who describe the game better than others (3). So, the style of commentating is a factor for the overall radio experience, and could sometimes give no clue what was happening in the game.

Information Load: Participants commented how the SMB augmentation provided by our prototype did not overwhelm or hinder their enjoyment of the experience (7). However, some participants mentioned high concentration levels for 3D and AD, but described how their cognitive load reduced as points were watched (2). Some also reported that they would use ADs by themselves at home, and 3D at home (3). ADs are preferred to get detailed information and keep up with every shot and its location (4). 3D audio is treated as a presentation condition which can be used when they want to get detailed information as well as not distracting friends (3).

Spectator Experience: Participants were asked about the context in which they want to use each of the presentation conditions. They reported that they would use radio with headphones when at home, at sports venues, outside in public spaces, or with others (11). Regarding the social context, it was surprising to see that participants were worried for other people in the same environment if they use radio and audio descriptions. They commented that these would create noise for their friends, and would ruin their overall experience (3). However, they said that this would not be an issue if they could use 3D audio as this would reduce the need for excessive commentary which participants perceived as the main factor creating noise and distractions for their friends (3).

Accuracy and Consistency of the Augmentation: Participants also commented on the amount of information given by each of the presentation conditions. Some said that radio broadcasts were too busy to include any other information such as the location of the ball and the player (6). ADs helped with the ball position, adding similar information about players, umpire, and the crowd would be distracting (7). 3D audio gave more space for other sounds that can be used for the players, umpire, and the crowd (6).

Usefulness of the Augmentation: 3D audio was found to be a presentation condition that participants had the least positive experience with. They liked the subtle feedback and how it did not remove the experience of the crowd noise and other broadcasting sounds such as calls for balls going out of bounds (3). Some participants commented that it would be useful for people who want to experience a new way of watching tennis matches, pointing out that adjustment of 3D audio could improve the experience (2).

7 Discussion

It's clear there is a need for improving the spectator experience for blind people. Although radio is the traditional way of consuming SMB, it cannot provide the information that blind spectators want to know. Our participants could not fully comprehend what was happening in the games with radio commentary provided. High speed speech is intelligible by blind people [30], and can be improved further with training [45], however the information content of fast speech is typically limited. In SMB the pace is fast and the information is vast, adding an extra barrier to consumption for blind people. Our presentation conditions provide these missing information in an accessible way **(RQ 2)**, yet participants still struggle to fully interpret the shots made.

It is surprising that none of our presentation conditions had a higher preference than the others **(RQ 1)**. They were equally as poor in conveying the details of what was happening in the SMB. Though our audio descriptions were developed to ensure direct mapping with the tennis court, and were refined through pilot testing, perhaps they remained unclear. Rapid successive shots from players may have led to short time intervals between bounces of the ball, which in turn would reduce the time between descriptions. Another issue may be the synthesized nature of descriptions: although generally considered acceptable, human speech is both preferred and easier to understand for audio descriptions [23]. We used a generic HRTF for rendering 3D audio because it was unfeasible to capture individual HRTFs. Although 'good enough' [1], this approach has its disadvantages as some sounds remain difficult to localize. More advanced techniques exist for synthesizing 3D sounds [6] may alleviate some of these problems and produce better 3D audio SMB augmentation.

Although radio was a popular medium amongst our participants, its information bandwidth is already occupied with dense commentary. This is a common disadvantage for blind people [29]. The commentaries help blind spectators to understand how the game is played on a higher level while omitting details such as the ball placements, player locations, and shot types. Our results demonstrate how radio broadcasting is inadequate in addressing the needs of blind spectators. Although it is very popular and the most relied on medium, it falls short in delivering game play details. In contrast TV broadcasts have more space to augment the experience. Tennis is one of the sports where the crowd and the commentators do not speak during game play, providing opportunities for augmenting the broadcast by adding audio description and/or 3D audio. However limitations for concurrent audio and speech should be considered [15, 16, 4, 8]. Color commentators typically focus on building the buzz and excitement around the event, emphasizing the reaction of the crowd, and therefore their commentary may only capture what the commentators themselves perceive as significant, restricting the listener's experience to only events commentators acknowledge. Thus there are limitations to consider regarding the type of the commentary which can prime listeners, and further inhibits understanding [13]. Also blind spectators may confuse broadcast effect sounds with game events, having detrimental effects on attention and the spectator experience [25].

We note the striking contrast between a qualitative preference towards radio broadcast with inconclusive quantitative results. Results from our user evaluation show no difference in performance (quantitative) when SMB is augmented with things participants say they want (qualitative). Though radio was preferred, it did not improve participants' accuracy in interpreting what is happening in the game **(RQ 2)**. Perhaps our participants simply regressed to their comfort zone: radio is familiar and they understand how it operates, where 3D audio is difficult to interpret at first and requires some training to acquire competency in using it **(RQ 1)**. Participants found the augmentations difficult to interpret at first, though this may have been due to unfamiliarity with 3D audio and the content of the audio descriptions. Previous work has highlighted how, particularly for information gathering tasks, efficiency with technological use is key for blind people [40]. We can see this insight replicated in our results: participants described how they could see the benefit of an augmented broadcast channel providing discrete spatiotemporal information regarding the ball and player movement, and would likely use an augmented broadcast channel similar to our system if improvements were made to the design **(RQ 2)**. Given the limitations of the de facto system (radio), our attempt to improve the quality of life for participants, however small it may seem, was met with much praise from our participants.

Of upmost importance to our participants was the need to be on a level spectating field as their sighted peers. Although participants enjoyed the benefits of a social context with respect to bespoke commentary, at the same time there is a strong desire for independence **(RQ 1)**. They want to be able to enjoy the SMB with others: this requires a level of independent consumption so they can feel part of the group and engage in social discussions about the SMB while simultaneously consuming it. Participants said they would follow up post-match statistics to ensure they understood the match outcome correctly, enabling them to engage in social conversation with sighted peers. Delivering real time information to blind spectators in a format they can understand will help elevate them to the same level as their peers, and include them in the discussion so they may enjoy SMB.

The results of our questionnaire and evaluations support this, indicating how blind spectators want to experience what sighted people experience. We would like to finish with the following quote which captures this succintly:

> "I would like to hear the details that sighted people can see." (P1)

8 Limitations

We could only use two blind spectators for the pilot studies. We could not do a comprehensive pilot study due to logistic reasons. Our initial attempts with pilot participants to do a remote study overwhelmed them [3]. We acknowledge that this imposed limitations on the development of the prototype. However, we believe that our iterative development that includes multiple evaluations provided a reasonable final prototype. We also did not investigate sustained consumption of the SMB. We would like to identify the features of the best possible prototype system before further investigations. We believe this type of consumption requires more control over the augmentations added to the original broadcasts by the users. We listed supporting social consumption as one of the main motivations. Though pursuing this motivation in a formal study was out of scope for our work here, our prototype system, which fulfills the requirements of effective conveyance of information, serves as the prerequisite for future social consumption studies. SMB consumption is a complex experience, with many factors driving people towards consuming content. In this early stage we have focused on developing a system that is up to the challenge of delivering accurate information to create a good experience for spectators. We acknowledge that more elaborate studies are needed to investigate factors such as color commentary engagement, social spectating, and why people prefer one presentation condition over another.

9 Conclusion

Our main contribution is providing insights into how radio broadcasting falls short of presenting a sport event to blind spectators. We have presented our work exploring how blind spectators approach sports media broadcast content. We built a prototype system for augmenting SMB content to enhance the spectator experience of blind and low vision people, and evaluated our prototype in a controlled user evaluation. We also conducted an online questionnaire to understand the social and technological context in which blind people consume SMB. Results showed a contrast between qualitative and quantitative responses in peoples' preference regarding augmented broadcast channels and traditional broadcast commentary. Though participants would repeatedly tout the benefit of radio over other ways to present information in the sport, they were no better at answering basic questions regarding the outcome of a game when listening to radio compared to augmented broadcast channels. Care must be taken when designing augmentations so as not to interfere with other aspects of the SMB itself and the external social context in which SMB spectating takes place. We hope our work raises a new design challenge for HCI, and provides principles to designers, engineers, researchers, and broadcasters in the HCI community producing more inclusive broadcast material.

Acknowledgments

Cagatay Goncu is supported by the Australian Research Council (ARC) grant DE180100057, and thanks the Faculty of Information Technology at Monash University for their support. Daniel J. Finnegan thanks the School of Computer Science & Informatics at Cardiff University for their continued support. We also thank Tennis Australia, particularly the Game Insight Group members Machar Reid, Touqeer Ahson and Stephanie Kovalchik for their support, and finally Kate Stephens for providing insightful feedback during development.

References

1. Berger, C.C., Gonzalez-Franco, M., Tajadura-Jiménez, A., Florencio, D., Zhang, Z.: Generic HRTFs May be Good Enough in Virtual Reality. Improving Source Localization through Cross-Modal Plasticity. Frontiers in Neuroscience **12** (2018). https://doi.org/10.3389/fnins.2018.00021
2. Bielli, S., Harris, C.G.: A mobile augmented reality system to enhance live sporting events. In: Proceedings of the 6th Augmented Human International Conference. pp. 141–144. AH '15, Association for Computing Machinery, New York, NY, USA (Mar 2015). https://doi.org/10.1145/2735711.2735836
3. Brinkley, J., Huff, E.W., Boateng, K.: Tough but Effective: Exploring the use of Remote Participatory Design in an Inclusive Design Course Through Student Reflections. In: Proceedings of the 52nd ACM Technical Symposium on Computer Science Education. pp. 170–176. SIGCSE '21, Association for Computing Machinery, New York, NY, USA (Mar 2021). https://doi.org/10.1145/3408877.3432527
4. Cherry, E.C.: Some Experiments on the Recognition of Speech, with One and with Two Ears (1953). https://doi.org/10.1121/1.1907229
5. Cheung, G., Huang, J.: Starcraft from the stands: understanding the game spectator. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 763–772. CHI '11, Association for Computing Machinery, New York, NY, USA (May 2011). https://doi.org/10.1145/1978942.1979053
6. Cobos, M., Lopez, J., Spors, S.: A Sparsity-Based Approach to 3D Binaural Sound Synthesis Using Time-Frequency Array Processing. EURASIP Journal on Advances in Signal Processing **2010**(1), 415840 (Sep 2010). https://doi.org/10.1155/2010/415840
7. Dong, M., Wang, H., Guo, R.: Towards understanding the differences of using 3D auditory feedback in virtual environments between people with and without visual impairments. In: 2017 IEEE 3rd VR Workshop on Sonic Interactions for Virtual Environments (SIVE). pp. 1–5 (Mar 2017). https://doi.org/10.1109/SIVE.2017.7901608
8. Donker, H., Klante, P., Gorny, P.: The design of auditory user interfaces for blind users. In: Proceedings of the second Nordic conference on Human-computer interaction. pp. 149–156 (2002). https://doi.org/10.1145/572020.572038
9. Eastman, S.T., Land, A.M.: The Best of Both Worlds: Sports Fans Find Good Seats at the Bar. Journal of Sport and Social Issues **21**(2), 156–178 (May 1997). https://doi.org/10.1177/019372397021002004, publisher: SAGE Publications Inc
10. Ferrand, S., Alouges, F., Aussal, M.: An electronic travel aid device to help blind people playing sport. IEEE Instrumentation Measurement Magazine **23**(4), 14–21 (Jun 2020). https://doi.org/10.1109/MIM.2020.9126047, conference Name: IEEE Instrumentation Measurement Magazine
11. Francombe, J., Brookes, T., Mason, R.: Evaluation of Spatial Audio Reproduction Methods (Part 1): Elicitation of Perceptual Differences. Journal of the Audio Engineering Society **65**(3), 198–211 (Mar 2017). https://doi.org/10.17743/jaes.2016.0070
12. Funakoshi, R., Okudera, Y., Koike, H.: Synthesizing Pseudo Straight View from A Spinning Camera Ball. In: Proceedings of the 7th Augmented Human International Conference 2016. pp. 1–4. AH '16, Association for Computing Machinery, New York, NY, USA (Feb 2016). https://doi.org/10.1145/2875194.2875236
13. Goldschmied, N., Sheptock, M., Kim, K., Galily, Y.: Appraising Loftus and Palmer (1974) Post-Event Information versus Concurrent Commentary in the Context

of Sport. Quarterly Journal of Experimental Psychology **70**(11), 2347–2356 (Nov 2017). https://doi.org/10.1080/17470218.2016.1237980, publisher: SAGE Publications

14. Gregory Appelbaum, L., Cain, M.S., Darling, E.F., Stanton, S.J., Nguyen, M.T., Mitroff, S.R.: What is the identity of a sports spectator? Personality and Individual Differences **52**(3), 422–427 (Feb 2012). https://doi.org/10.1016/j.paid.2011.10.048
15. Guerreiro, J.: Towards screen readers with concurrent speech: where to go next? ACM SIGACCESS Accessibility and Computing (115), 12–19 (2016). https://doi.org/10.1145/2961108.2961110, publisher: ACM New York, NY, USA
16. Guerreiro, J., Gonçalves, D.: Scanning for digital content: How blind and sighted people perceive concurrent speech. ACM Transactions on Accessible Computing (TACCESS) **8**(1), 1–28 (2016). https://doi.org/10.1145/2822910, publisher: ACM New York, NY, USA
17. Hermann, T., Zehe, S.: Sonified Aerobics - Interactive Sonification of Coordinated Body Movements (Jun 2011), `https://smartech.gatech.edu/handle/1853/51764`, accepted: 2014-05-21T16:35:15Z Publisher: International Community for Auditory Display
18. Huggins, M.: BBC Radio and Sport 1922–39. Contemporary British History **21**(4), 491–515 (Dec 2007). https://doi.org/10.1080/13619460601060512, publisher: Routledge
19. Ichiki, M., Shimizu, T., Imai, A., Takagi, T., Iwabuchi, M., Kurihara, K., Miyazaki, T., Kumano, T., Kaneko, H., Sato, S., Seiyama, N., Yamanouchi, Y., Sumiyoshi, H.: Study on Automated Audio Descriptions Overlapping Live Television Commentary. In: Computers Helping People with Special Needs. pp. 220–224. Springer International Publishing (2018), `https://link.springer.com/chapter/10.1007%2F978-3-319-94277-3_36`
20. Kamel, H.M., Landay, J.A.: A study of blind drawing practice: creating graphical information without the visual channel. In: Proceedings of the fourth international ACM conference on Assistive technologies. pp. 34–41. Assets '00, Association for Computing Machinery, New York, NY, USA (Nov 2000). https://doi.org/10.1145/354324.354334
21. Kim, J., Kim, Y., Kim, D.: Improving well-being through hedonic, eudaimonic, and social needs fulfillment in sport media consumption. Sport Management Review **20**(3), 309–321 (Jun 2017). https://doi.org/10.1016/j.smr.2016.10.001
22. Kim, S., Lee, K.p., Nam, T.J.: Sonic-Badminton: Audio-Augmented Badminton Game for Blind People. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. pp. 1922–1929. CHI EA '16, Association for Computing Machinery, New York, NY, USA (May 2016). https://doi.org/10.1145/2851581.2892510
23. Kobayashi, M., O'Connell, T., Gould, B., Takagi, H., Asakawa, C.: Are synthesized video descriptions acceptable? In: Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility. pp. 163–170. ASSETS '10, Association for Computing Machinery, New York, NY, USA (Oct 2010). https://doi.org/10.1145/1878803.1878833
24. Kumano, T., Ichiki, M., Kurihara, K., Kaneko, H., Komori, T., Shimizu, T., Seiyama, N., Imai, A., Sumiyoshi, H., Takagi, T.: Generation of Automated Sports Commentary from Live Sports Data. In: 2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). pp. 1–4 (Jun 2019). https://doi.org/10.1109/BMSB47279.2019.8971879

25. Kwon, Y.S., Lee, S.E.: Cognitive Processing of Sound Effects in Television Sports Broadcasting. Journal of Radio & Audio Media **27**(1), 93–118 (Jan 2020). https://doi.org/10.1080/19376529.2018.1541899
26. Lee, G., Bulitko, V., Ludvig, E.A.: Automated Story Selection for Color Commentary in Sports. IEEE Transactions on Computational Intelligence and AI in Games **6**(2), 144–155 (Jun 2014). https://doi.org/10.1109/TCIAIG.2013.2275199, conference Name: IEEE Transactions on Computational Intelligence and AI in Games
27. Lee, M., Kim, D., Williams, A.S., Pedersen, P.M.: Investigating the Role of Sports Commentary: An Analysis of Media-Consumption Behavior and Programmatic Quality and Satisfaction. Journal of Sports Media **11**(1), 145–167 (Jul 2016). https://doi.org/10.1353/jsm.2016.0001, `https://muse.jhu.edu/article/626349`, publisher: University of Nebraska Press
28. Lessel, P., Vielhauer, A., Krüger, A.: 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. pp. 1571–1576. CHI '17, Association for Computing Machinery, New York, NY, USA (May 2017). https://doi.org/10.1145/3025453.3025708
29. Lopez, M., Kearney, G., Hofstädter, K.: Audio Description in the UK: What works, what doesn't, and understanding the need for personalising access. British Journal of Visual Impairment **36**(3), 274–291 (Sep 2018). https://doi.org/10.1177/0264619618794750, `https://doi.org/10.1177/0264619618794750`, publisher: SAGE Publications Ltd
30. Moos, A., Trouvain, J.: Comprehension of Ultra-Fast Speech–Blind vs. 'Normally Hearing' People p. 4 (2007), `http://www.icphs2007.de/conference/Papers/1186/1186.pdf`
31. Petrie, H., Hamilton, F., King, N., Pavan, P.: Remote usability evaluations with disabled people. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1133–1141. CHI '06, Association for Computing Machinery, New York, NY, USA (Apr 2006). https://doi.org/10.1145/1124772.1124942
32. Pettitt, B., Sharpe, K., Cooper, S.: AUDETEL: Enhancing television for visually impaired people. British Journal of Visual Impairment **14**(2), 48–52 (May 1996). https://doi.org/10.1177/026461969601400202, `https://doi.org/10.1177/026461969601400202`, publisher: SAGE Publications Ltd
33. Pizzo, A.D., Na, S., Baker, B.J., Lee, M.A., Kim, D., Funk, D.C.: eSport vs. Sport: A Comparison of Spectator Motives. Sport Marketing Quarterly **27**(2) (2018)
34. Popovici, I., Vatavu, R.D.: Towards Visual Augmentation of the Television Watching Experience: Manifesto and Agenda. In: Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video. pp. 199–204. TVX '19, Association for Computing Machinery, New York, NY, USA (Jun 2019). https://doi.org/10.1145/3317697.3325121, `http://doi.org/10.1145/3317697.3325121`
35. Rafey, R.A., Gibbs, S., Hoch, M., Le Van Gong, H., Wang, S.: Enabling custom enhancements in digital sports broadcasts. In: Proceedings of the sixth international conference on 3D Web technology. pp. 101–107. Web3D '01, Association for Computing Machinery, New York, NY, USA (Feb 2001). https://doi.org/10.1145/363361.363384, `https://doi.org/10.1145/363361.363384`
36. Ramos, M.: The emotional experience of films: does Audio Description make a difference? The Translator **21**(1), 68–94 (Jan 2015). https://doi.org/10.1080/13556509.2014.994853

37. Reeves, S., Benford, S., O'Malley, C., Fraser, M.: Designing the spectator experience (2005). https://doi.org/10.1145/1054972.1055074
38. Saidi, N.L.: Sound guided football/basketball game for blind people (Oct 2007), https://patents.google.com/patent/US20070238557A1/en
39. Savery, R., Ayyagari, M., May, K., Walker, B.N.: Soccer sonification: Enhancing viewer experience. Georgia Institute of Technology (2019), http://hdl.handle.net/1853/61512
40. Shinohara, K., Tenenberg, J.: A blind person's interactions with technology. Communications of the ACM **52**(8), 58–66 (Aug 2009). https://doi.org/10.1145/1536616.1536636, https://doi.org/10.1145/1536616.1536636
41. Spelke, E., Hirst, W., Neisser, U.: Skills of divided attention. Cognition **4**(3), 215–230 (Jan 1976). https://doi.org/10.1016/0010-0277(76)90018-4
42. Swaminathan, M., Pareddy, S., Sawant, T.S., Agarwal, S.: Video Gaming for the Vision Impaired. In: Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. pp. 465–467. ASSETS '18, Association for Computing Machinery, New York, NY, USA (Oct 2018). https://doi.org/10.1145/3234695.3241025
43. Toupal, R., Schmid, D.: System and method for automatically generating a narrative report of an event, such as a sporting event (Dec 2005), https://patents.google.com/patent/US6976031B1/en
44. Trail, G.T., Kim, Y.K.: Factors influencing spectator sports consumption: NCAA women's college basketball. International Journal of Sports Marketing & Sponsorship **13**(1) (2011)
45. Walker, B.N., Lindsay, J., Nance, A., Nakano, Y., Palladino, D.K., Dingler, T., Jeon, M.: Spearcons (Speech-Based Earcons) Improve Navigation Performance in Advanced Auditory Menus. Human Factors **55**(1), 157–182 (Feb 2013). https://doi.org/10.1177/0018720812450587, publisher: SAGE Publications Inc
46. Wedoff, R., Ball, L., Wang, A., Khoo, Y.X., Lieberman, L., Rector, K.: Virtual Showdown: An Accessible Virtual Reality Game with Scaffolds for Youth with Visual Impairments. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–15. CHI '19, Association for Computing Machinery, New York, NY, USA (May 2019). https://doi.org/10.1145/3290605.3300371
47. Zillmann, D., Bryant, J., Sapolsky, B.S.: Enjoyment from sports spectatorship. Sports, games, and play: Social and psychological viewpoints **2**, 241–278 (1989), https://www.taylorfrancis.com/books/e/9780203728376/chapters/10.4324/9780203728376-14
48. Zollmann, S., Langlotz, T., Loos, M., Lo, W.H., Baker, L.: ARSpectator: Exploring Augmented Reality for Sport Events. In: SIGGRAPH Asia 2019 Technical Briefs. pp. 75–78. SA '19, Association for Computing Machinery, New York, NY, USA (Nov 2019). https://doi.org/10.1145/3355088.3365162