VStroll: An Audio-based Virtual Exploration to Encourage Walking among People with Vision Impairments

Gesu India*
gesuindia@gmail.com
Microsoft Research
Bangalore, India

Mohit Jain*
mohja@microsoft.com
Microsoft Research
Bangalore, India

Pallav Karya
pallavkaria@gmail.com
Microsoft Research
Bangalore, India

Nirmalendu Diwakar
nirmalendu.diwakar@gmail.com
Microsoft Research
Bangalore, India

Manohar Swaminathan
manohar.swaminathan@microsoft.com
Microsoft Research
Bangalore, India

ABSTRACT

Current infrastructure design, discouragement by parents, and lack of internal motivation act as barriers for people with visual impairments (PVIs) to perform physical activities at par with sighted individuals. This has triggered accessible exercise technologies to be an emerging area of research. However, most current solutions have either safety concerns and/or are expensive, hence limiting their mass adoption. In our work, we propose *VStroll*, a smartphone app to promote walking among PVIs, by enabling them to virtually explore real-world locations, while physically walking in the safety and comfort of their homes. Walking is a cheap, accessible, and a common physical activity for people with blindness. VStroll has several added features, such as places-of-interest (POI) announcement using spatial audio and voice input for route selection at every intersection, which helps the user to gain spatial awareness while walking. To understand the usability of VStroll, 16 participants used our app for five days, followed by a semi-structured interview. Overall, our participants took 253 trips, walked for 50.8 hours covering 121.6 kms. We uncovered novel insights, such as discovering new POIs and fitness-related updates acted as key motivators, route selection boosted their confidence in navigation, and spatial audio resulted in an immersive experience. We conclude the paper with key lessons learned to promote accessible exercise technologies.

CCS CONCEPTS

• **Human-centered computing → Accessibility technologies**; *Empirical studies in HCI*.

KEYWORDS

accessibility, blind, smartphone, exercise, maps, point of interest, spatial audio, navigation, walk

*These two authors contributed equally.

ASSETS '21, October 18–22, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8306-6/21/10…$15.00
https://doi.org/10.1145/3441852.3471206

ACM Reference Format:
Gesu India*, Mohit Jain*, Pallav Karya, Nirmalendu Diwakar, and Manohar Swaminathan. 2021. VStroll: An Audio-based Virtual Exploration to Encourage Walking among People with Vision Impairments. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21), October 18–22, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3441852.3471206

1 INTRODUCTION

People with vision impairments (PVIs) engage in less physical exercises, and therefore have inferior physical fitness compared to sighted adults [5, 26, 27, 54]. On analyzing National Health Interview Survey data, Weil *et al.* found that the odds of being obese or overweight for persons with vision impairments were 1.5 times greater than sighted individuals [50]. Several studies have been conducted to understand concerns, motivations, barriers, and experiences with exercise in the blind population, and also understand the role of technology in encouraging exercise among them [15, 39, 47]. PVIs face several barriers to participate in physical exercises, including inaccessible fitness infrastructures, discouragement by parents due to safety concerns, limited social opportunities to exercise with a partner, lack of experience and interest, and misperceptions related to the importance of exercise [15, 27, 29, 39, 40, 44, 47].

This has triggered accessible exercise technologies to be an emerging area of research in the HCI and accessibility community [38, 57, 58]. To promote physical activities, several sports have been invented and/or made accessible for people with blindness, such as goalball [8], beep baseball, and blind cricket. All these sports use a combination of audio and vibrotactile-based stimuli to replace the visual stimuli. However, participation in these sports is very limited, mainly due to safety concerns and perceived requirement of athletic abilities and competence [47].

Exergames can mitigate these barriers, as they combine playful video gaming with exercise and can be played at the comfort of home, thus motivating PVIs to engage in more exercising. Several accessible exergames have been developed, which are either accessible versions of famous exergames for sighted individuals (*e.g.*, VI-Tennis [32], Blind Hero [57]), or exergames specifically designed for PVIs (*e.g.*, Eyes-Free Yoga [38], Pet-N-Punch [34]). However, these exergames usually require expensive hardware and/or accessible infrastructure, thus limiting their adoption.

Prior work found walking to be the most common physical activity among people with blindness [15, 39], as walking neither requires learning any special skills, nor requires a gym membership or expensive equipment. The Centers for Disease Control and Prevention recommends 150 minutes/week of brisk walking (*i.e.*, 20-30 minutes/day) for adults to meet their physical exercise needs [6, 46]. However, poor road infrastructures, lack of walking partners, and most importantly, the monotonous nature of walking, act as hindrances to pursue walking on a regular basis [15, 39].

To solve these issues, we propose *VStroll*, a smartphone app to promote walking among PVIs, by enabling them to virtually explore a real-world location while walking. The user can choose any location and start exploring it virtually, listening to spatial information about points-of-interest (POIs), while they physically walk in the safety and comfort of their homes. Features such as POI announcements using binary spatial audio, voice input for road selection at every intersection, and regular updates related to fitness parameters (like calories burned, distance traveled) were added to keep the participants engaged. Accruing spatial information of POIs have been found to help people with blindness in navigation and building mental map of the real-world [14, 22, 55]. Assistive technology based solutions have been proposed to provide PVIs with spatial information while traveling in physical world (*e.g.*, Nearby Explorer, Lazarillo app), or to virtually plan their navigation in advance using turn-by-turn instructions and POI announcements [14, 55]. However, none of these systems enable the user to *freely explore* a real-world location in a virtual mode.

To understand the usability of VStroll, 16 participants used the app for five days, followed by a post-usage semi-structured interview. Overall, our participants took 253 trips, walked for 50.8 hours covering 121.6 kms. We found novel insights, including spatial audio resulting in immersive experience, discovering new POIs and fitness-related updates as key motivators, road selection providing autonomy to boost navigation confidence, nostalgia driving virtual visits to childhood locations, increased spatial knowledge, and comparison of VStroll with various fitness, entertainment and navigation apps. We conclude the paper with lessons learned to develop solutions that can enable and motivate exercising among the blind population, with respect to gamification, prominence-based POI, and privacy concerns.

2 RELATED WORK

Our work is mainly informed by two areas of relevant research: motivate exercising and gaining spatial information among the blind population. Below we discuss prior works in these areas, with a note that none of them attempt to combine motivating physical exercise by enabling virtual exploration of the real-world, which is our key contribution.

2.1 Physical Exercising among PVIs

Exercise requirements, barriers, and ways to motivate exercising among people with visual impairments have been widely discussed [5, 15, 39, 47, 50]. Factors such as the state of infrastructure, psychosocial barriers, and physical constraints disable people with visual impairments to perform exercises at par with the sighted population [15, 25]. Although infrastructure barriers can be handled by inclusive design, internal motivation still remains one of the most critical hurdles in engaging with physical activities [10, 47]. Exclusion from the inaccessible constructs of society, like physical education classes and gyms, in addition to little support at home to be more physically active contribute to PVIs perceived lack of competence to do exercise, which in turn has a detrimental effect on their motivation [24, 30, 47].

Several methods have been explored to motivate PVIs to participate in physical exercises. Common sports for sighted users have been adapted for PVIs using auditory and haptic feedback, *e.g.*, blind cricket, beep baseball, etc. Technologies have also been proposed to enhance player's performance, such as Goby for swimming [37] and Running Guide [58]. Goby uses a downward-facing camera to track the swimmer's position and provides audio feedback, while Running Guide uses wearable glasses embedded with GPS sensors to help runners track the route by receiving spatial 3D auditory and vibrotactile feedback. Sports targeted towards blind players have also been invented, such as Goalball [8] and Polybat [53]. Due to fear of injury, parents have been found to be apprehensive of their children playing such sports [47].

Thus, more recently, accessible exergames have been developed, as they are safe and can be played indoors. Exergames combine an immersive video gaming experience with physical exertion [43]. Accessible exergames are either accessible versions of existing exergames [32, 33, 35, 57], or exergames developed originally for PVIs [34, 38]. For instance, VI-Tennis [32], VI-Bowling [33], Skiing [35], and Blind Hero [57], are accessible alternatives of Wii Tennis, Wii Bowling, Wii Ski Slalom, and Guitar Hero, respectively. In VI-Tennis [32], players receive audio feedback notifying the pace and distance of an approaching ball, and vibrotactile feedback (on Wii remote) indicating the best time to hit the ball; Blind Hero [57] uses a glove with motors attached to each finger-tip, to transform the rhythm-related visual information of Guitar Hero into tactile feedback. To generalize, Morelli *et al.* [36] proposed a technique to make any existing exergame accessible, without changing the source code. The proposed solution uses real-time video analysis for detection and substitution of visual cues with vibrotactile cues, and game play using whole body gestures on Microsoft Kinect.

Novel exergames have been developed specifically for PVIs [34, 38], relying on audio and/or tactile stimuli. Pet-N-Punch [34] utilizes a Whac-A-Mole style game play using two Wii remotes, wherein the player receives audio feedback while searching for rodents, and tactile feedback on hitting rodents. Similarly, Rector *et al.* created Eyes-Free Yoga [38] to make yoga exercises accessible for PVIs, using Microsoft Kinect's Skeletal Tracking to detect player's yoga poses and provide auditory feedback similar to a yoga instructor. All these accessible exergames require expensive gaming hardware, which hinders their widespread adoption.

On the other hand, walking is a cheap accessible exercise; 20-30 minutes of daily walking meets the physical needs of the human body [6]. Though walking is a common physical activity for people with blindness [15, 39], it is not practised regularly due to a variety of reasons, such as poor road infrastructures, lack of walking partners, and its monotonous nature [15, 39]. As a solution, prior work recommends gamified walking, for instance, by incorporating a talking pedometer for PVIs [12, 28]. A wearable talking pedometer showed positive effect on 22 children with vision impairments, as they set challenging goals to increase their daily physical activity levels, thus changing their walking behaviour [28]. Similarly, wearables (*e.g.*, Fitbit, Nike FuelBand) and smartphone apps (*e.g.*, Step Tracker) tracking physical activities have shown positive impact on sighted individuals [13], as access to fitness parameters—step count, calories burned, pulse rate, sleep cycle, *etc*—have been found to encourage users.

In this work, we build upon the existing literature to develop a solution to motivate PVIs to walk by providing them access to fitness parameters, along with a novel added component of virtual exploration while walking.

2.2 Spatial Learning among the Blind Population

In the absence of visual feedback, PVIs learn about their physical environment by gathering spatial information (related to navigation routes, landmarks, and points-of-interest) to develop a mental map of the physical space [14, 31, 49]. Several assistive technologies have been developed aiming to help PVIs by providing them spatial information: (a) while navigating in the physical world, or (b) using virtual navigation.

Smartphone navigation apps (such as Google Maps, Bing Maps) assist PVIs using audio-based turn-by-turn navigation instructions. However, they do not provide information about nearby places-of-interest (POIs). Awareness about nearby POIs can help PVIs in developing a better spatial understanding of a location [14]. To fill this gap, smartphone apps, such as Lazarillo, Nearby Explorer, *etc*., provide accessible updates about POIs within a fixed radius around the user, thus helping the user to locate specific POIs, orient themselves, and navigate. Similarly, Microsoft's Soundscape [1] uses spatial 3D audio cues to inform about nearby POIs. While these technologies enable users to explore their current physical location, they do not allow virtual exploration.

Another set of technologies focus on enabling users to navigate virtually and gather spatial awareness about any physical location. They help PVIs to plan their navigation in advance, thus boosting their confidence for a physical visit [14, 55]. Tactile maps are an effective way to learn spatial relationships [31, 51]. However, printing them is expensive, and they are neither interactive nor provide up-to-date information [42]. Yatani *et al.* [55] built SpaceSense, which provides turn-by-turn instructions to virtually navigate from a chosen start location to any destination, along with tactile feedback conveying direction information of the destination or any POIs along the route. Similarly, VirtualLeap and VirtualWalk [14] are two smartphone-based virtual navigation systems that provides turn-by-turn navigation and announces enroute POIs information (location, orientation, *etc.*). VirtualLeap allows the user to jump through turn-by-turn instructions, POIs, and road intersections, while VirtualWalk simulates step-by-step walking at variable speed. These systems [11, 14, 55] were found to be helpful in creating an accurate mental representation of the route, including the locations of the POIs. None of these systems allow *free* exploration, as they require a start location and destination as input, and provide guided step-by-step navigation as output.

Prior work enabling free virtual exploration relied heavily on visual cues, thus making them inaccessible for the blind. Shih [45] created a 3D virtual environment integrated with Google Street View for exploring London. Christou *et al.* [7] and Roes *et al.* [41] enabled virtual exploration of cultural heritage sites and museums, respectively.

In contrast, VStroll emulates true virtual exploration using a smartphone, wherein the user enters a start location and decides the navigation route as she/he walks, by telling at each intersection which road to take next.

3 VSTROLL DESIGN

We decided to develop the VStroll app on the Android platform, primarily because Android has >70% of the worldwide market share [2]. To achieve an ideal virtual exploration, enabling a user to start the trip at any chosen physical location and stroll around virtually at will, we designed the following system.

3.1 User Interface

The VStroll app consists of multiple screens, including Start Location Selection, Ongoing Trip, History, Settings, and Help (Figure 1). Below are the key aspects of VStroll.

Calibration: On launching the app for the first time, it guides the user to calibrate the sensitivity of step detection. To do so, the user needs to click the 'Start Calibration' button, put the phone in a (bottomwear) pocket, walk for 15 steps, and then click the 'Calibrate' button. The computed sensitivity value is stored in the phone, and remains consistent for the subsequent app launches. While using VStroll, if the user feels that the steps are getting counted incorrectly, she/he can re-calibrate anytime from the Settings screen.

Start Location: After calibration, the user needs to input the start location (Figure 1a). The start location text box has the Google Maps auto-complete feature enabled to minimize typing. By default, the user's current location is the start location. The entered start location is retained between app usage, *i.e.*, from the second trip onwards, the default start location is the chosen start location of the previous trip. After selecting a start location, the user needs to click the 'Start' button (Figure 1a). It guides the user to put the phone in a pocket and start walking, and moves to the Ongoing Trip screen (Figure 1b). In the back-end, the user is placed virtually on the nearest road from the entered start location, and the user starts walking on the left/right side of the road, depending on the user's country driving side. The app informs the user, '*You are walking on X road towards Y intersection*'. Whenever walking steps get detected, a background wind sound is played to provide auditory feedback. We experimented with various background sounds in our pilot study, including the sound of foot steps. We found our participants altering their usual walking pace to match their steps with the step sound. To avoid this, the wind sound was selected, which can be disabled from the Settings screen.

POI: On crossing any point-of-interest (POI, as in restaurant, shop, office, bank, school, historical site, place of worship, *etc.*), certain details of the POI gets announced, *e.g.*, '*X restaurant serving Y cuisine*', '*X, a Y shop*', '*X bank ATM*', '*X temple worshipping Y god*' (Figure 1b). POIs within 60 meters of the user's virtual location get announced. The POIs on the OpenStreetMaps are marked at the center of their location rather than along the edge of the road. With trial-and-error, a radius of 60 meters was determined, which allowed

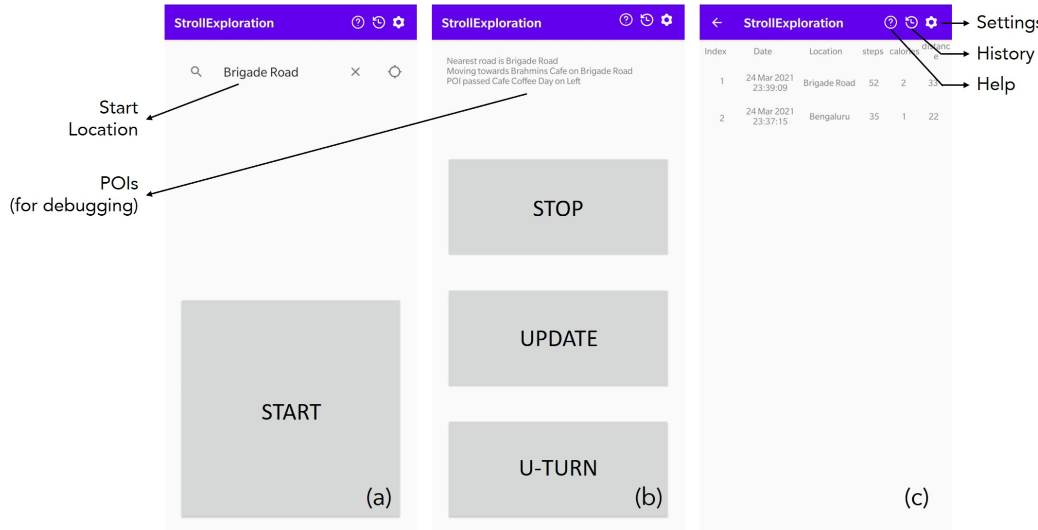

Figure 1: User Interface of the VStroll app: (a) Start Location Selection screen, (b) Ongoing Trip screen, (c) History screen.

our algorithm to successfully cover POIs located on both sides of the road in most cases. A POI announcement is preceded by a 'beep' sound to help users distinguish between the announcements of multiple (two or more) closely spaced POIs. The POIs are played using binary spatial audio, such that the announcement is audible only in left/right ear, in order to make the user spatially aware that the POI is on the left/right side of the road. (Note: Except POI, all other information, instruction and fitness-related audio gets played equally in both left and right audio channels.) Prior research has shown that PVIs at times prefer wearing headphones in only one ear, keeping the other ear free to listen to environment sounds [19, 23]. Using Settings, the user can disable spatial audio, which instead adds '*to your left*' or '*to your right*' to each POI announcement (similar to [14]).

We experimented with a full spatial audio engine (similar to [1]) to provide both orientation and distance audio cues. In our initial pilot study, our participants complained about information overload, and were found struggling to listen to farther POIs. Thus we decided to use a simplified left-right binary spatial audio. Moreover, we (arbitrarily) used male voice for POI announcements and female voice for the rest of the announcements, to separate POIs audio from the rest.

If the nearest POI is more than 100 meters away, VStroll provides *filler information* (like '*You are on X street*', '*You have walked X steps*', '*You are X meters away from Y intersection*', '*You have burned Y calories*') to avoid long periods of silence. On the other hand, if a road has several POIs, a few might get missed, *i.e.*, if the user passes another POI during an ongoing POI announcement, the new POI is not announced, instead the system plays a '*tuck*' (missed POI) sound to notify the user. If no steps get detected for 5 sec or more, a periodic message, '*please continue walking*', is played nudging the user to move. Other *filler information* gets announced as well, till the user resumes walking.

Intersection: On reaching an intersection of three or more roads, VStroll asks the user to choose a direction, *e.g.*, '*You have reached an intersection with three roads, please speak your choice after the beep... left on X road, U-Turn on Y road, right towards Z POI.* (beep sound)'. Many roads are *unnamed* on OpenStreetMap, in such cases, we use the nearest POI on that road as a marker. After the beep sound, the app waits for the user's voice input. We decided to use voice-based input to minimize the need of taking the phone out of the pocket for interaction while walking. Moreover, PVIs have been found to often use voice as input [4, 19]. The user needs to speak one of the keywords, *i.e.*, left, U-Turn, right, slight left, *etc.* (Note: The user does not have to change their walking direction in the physical world). In case of wrong/no input, the app provides one more chance by repeating the road selection message and beep sound. After that, it randomly auto-selects one of the roads.

Miscellaneous: The Ongoing Trip screen (Figure 1b) has three buttons: (a) Stop: it stops the current trip and takes the user to the Start Location Selection screen, (b) Update: it provides fitness-related information about the ongoing trip, such as number of steps taken, calories burned, and distance travelled, and (c) U-Turn: it makes the virtual user walk in the opposite direction. Details about the trip get saved and can be accessed on the History screen (Figure 1c), which shows date, start location, trip duration, number of steps, calories burned, and distance walked, for each trip. The Settings screen allows (re-)calibration, enabling/disabling directional audio, and enabling/disabling background sound.

3.2 Implementation Details

We used a combination of Google Maps (for auto-complete start location) and OpenStreetMaps (OSM, for POIs and road graph). OSM Overpass API provides access to a geographic area as a graph using nodes for intersections and edges for streets. While walking, if the user reaches close to the boundary of the accessed region in the virtual world, the app dynamically fetches adjacent region(s) from OSM. Google Maps API also offers similar functionalities (like POI, road graph), however in our pilot study, we found Google Maps to be very expensive, hence we switched to OSM. On the other hand, OSM has a few disadvantages, such as several roads are *unnamed*, POIs are sparser than Google Maps, *etc.*

For step detection, we relied on the smartphone's in-built accelerometer and gyroscope sensor data, as the step counter sensor is supported by Android API level 19 and above, which was not available on a few of our participant's phones. VStroll assumes a standard fixed stride length of 69 cm (similar to [14]), and based on the number of detected steps, it advances the user's position on the virtual map. The app was made accessible by TalkBack prompts announced using the Google TTS (Text-to-Speech) engine. All input and output events (including steps, POI announcements, voice inputs and user clicks) were logged and saved on the server for analysis.

4 STUDY DESIGN

To evaluate the usability of VStroll, we conducted an IRB-approved user study during Jan-Feb 2021 in India.

4.1 Procedure

Our study required signing a consent form, installing the VStroll app, receiving app usage related training, and a semi-structured interview after five days of app usage — all of which were conducted online, due to the COVID-19 induced restricted mobility.

Two NGOs working with PVIs in India helped us to recruit the participants, by circulating the recruitment email in their respective networks and WhatsApp groups. Participation criteria was that the individual should be above 18 years with vision impairments, have been using an Android smartphone with the help of TalkBack screen reader for two or more years, have Internet connectivity, have a pair of working head/ear-phone (with microphone), and have no additional physical/mental disability. We emailed the consent form to people who responded to our recruitment message. After they accepted it, we conducted an hour-long online training session, wherein we gave them a web-link to download and install VStroll, provided app details in simple words ("*VStroll helps in exploring the surroundings of any outdoor location, while walking indoors*"), followed by hands-on training on using the app.

The training comprised of: (a) sensor sensitivity calibration for accurate estimation of step count, (b) app usage related instructions (including selection of start location, POI announcements using binary spatial audio, voice input for the next road selection at every intersection, and details about Update, Help, Settings, U-Turn and History button), and (c) taking a short trip to introduce these features in practice. The participants were encouraged to ask questions during the training. At the end of this session, we collected their demographic information, exercise and navigation habits, and technology usage for exercising and navigation. Participants were asked to use VStroll in and around their homes for 20-30 mins daily, for the next five days. They were instructed to always use stereo head/ear-phones with a microphone while using the app. One of the authors sent a WhatsApp message daily to each participant reminding them to use the app. After five-days of app usage, a post-study telephonic interview was conducted to discuss participants' experience and feedback of using VStroll. During the interview, we asked them about the reasons to visit specific locations (virtually), their motivation behind using the app, challenges faced, and suggestions to improve the app.

The training and interview sessions were conducted in Hindi for 5 participants, and in English for 11 participants. After receiving verbal consent from the participants, the interviews were audio-recorded, and were transcribed using Tactiq.io (a browser plug-in for automatically generating live transcription). The auto-generated transcriptions were manually verified and corrected by the interviewer. On average, the interview lasted for ~40 mins. At the end of the interview, participants were asked to uninstall the VStroll app. Participants were paid Rs 250 (~4 USD) for participation.

4.2 Participants

Sixteen blind participants (5 female) with average age of 29.6±10.1 years participated in the study (Table 1). Eleven of them were blind from birth, while the rest lost their eyesight later due to accident/disease; on average, our participants have been blind for 26.4±12.1 years. Only one participant had a graduate degree, twelve had undergraduate degrees, and three passed high school. A majority of our participants (11) were students, while the rest were full-time employees working as teacher, banker, accessibility tester, etc. All the participants have been using a touchscreen smartphone with the help of TalkBack for more than 3 years (6.1±2.3 years). All were typical smartphone users with access to stable internet connection, and reported using their phone for 5.8±2.1 hours daily for making calls, watching videos, and using WhatsApp extensively. Based on the smartphone cost, five participants owned high-range smartphones (>$300), ten owned mid-range phones ($150-$300), and one owned low-range phone (<$150). Twelve participants reported engaging in some physical activity (including walking, yoga, stretching) for three or more times a week. All of them used Google Maps frequently for navigation. Participants were from across eleven different cities of India.

4.3 Data Analysis

We conducted a mixed-method analysis to systematically analyze the data. The log files were quantitatively analyzed. As we had logged data from only sixteen participants, we emphasize on the qualitative interview data. We subjected our interview data to open coding and categorized our codes to understand user behaviour. Two authors participated in the coding process and iterated upon the codes until consensus was reached. Over the course of analysis, they discussed coding plans, developed preliminary codebooks, reviewed the codebook, refined/edited codes, and finalized categories and themes. The first-level codes were very specific, such as "walking in international locations" and "step counter in smartwatches". After several rounds of iteration, the codes were condensed into high-level themes, such as "motivation for walking", "spatial knowledge acquisition", and "comparison with other apps" (refer codebook in Appendix A).

5 RESULTS

On an average, participants used the VStroll app for 6.7±2.6 days, taking a total of 253 trips (15.8±12.1 trips/participant), by walking for 50.8 hours (12.0±13.5 mins/trip) covering a distance of 121.6 kms

Table 1: Participants demography, along with log data (days used, trips taken, step count, distance walked, and calories burned). [Education–B: Bachelor's, M: Master's, HS: High School; Occupation–E: Employee, S: Student; Phone Model–S: Samsung, X: Xiaomi]

P.Id	Sex	Age	City	Education	Occupation	Blind Years	Phone Model	Talk-Back Years	VStroll Usage		Total Trips	Step count	Distance walked(km)	Calories burned
									(days)	(hrs)				
1	F	42	Mumbai	B	E	42	Moto G 5Plus	6	6	0.82	14	2281	1.5	88
2	M	44	Bangalore	B	E	44	S Galaxy A50	7	5	0.75	7	623	0.4	23
3	M	23	Nagpur	HS	S	6	X Poco F1	6	4	5.9	16	12518	8	496
4	M	31	Mumbai	B	S	31	One Plus 6	12	6	2.46	11	11512	7.4	456
5	F	22	Hyderabad	B	S	15	One Plus 7T	6	10	6.89	15	24589	15.7	977
6	M	39	Pune	B	E	21	S Galaxy A50	4	9	2.45	22	5353	3.4	206
7	F	53	Mumbai	B	E	53	Moto G 5G Plus	10	6	8.13	15	41096	26.3	1637
8	M	22	Ambala	B	S	18	Realme 3 RMX1825	3	12	2.18	9	5857	3.7	232
9	F	26	Mumbai	M	S	26	MI Note Pro 6	7	6	3.08	8	5068	3.24	200
10	M	21	Surat	B	S	21	S Galaxy M20	6	6	5.37	55	11551	7.4	446
11	M	24	Kolkata	B	S	20	MI Note 5	4	4	3.15	11	13357	8.5	530
12	F	25	Kerela	B	S	25	X Note 7 Pro	7	5	0.26	6	1032	0.7	40
13	M	26	Bangalore	B	S	26	S Galaxy J7 Pro	12	5	0.65	10	2608	1.66	101
14	M	18	Fatehpur	HS	S	18	Redmi Y2	4	7	3.39	30	29567	18.9	1169
15	M	22	Delhi	HS	S	22	S Galaxy M21	4	5	1.03	14	2526	1.6	98
16	M	35	Mumbai	B	E	35	Samsung M31	6	5	4.22	10	20640	13.2	820
Avg.		29.6				26.4		6.1	6.7	3.2	15.8	11886.1.1	7.6	469.9
Stdev.		10.1				12.1		2.3	2.6	2.4	12.1	11669.8	8.83	464.7

(0.48±0.7 km/trip). Overall, they found VStroll's user interface easy-to-learn, intuitive, and accessible. Participants walked in real-world spaces while virtually exploring real-world locations using VStroll. A majority of our participants (13) reported using VStroll while walking indoors in their balcony, living room, terrace, *etc.*, three participants walked outdoors in nearby open spaces and parks, and three participants walked on a treadmill. Depending on the space, participants mentioned walking back-and-forth (*e.g.*, wall-to-wall in the living room), or in circles (*e.g.*, on the terrace). None of the participants reported confusing their real-world walking direction with virtual walking direction. Since participants were walking in familiar surroundings, none of them used a white cane with VStroll. Below, we report the key findings from our study.

5.1 Motivation for Walking

Prior work has shown that fitness-related regular updates have positive motivational effects on the walking behavior among PVIs and sighted users [13, 28]. Similarly, in our study, participants reported that real-time updates about fitness parameters (such as step count, calories burned, and distance walked) kept them informed of their progress and boosted their confidence, hence motivated them to walk more. Five participants mentioned feeling more conscious and responsible towards their health after using VStroll. For instance,

> "*Mentioning calories in the Update button was very motivating for me. It made me more attentive towards my health... Also it encouraged me to walk more, mainly to burn more calories.*" – P6.

Six participants felt empowered upon realising their potential to walk long distances.

> "*I like walking, I have always liked walking... I got to walk a lot due to this app... I never knew I could walk so much, that too at home.*" – P10.

In total, our participants took 190,178 steps, and burned 7,519 calories by walking 121.6 kms. Participants received these fitness updates by clicking the Update button (Figure 1b), or as filler announcements (when they stopped walking or when the next POI was more than 100 meters away). Overall, participants clicked the Update button 731 times (2.9±5.8 times/trip). To avoid physical interaction with the phone while walking, two participants decided not to access the Update button, and tracked their progress by relying on the filler announcements.

A majority of our participants (10) used these updates to set and monitor their daily fitness goals, often aiming for ambitious goals (similar to [13, 28]). For instance, "*I aimed for losing 120-150 calories everyday.*" – P16. This is interesting as VStroll does not explicitly support goal setting and tracking. Participants kept track of their daily fitness goal using the History screen (Figure 1c), by accessing fitness reports of all their virtual trips. However, the current History screen is primitive, and participants recommended adding multiple features, such as adding accessible charts summarizing fitness progress over days, weeks and months. For three participants, VStroll played the role of a fitness instructor, "*a walking coach*", as it closely monitors their progress and "*nudges*" them to walk.

> "*Whenever I stopped walking, the app asked me to walk... This app makes me walk!*" – P15.

Apart from the health benefits, spatial information provided as POI, road name, and intersection related announcements not only kept our participants engaged in the virtual trip, but also encouraged them to walk more to gain spatial understanding and discover new POIs, fueling their curiosity for "*what's coming next*". P7 compared finding interesting POIs to finding treasures in a treasure hunt game. This led seven participants to walk longer than initially planned.

> "*This app does promote walking! I went out that day and I was very exhausted, so I thought I'd just walk for 20 minutes, as it was late at night. But I chose a very interesting location, Mysore Palace, and ended up walking for almost an hour. I felt like, you know, let me keep going, as I was listening to more and more interesting places* (POIs)." – P5.
>
> "*I am kind of a person who likes traveling a lot. As the app told me about new and new places* (POIs)*, like turn left and there's this road, so I was like, "oh, let's see*

> *what's on that road". Even if I originally planned on walking thousand steps, I would simply go on and on, and complete five thousand steps.*" – P3.

Participants found VStroll "*addictive*", as they wanted to explore one more road, reach the next intersection, or find an interesting POI, before ending the trip, thus resulting in long trips. In total, two trips (from Marrakesh and Kala Ghoda Cafe, Mumbai) were more than one hour long, and 34 trips were more than 30-mins long. A maximum of 77 POIs were heard in a single trip; overall, our participants listened to 928 POIs (3.7±8.0 POI/trip), and missed 209 POIs as they overlapped with other POI announcements.

Moreover, the emphasis of walking indoors really motivated our participants, as they have not previously realized that they can walk indoors for long distances.

> "*I have not walked for the last 4 years because my health was bad. Also I didn't have a good space to walk. I started walking when I started using this app. I am thinking of continuing walking even after this app.*" – P14.

Finally, three participants labelled themselves as "*travellers*", as they used to travel extensively before losing their sight. VStroll brought back their fondest travel memories. For instance:

> "*Travelling is very close to my heart... Before my accident where I lost my eyesight, I loved traveling a lot. Since this app enabled me to travel, even though virtually, I would definitely want to keep using it.*" – P3.

All these motivations led our participants to continue using the app even after the study ended. After the post-study interview, we asked each participant to uninstall the app. Surprisingly, the log data showed six participants using VStroll even after the exit interview for the next 2-8 days (4.7±2.5 extra days/participant taking 12.9±12.0 extra trips/participant). We believe this as the key testament to the success of our proposed solution. (Note: Due to resource constraints, including server infrastructure and Google Maps related fees, we were not able to support the app). *E.g.*,

> "*I never like to walk anywhere. Seriously, I'm very bored to walk. But with this app, I started to like walking. Now I am very happy to find time to walk, like let's go and walk a little, or maybe do a few rounds in the house. So, now if this app goes away, I will also lose my newly developed habit of walking.*" – P10.

5.2 Immersive Virtual Exploration

A majority of the participants (12) found VStroll's virtual trips to be "*highly immersive*", making the experience very engaging. The frequent POI-, fitness- and trip-related announcements shifted participants' focus away from the monotony of walking, and towards the virtual exploration. Participants perceived themselves as "*virtual avatars*", exploring real-world locations in an accessible virtual environment:

> "*Like how sighted people use VR to explore places from home, I can also travel... in this virtual avatar.*" – P4.

Participants discussed several factors that contributed to the immersive experience, including spatial audio for POI announcement, voice input for road selection, and background sound. First, six participants appreciated the usage of binary spatial audio, as it saves time by not announcing "*on your left/right*", resulting in more POI announcements and less missed POIs. Moreover, it makes the experience immersive, for instance:

> "*Listening to the announcement is a very immersive experience... In future, if I am actually walking there in Venice or Manali, I would be able to tell people walking with me about locations that will be coming on that road, on the left or right side of the road.*" – P3.

In spite of enjoying the current POI announcement using Google TTS, four participants suggested enabling use of other TTS engines. This is in line with prior findings that most blind users switch between multiple TTS engines frequently [19]. Moreover, spatial audio required the user to wear head/ear-phones, blocking the environment sound, which is not preferable [19, 23], though none of our participants disabled the spatial audio mode from the Settings screen. Interestingly, three participants complained that they were not able to identify which "*earphone was right or left*", creating uncertainty in the POIs orientation.

Second, participants described the voice-based interaction for road selection at every intersection as "*a conversation with a virtual avatar*" (P10). Road selection provided participants autonomy of choosing their own route during virtual travel. Such autonomy is crucial as it helps PVIs to be independent and raise their confidence [28]; prior work points that people who are blind rely on sighted help for navigation, due to cultural, parental and infrastructural reasons [22, 52].

> "*I don't need any guidance, guide or sighted person or friend to travel... that's really good.*" – P10.
> "*Well, exploration offline is impossible, as there should always be someone* (sighted) *along with me who can tell me what's around me and then I would enjoy... Otherwise, I have to somehow know in advance where I am going and the places that are coming, which was possible with this app, you know.*" – P6.

Moreover, the U-Turn feature further helped the autonomy of route selection, as participants used the U-Turn button often (overall, 156 times) to walk back to the previous intersection and choose a different road.

Voice input also helped participants to interact with the app hands-free, without taking the phone out of their pockets, thus enabling them to enjoy an uninterrupted immersive walk.

> "*It gives you an option to just speak, rather than searching on the screen. So, no need to look for buttons and things like that, you just have to say the direction. That's very nice.. very good.*" – P2.

To add, P3 suggested using voice commands for regular fitness updates, instead of accessing the on-screen Update button. However, P12 was hesitant to use voice input, due to privacy concerns (similar to [19, 56]).

Third, six participants attributed immersiveness to the background wind sound. For instance, P7 described feeling the wind during a virtual trip around a beach due to the background sound:

> "*When I tried the Nariman Point trip, I felt the sea because of the background sound...* (chuckled) *though*

> *later I figured out that that sound was for all the trips. It would have been an even more immersive experience with location specific background sound.*" – P7.

Apart from the virtual travel, two participants compared VStroll experience with their prior gaming experiences. P13 found the idea of "*free exploration*" similar to Nintendo's go-kart racing game, wherein the player drives through new places, while P7 related discovering POIs with finding treasures, thus gamifying the VStroll experience.

Our participants mentioned the added benefit of virtual travel, as an accessible, safer, and hassle-free alternative to real-world travel, especially during times of COVID-19 imposed restricted mobility. For instance, six participants took a virtual trip of their University as they "*would have liked to be on the campus, but cannot be right now coz of COVID.*" – P4.

5.3 Spatial Knowledge Acquisition

Participants acquired spatial knowledge about the locations they virtually explored using VStroll. Announcements related to the road and intersections helped the participants create a mental map of the real-world location they were exploring, while POIs made them more aware of "*essential places*".

To begin with, 14 participants explored their current neighbourhood (by selecting '*current location*' as the start location) in 70 trips, mainly to learn more about their neighbourhood, and validate if VStroll is indeed working, *i.e.*, announcing the correct road name, intersection, and POIs. Hearing familiar POIs increased participants' trust in VStroll. Surprisingly, 12 participants reported serendipitously discovering new POIs, unfamiliar roads, and shorter routes to previously known POIs. For instance, P2 mentioned learning about a tailor shop next to his house where he has been staying for the last two years. This also highlights that PVIs have limited spatial knowledge of their surroundings [14, 55]. Participants encountered ATMs, medical shops, and other essential POIs which in future can help with their daily needs.

> "*Actually, I never knew that I have some Ayurvedic clinic and some dental hospital right here.. A few months back, my mother was looking for an Ayurvedic doctor. I searched on maps, I could not find any... I suddenly found it on your app and I told that to my mother and she was very happy... I am living in this place by birth, so more than 20 years now, still I didn't know.*" – P10.

After exploring their neighbourhood, most participants extended their virtual trips to their workplace (such as their office or University) and locations that they have visited/stayed in the past. Seven participants became nostalgic during the interview, describing fond memories of visiting locations where they spent their childhood, did their schooling, had a memorable vacation, or their parents/friends home. For instance:

> "*I have lived all my life there. There were so many familiar places that were announced... It reminded me of my Bombay University days... It was very exciting. When I used to take the metro from Chhatrapati Shivaji Metro Station... That was my best trip!*" – P7.

In total, 41 trips started from an educational institution (*e.g.*, IIT Bombay, Kakkavayal GHS School, University of Illinois at Chicago) and 23 trips from a workplace (*e.g.*, MDH Gurgaon, R.S. Brothers AS Rao Nagar). Even while exploring these "*known*" locations, it helped the participants increase their spatial knowledge. For instance, when taking a virtual trip of her University, P12 learned about the "*circular layout of the campus..*" and "*one can take any route to reach the gate... all roads lead to the main gate!*". Moreover, seven participants mentioned visiting a future workplace to increase their "*familiarity*", be "*more confident*" for the physical visit, and help with their travel plans, in general. For instance, P4 virtually explored the MIT campus to identify landmarks, shops, and "*places of utilitarian needs*". This finding is in line with prior work on gaining spatial knowledge using virtual navigation [14], where learning turn-by-turn instructions and POIs during virtual navigation made participants feel more confident about physically visiting the location.

Participants explored several exotic locations as well; out of the total 253 trips, 43 trips were international trips in cities like Amsterdam, Venice, Florida, Marrakesh, Paris, Mecca, etc., and to specific landmarks, like NASA Mission Control Center, Wall Street, CN Tower, etc. Participants wanted to travel and explore these locations, but due to socio-economic constraints, "*virtual travel seems to be the next best option!*". Moreover, participants noticed that international locations have no 'unnamed roads' and higher number of POIs, thus resulting in more engaging experience:

> "*That Wall street trip, I can say, that was quite perfect! In Bangalore and Bombay, at various intersections it says unnamed roads but that didn't happen in Wall Street. All roads are properly labeled... I liked the New York trip, because in that trip, it announced everything... it announced so many popular places on that road and there were many shops and all, it announced everything.*" – P11.

However, not all international trips were well received, and the main culprit was sparsity of POIs. For instance:

> "*I am very much interested in NASA... I thought that I'll get to know, get to hear a lot of things that I usually come across while reading many articles about it. But obviously those buildings are very far away since that is a very huge campus. I was also thinking of roaming around in Stanford, but then I checked that the campus is like over 8,000 acres. I would have to drive around there!*" – P3.

This problem can be solved by enabling walking at variable speed, similar to VirtualWalk [14]. In our VStroll design, the default start location is the entered start location of the previous trip. We found that for 113 (44.7%) trips, a new start location was entered, while for 70 (27.7%) trips, the previously entered start location was re-used. This highlights that people were more interested and excited about exploring newer locations.

Overall, participants appreciated the spatial learning from VStroll as it reduces their reliance on sighted assistance during real-world navigation. Moreover, it was found to be more accurate and accessible by our participants:

> "*A sighted person will tell me to go straight, turn left, turn right... they will not tell me how much I need to walk to reach that turn, or they will not tell me the landmark before the turn.*" – P6.

Prior work states that knowledge about POIs before a turn is a key information for navigation among PVIs [14]. Our participants were highly excited to learn about new POIs; four participants even mentioned sharing their newly-learned spatial knowledge with their sighted peers and family members. "*Once in a while, it is really nice to be able to share with my sighted friends, "Oh, there is this nice restaurant that I know. We can go there."* " – P3.

5.4 Comparison with Fitness, Entertainment and Navigation Apps

Participants compared VStroll with fitness tracking related phone apps and technologies, entertainment apps which they use while exercising, and navigation-cum-exploration apps that they use during traveling. Though we did not introduce VStroll as a fitness tracker app, a majority of our participants (9) found easy access to fitness tracking parameters in VStroll (such as distance walked, step count, and calories burned) similar to their current fitness apps and/or wearables. Seven participants mentioned tracking their daily step count using these apps: Pedometer (3 participants), Nike Run Club (2), MI Steps Tracker (1), and Impact (1); two participants used fitness (wrist) bands: FitBit (1) and Apple Watch (1); four participants used a fitness equipment: treadmill (3) and exercise bike (1). The four participants using fitness equipment mentioned that VStroll provided "*accessible step count updates*", which was not available in their equipment. For instance, P5 required sighted help to read step count and calories burned on his treadmill, hence she started using VStroll while walking on the treadmill to easily access these fitness parameters. However, this direct comparison with existing fitness tracking approaches led to comparison of VStroll's step counting accuracy:

> "*the Pedometer app is less reliable than VStroll*" – P10,
> "*my Apple watch is more accurate*" – P3.

Based on their past experience with other fitness trackers, five participants suggested adding specific features to VStroll, such as set and monitor fitness goals, receive daily updates on the progress, and to share their progress on social media.

Second, participants compared the "*constant commentary*" on VStroll with music apps. Twelve participants mentioned using an entertainment app while walking, in order to listen to songs, podcasts or talk shows on YouTube (5 participants), Spotify (3), Amazon Music (2), Audible (1), and Castbox (1), Most of them admitted that unlike VStroll, their music/podcast apps did not motivated them to walk, and lacked interactivity.

> "(Before VStroll,) *when I was walking, I mostly focused on listening to these podcasts... With this app* (VStroll), *because I continuously got feedback on what was around me, I was always occupied. I was like, "okay in 150 meters, there is a landmark on my left, so let me walk to reach there". This kind of nudged me to walk more. When I am listening to a podcast, it's not nudging me to walk more or something, right?*" – P4.

Third, eight participants compared VStroll to navigation apps (*e.g.*, Google Maps), and six participants found it similar to apps used for learning about nearby places when travelling (*e.g.*, Nearby Explorer, Lazarillo). All our participants used Google Maps for real-world navigation, four of them mentioned VStroll provided precise and detailed POI update, which is lacking in Google Maps:

> "*Basically, this app is really helpful if I want to explore locations that I'm not familiar with. It's really great to know each landmark, each junction... it's very helpful. With Google Maps, I cannot hear this much of precise and smaller details.*" – P13.

Three participants were interested in using VStroll, instead of Google Maps, to prepare for their future navigation plans: "*I definitely felt that if I have to go somewhere, I can use this app beforehand... to plan, and it will help me.* " – P14. Similarly, participants having experience with other nearby POIs announcement apps (like Nearby Explorer, Lazarillo) found VStroll to provide more precise and descriptive spatial information. Hence, they wanted to use VStroll even when they are physically traveling on road, not limiting VStroll to virtual walking: "*It would be better, if I could use this for real travelling as well.*" – P9.

6 DISCUSSION

In this paper, we discuss the design of VStroll, an Android application that aims to motivate people with vision impairments to walk, by enabling them to virtually explore real-world locations using an accessible interface, while physically walking in the safety and comfort of their homes. With that, VStroll solves the two key hurdles in walking regularly among PVIs: poor road infrastructures and monotonous nature of walking. Compared to prior work in virtual exploration for sighted individuals [7, 41, 45], VStroll has several added accessible features—POI announcements using spatial audio, voice input for route selection, and frequent step count related updates—which helps the user to gain spatial awareness while walking. A user study with 16 PVIs showed that participants enjoyed using VStroll, and took 253 trips, walked for 50.8 hours, and covered a distance of 121.6 kms. Participants reported VStroll to be engaging, immersive, safe, and accessible. Most importantly, VStroll motivated the participants to walk more, as we found participants using the app even after the study ended. Moreover, similar to prior work on cognitive effects of virtual navigation [11, 14, 55], VStroll increased participants' spatial knowledge, as they discovered POIs and shorter routes even in familiar locations, including their neighbourhood and University campus. VStroll enabled participants to *freely explore* any real-world location of their choice, which was lacking in prior works [14, 55]. Participants explored a variety of present (such as current location, workplace), past (such as hometown, school, memorable vacation spot) and future (such as International trip) locations, and referred to VStroll as *virtual reality for the blind*. Finally, regular updates to fitness parameters, like step count, calories burned, etc., encouraged participants to define ambitious fitness goals, and made them conscious about physical fitness. Below, we discuss key lessons learned.

6.1 Lessons Learned

Immersive Experience: An assistive technological solution uses a combination of audio and vibrotactile-based stimuli to replace the visual stimuli. To further enhance the immersive experience of VStroll, we need to better utilize both the stimuli—by varying the background sound and by exploiting the tactile feedback. The background sound can be context dependent and change dynamically based on the user's virtual location and/or nearby POIs. For instance, wind sound when walking next to an ocean, traffic sounds on a busy street, and prayer and chants sound when crossing a temple POI. The tactile feedback can be utilized for multiple purposes—to communicate suggestions for route selection based on the number of POIs on each road (*e.g.*, single vibration for the left road, double vibration for right), to provide walking speed related feedback (*e.g.*, continuous vibration till the user reaches the desired walking speed). Moreover, as it is a virtual trip, the user should be able to 'Pause' the trip, and 'Resume' it the next day from the same virtual location where she/he paused. These added features can help in creating a true virtual reality experience for PVIs.

Gamification: Though participants gamified VStroll by setting fitness goals and equating discovering POIs to finding treasures, it was not intended as per our design. There are several ways to gamify VStroll, such as providing an option to set walking goals and track progress, rewarding user with badges (*e.g.*, '1 km walker', 'Paris Explorer', 'Super Speed', *etc.*) on accomplishing system-defined goals and a way to share it on social media, and having a region-specific daily leaderboard to promote healthy competition. Moreover, a treasure hunt game can be integrated in VStroll, wherein the user needs to reach a POI (located in a 1-km radius of the user) within a specific time constraint. A more formal approach would be to gamify the technological solution following the principles of Ludic Design for Accessibility [16, 48].

Privacy Concerns: Voice input enabled hands-free uninterrupted walk, however a few participants raised privacy concerns in using voice-based interactions (similar to [19, 21]). To solve it, additional interaction options for route selection should be added, such as using on-screen gestures, phone tilt [14], bezel menu [18], or body movement. With the smartphone in pocket, in-built accelerometer and gyroscope sensor data can be utilized to classify if the user has taken a left/right/U-Turn in the physical world. This body movement gesture as input should be considered only after the beep sound, to avoid false positives. Moreover, privacy terms and information related to the data being collected should be available in accessible language, as recommended by Ahmed *et al.* [3].

Prominence-based POI: For our implementation, we used a combination of Google Maps and OSM: Google Maps for auto-complete start location, and OSM for POIs and road network. As mentioned before, though Google Maps APIs are expensive, it provides a richer set of POIs and minimal number of unnamed roads. Moreover, for each POI, Google Maps APIs provide several other information, including short description, user rating, and number of raters. This information can be used to make the POI announcement richer with added description, and also help in computing the 'prominence' of a POI using the rating and rater's information. In case of a street with several POIs (*e.g.*, Times Square), more prominent POIs should be more likely to be announced, thus helping the user to have more accurate spatial awareness. Such a prominence-based (customizable) POI announcer resembles a real-world tourist guide, and can be enabled using a conversational agent which continuously interacts with the user [20].

Crowdsourcing: Missing button labels is a common accessibility problem in smartphone apps, and PVIs have been found to meticulously label each button [19]. Similarly, VStroll users might know the name of a few 'unnamed roads' or 'unnamed intersections', or know details about a missing POI at a specific location. Even outdated information in OSM is commonplace. Hence, the future version of the app should allow the users to improve OSM street name and POI data, by collecting this information in a crowd-sourced manner from PVIs. Moreover, there should be a way to bookmark any POI, maybe using voice input, which can also help in determining the prominence of a POI.

6.2 Limitations

We acknowledge several limitations of this work. First, the results should be interpreted based on some specificities of the study, including the relatively small sample size and users' first encounter with VStroll. The positive user responses and other conclusions should be validated by future, preferably longitudinal studies. Second, despite the actual curves of the road, VStroll assumes every road between two intersections as a straight road, therefore users are unable to distinguish between a completely straight road and one that is somewhat curved. Third, as all our participants were blind and relied completely on TalkBack, it is hard to generalize our findings to low vision users, who have been found to use a combination of TalkBack, zoom magnification, and close working distance to the screen [17]. Fourth, the small sample size limited our analyses. A larger number of participants is required to identify broader trends. Finally, as all the participants were from India, a few of our findings may have roots in the sociocultural context. However, India has one-third of the world's blind population [9], making these findings valuable for a large target audience.

7 CONCLUSION

We proposed VStroll, a smartphone app to promote walking among people with blindness. It enables the user to virtually explore real-world locations while walking. The user can select a location, and start virtually exploring it on VStroll, while listening to the spatial information about points-of-interest, as they continue to walk. The app also enables POI announcements using binary spatial audio, turn selection at road intersection using voice input,and regular fitness updates. We evaluated VStroll with 16 PVIs, wherein participants used the app for five days, followed by an interview. We found several novel insights, including key motivating factors, spatial audio resulting in immersive experience, and road selection boosting navigation confidence by providing autonomy and independence. To conclude, VStroll motivated users to walk more, along with gaining spatial knowledge.

REFERENCES

[1] 2018. *Microsoft Soundscape: A map delivered in 3D sound.* Retrieved Septemeber 1, 2020 from https://www.microsoft.com/en-us/research/product/soundscape/
[2] 2021. *Mobile Operating System Market Share Worldwide.* Retrieved Apr 14,2021 from https://gs.statcounter.com/os-market-share/mobile/worldwide

[3] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. 2015. Privacy Concerns and Behaviors of People with Visual Impairments. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3523–3532. https://doi.org/10.1145/2702123.2702334

[4] Shiri Azenkot and Nicole B. Lee. 2013. Exploring the Use of Speech Input by Blind People on Mobile Devices. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (Bellevue, Washington) *(ASSETS '13)*. Association for Computing Machinery, New York, NY, USA, Article 11, 8 pages. https://doi.org/10.1145/2513383.2513440

[5] Michele Capella-McDonnall. 2007. The Need for Health Promotion for Adults who are Visually Impaired. *Journal of Visual Impairment & Blindness* 101, 3 (2007), 133–145. https://doi.org/10.1177/0145482X0710100302 arXiv:https://doi.org/10.1177/0145482X0710100302

[6] CDC. 2019. How much physical activity do adults need? Retrieved Feb 1, 2020 from https://www.cdc.gov/physicalactivity/basics/adults/index.htm

[7] Chris Christou, Cameron Angus, Celine Loscos, Andrea Dettori, and Maria Roussou. 2006. A Versatile Large-Scale Multimodal VR System for Cultural Heritage Visualization. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (Limassol, Cyprus) *(VRST '06)*. Association for Computing Machinery, New York, NY, USA, 133–140. https://doi.org/10.1145/1180495.1180523

[8] Tuncay Çolak, Belgin Bamaç, Mensure Aydin, Bergün Meriç, and Aydin Özbek. 2004. Physical fitness levels of blind and visually impaired goalball team players. *Isokinetics and exercise science* 12, 4 (2004), 247–252.

[9] HT Correspondent. 2017. *Number of blind to come down by 4M as India set to change blindness definition.* Retrieved Dec 1, 2019 from https://bit.ly/3aXHKB7

[10] Adam Dwyer. 2017. Factors That Increase Physical Activity in Youth Who Are Visually Impiared. *Kinesiology, Sport Studies, and Physical Education Synthesis Projects* (May 2017). https://digitalcommons.brockport.edu/pes_synthesis/22

[11] Agebson Rocha Façanha, Ticianne Darin, Windson Viana, and Jaime Sánchez. 2020. O&M Indoor Virtual Environments for People Who Are Blind: A Systematic Literature Review. *ACM Trans. Access. Comput.* 13, 2, Article 9a (Aug. 2020), 42 pages. https://doi.org/10.1145/3395769

[12] John Foley, Lauren Lieberman, and Barbara Wood. 2008. Teaching Strategies With Pedometers for All Children. *Re:view: Rehabilitation and Education for Blindness and Visual Impairment* 39 (12 2008), 206–212. https://doi.org/10.3200/REVU.39.4.206-212

[13] Thomas Fritz, Elaine M. Huang, Gail C. Murphy, and Thomas Zimmermann. 2014. Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 487–496. https://doi.org/10.1145/2556288.2557383

[14] João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2017. Virtual Navigation for Blind People: Building Sequential Representations of the Real-World. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) *(ASSETS '17)*. Association for Computing Machinery, New York, NY, USA, 280–289. https://doi.org/10.1145/3132525.3132545

[15] Gesu India, Mohit Jain, and Manohar Swaminathan. 2021. Understanding Motivations and Barriers to Exercise among People with Vision Impairments in India. In *Human-Computer Interaction – INTERACT 2021*. Springer Berlin Heidelberg, Berlin, Heidelberg.

[16] Gesu India, Vidhya Y, Aishwarya O, Nirmalendu Diwakar, Mohit Jain, Aditya Vashistha, and Manohar Swaminathan. 2021. Teachers' Perceptions around Digital Games for Children in Low-resource Schools for the Blind. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Virtual) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, 12 pages. https://doi.org/10.1145/3411764.3445194

[17] Danielle Irvine, Alex Zemke, Gregg Pusateri, Leah Gerlach, Rob Chun, and Walter M. Jay. 2014. Tablet and Smartphone Accessibility Features in the Low Vision Rehabilitation. *Neuro-Ophthalmology* 38, 2 (2014), 53–59. https://doi.org/10.3109/01658107.2013.874448 arXiv:https://doi.org/10.3109/01658107.2013.874448

[18] Mohit Jain and Ravin Balakrishnan. 2012. User Learning and Performance with Bezel Menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2221–2230. https://doi.org/10.1145/2207676.2208376

[19] Mohit Jain, Nirmalendu Diwakar, and Manohar Swaminathan. 2021. Smartphone Usage by Expert Blind Users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Virtual) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, 12 pages. https://doi.org/10.1145/3411764.3445074

[20] Mohit Jain, Pratyush Kumar, Ramachandra Kota, and Shwetak N. Patel. 2018. Evaluating and Informing the Design of Chatbots. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 895–906. https://doi.org/10.1145/3196709.3196735

[21] Mohit Jain, Rohun Tripathi, Ishita Bhansali, and Pratyush Kumar. 2019. Automatic Generation and Evaluation of Usable and Secure Audio ReCAPTCHA. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 355–366. https://doi.org/10.1145/3308561.3353777

[22] Vaishnav Kameswaran, Jatin Gupta, Joyojeet Pal, Sile O'Modhrain, Tiffany C. Veinot, Robin Brewer, Aakanksha Parameshwar, Vidhya Y, and Jacki O'Neill. 2018. 'We Can Go Anywhere': Understanding Independence through a Case Study of Ride-Hailing Use by People with Visual Impairments in Metropolitan India. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 85 (Nov. 2018), 24 pages. https://doi.org/10.1145/3274354

[23] Shaun K. Kane, Chandrika Jayant, Jacob O. Wobbrock, and Richard E. Ladner. 2009. Freedom to Roam: A Study of Mobile Device Adoption and Accessibility for People with Visual and Motor Disabilities. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, Pennsylvania, USA) *(ASSETS '09)*. Association for Computing Machinery, New York, NY, USA, 115–122. https://doi.org/10.1145/1639642.1639663

[24] Francis Kozub. 2006. Motivation and Physical Activity in Adolescents With Visual Impairments. *Re:view: Rehabilitation and Education for Blindness and Visual Impairment* 37 (01 2006), 149–160. https://doi.org/10.3200/REVU.37.4.149-160

[25] Lauren Lieberman, Cathy Houston-Wilson, and Francis Kozub. 2002. Perceived Barriers to Including Students with Visual Impairments in General Physical Education. *Adapted Physical Activity Quarterly* (Jan. 2002). https://digitalcommons.brockport.edu/pes_facpub/21

[26] Lauren J. Lieberman, Heidi Byrne, Craig O. Mattern, Celia A. Watt, and Margarita Fernández-Vivó. 2010. Health-Related Fitness of Youths with Visual Impairments. *Journal of Visual Impairment & Blindness* 104, 6 (2010), 349–359. https://doi.org/10.1177/0145482X1010400605 arXiv:https://doi.org/10.1177/0145482X1010400605

[27] Lauren J. Lieberman and Elaine McHugh. 2001. Health-Related Fitness of Children who are Visually Impaired. *Journal of Visual Impairment & Blindness* 95, 5 (2001), 272–287. https://doi.org/10.1177/0145482X0109500503 arXiv:https://doi.org/10.1177/0145482X0109500503

[28] Lauren J Lieberman, Moira E Stuart, Karen Hand, and Barbara Robinson. 2006. An investigation of the motivational effects of talking pedometers among children with visual impairments and deaf-blindness. *Journal of Visual Impairment & Blindness* 100, 12 (2006), 726–736.

[29] Fatti Longmuir. 1998. Considerations for Fitness Appraisal, Programming, and Counselling of Individuals With Sensory Impairments. *Canadian Journal of Applied Physiology* 23, 2 (1998), 166–184. https://doi.org/10.1139/h98-011 arXiv:https://doi.org/10.1139/h98-011 PMID: 9578956.

[30] José Marmeleira, Luis Laranjo, Olga Marques, and Catarina Pereira. 2014. Physical Activity Patterns in Adults Who Are Blind as Assessed by Accelerometry. *Adapted Physical Activity Quarterly* 31, 3 (July 2014), 283–296. https://doi.org/10.1123/apaq.2013-0039

[31] Susanna Millar. 1994. *Understanding and representing space: Theory and evidence from studies with blind and sighted children.* Clarendon Press/Oxford University Press.

[32] Tony Morelli, John Foley, Luis Columna, Lauren Lieberman, and Eelke Folmer. 2010. VI-Tennis: A Vibrotactile/Audio Exergame for Players Who Are Visually Impaired. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games* (Monterey, California) *(FDG '10)*. Association for Computing Machinery, New York, NY, USA, 147–154. https://doi.org/10.1145/1822348.1822368

[33] Tony Morelli, John Foley, and Eelke Folmer. 2010. VI-Bowling: A Tactile Spatial Exergame for Individuals with Visual Impairments. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility* (Orlando, Florida, USA) *(ASSETS '10)*. Association for Computing Machinery, New York, NY, USA, 179–186. https://doi.org/10.1145/1878803.1878836

[34] Tony Morelli, John Foley, Lauren Lieberman, and Eelke Folmer. 2011. Pet-N-Punch: Upper Body Tactile/Audio Exergame to Engage Children with Visual Impairments into Physical Activity. In *Proceedings of Graphics Interface 2011* (St. John's, Newfoundland, Canada) *(GI '11)*. Canadian Human-Computer Communications Society, Waterloo, CAN, 223–230.

[35] Tony Morelli, John Foley, Lauren Lieberman, and Eelke Folmer. 2014. An Exergame to Improve Balance in Children who are Blind. In *Proceedings of the 9th International Conference on the Foundations of Digital Games*.

[36] Tony Morelli and Eelke Folmer. 2011. Real-Time Sensory Substitution to Enable Players Who Are Blind to Play Video Games Using Whole Body Gestures. In *Proceedings of the 6th International Conference on Foundations of Digital Games* (Bordeaux, France) *(FDG '11)*. Association for Computing Machinery, New York, NY, USA, 147–153. https://doi.org/10.1145/2159365.2159385

[37] Annika Muehlbradt, Varsha Koushik, and Shaun K. Kane. 2017. Goby: A Wearable Swimming Aid for Blind Athletes. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) *(ASSETS '17)*. Association for Computing Machinery, New York, NY, USA, 377–378. https://doi.org/10.1145/3132525.3134822

[38] Kyle Rector, Cynthia L. Bennett, and Julie A. Kientz. 2013. Eyes-Free Yoga: An Exergame Using Depth Cameras for Blind & Low Vision Exercise. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (Bellevue, Washington) *(ASSETS '13)*. Association for Computing Machinery,

New York, NY, USA, Article 12, 8 pages. https://doi.org/10.1145/2513383.2513392
[39] Kyle Rector, Lauren Milne, Richard E. Ladner, Batya Friedman, and Julie A. Kientz. 2015. Exploring the Opportunities and Challenges with Exercise Technologies for People Who Are Blind or Low-Vision. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 203–214. https://doi.org/10.1145/2700648.2809846
[40] James H. Rimmer. 2006. Building Inclusive Physical Activity Communities for People with Vision Loss. *Journal of Visual Impairment & Blindness* 100, 1_suppl (2006), 863–865. https://doi.org/10.1177/0145482X0610001S10 arXiv:https://doi.org/10.1177/0145482X0610001S10
[41] Ivo Roes, Natalia Stash, Yiwen Wang, and Lora Aroyo. 2009. *A Personalized Walk through the Museum: The CHIP Interactive Tour Guide*. Association for Computing Machinery, New York, NY, USA, 3317–3322. https://doi.org/10.1145/1520340.1520479
[42] Jonathan Rowell and Simon Ungar. 2005. Feeling our way: tactile map user requirements-a survey. In *International Cartographic Conference, La Coruna*.
[43] Victoria Schwanda, Steven Ibara, Lindsay Reynolds, and Dan Cosley. 2011. Side Effects and "Gateway" Tools: Advocating a Broader Look at Evaluating Persuasive Systems. 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, 345–348. https://doi.org/10.1145/1978942.1978991
[44] Deborah R. Shapiro, Aaron Moffett, Lauren Lieberman, and Gail M. Dummer. 2005. Perceived Competence of Children with Visual Impairments. *Journal of Visual Impairment & Blindness* 99, 1 (2005), 15–25. https://doi.org/10.1177/0145482X0509900103 arXiv:https://doi.org/10.1177/0145482X0509900103
[45] Ya-Chun Shih. 2015. A virtual walk through London: culture learning through a cultural immersion experience. *Computer Assisted Language Learning* 28, 5 (2015), 407–428. https://doi.org/10.1080/09588221.2013.851703 arXiv:https://doi.org/10.1080/09588221.2013.851703
[46] Heidi I. Stanish. 2004. Accuracy of Pedometers and Walking Activity in Adults with Mental Retardation. *Adapted Physical Activity Quarterly* 21, 2 (2004), 167 – 179. https://doi.org/10.1123/apaq.21.2.167
[47] Moira E. Stuart, Lauren Lieberman, and Karen E. Hand. 2006. Beliefs about Physical Activity among Children who are Visually Impaired and their Parents. *Journal of Visual Impairment & Blindness* 100, 4 (April 2006), 223–234. https://doi.org/10.1177/0145482X0610000405
[48] Manohar Swaminathan and Joyjeet Pal. 2020. Ludic Design for Accessibility in the Global South. *Accessible Technology and the Developing World* (2020).
[49] Catherine TB and Florence G. 1997. Representation of space in blind persons: vision as a spatial sense? *Psychological bulletin* 121, 1 (1997).
[50] Evette Weil, Melissa Wachterman, Ellen P. McCarthy, Roger B. Davis, Bonnie O'Day, Lisa I. Iezzoni, and Christina C. Wee. 2002. Obesity among Adults with Disabling Conditions. *JAMA* 288, 10 (Sep 2002), 1265–1268.
[51] William R Wiener, Richard L Welsh, and Bruce B Blasch. 2010. *Foundations of orientation and mobility*. Vol. 1. American Foundation for the Blind.
[52] Michele A. Williams, Amy Hurst, and Shaun K. Kane. 2013. "Pray before You Step out": Describing Personal and Situational Blind Navigation Behaviors. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (Bellevue, Washington) *(ASSETS '13)*. Association for Computing Machinery, New York, NY, USA, Article 28, 8 pages. https://doi.org/10.1145/2513383.2513449
[53] DC Williamson. 2000. Polybat and table cricket: from adaptations to sport status. *British Journal of Teaching Physical Education* 31, 2 (2000), 16–18.
[54] L. Wyatt and G.Y. Ng. 1997. The Effect of Visual Impairment on the Strength of Children's Hip and Knee Extensors. *Journal of Visual Impairment & Blindness* 91, 1 (1997), 40–46. https://doi.org/10.1177/0145482X9709100107 arXiv:https://doi.org/10.1177/0145482X9709100107
[55] Koji Yatani, Nikola Banovic, and Khai Truong. 2012. SpaceSense: Representing Geographical Information to Visually Impaired People Using Spatial Tactile Feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 415–424. https://doi.org/10.1145/2207676.2207734
[56] Hanlu Ye, Meethu Malu, Uran Oh, and Leah Findlater. 2014. Current and Future Mobile and Wearable Device Use by People with Visual Impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3123–3132. https://doi.org/10.1145/2556288.2557085
[57] Bei Yuan and Eelke Folmer. 2008. Blind Hero: Enabling Guitar Hero for the Visually Impaired. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility* (Halifax, Nova Scotia, Canada) *(ASSETS '08)*. Association for Computing Machinery, New York, NY, USA, 169–176. https://doi.org/10.1145/1414471.1414503
[58] Yancong Zhu, Cong Wang, Wei Liu, and Yi Lv. 2019. Running Guide: Design of a Marathon Navigation System for Visually Impaired People. In *Proceedings of the Seventh International Symposium of Chinese CHI* (Xiamen, China) *(Chinese CHI '19)*. Association for Computing Machinery, New York, NY, USA, 7–15. https://doi.org/10.1145/3332169.3333579

A CODEBOOK

Table 2: Codebook from our analysis of interview transcripts. The codebook shows four themes (bold), 25 codes, prevalence (%) for each theme, and the total count of each theme and code.

Theme / Codes	Count	Theme / Codes	Count
Immersive virtual exploration (35.8%)	**228**	**Motivation for walking (27.3%)**	**174**
Immersiveness	89	Engagement	91
POI announcements	22	Exercise behaviour	11
Voice-based road selection	30	Step counter	26
Background wind sound	23	Fitness-related benefits	24
Travel experiences	22	History of trips	14
Fitness updates	15	App reinforcements	8
Screen readers	12	**Spatial knowledge acquisition (28.8%)**	**183**
U-turn	9	Mental model of the route	49
Binary spatial audio	6	POI announcements	34
Comparison with other apps (8%)	**51**	Start locations	32
Apps for exercises	17	Exquisite trips	31
Apps for entertainment	14	Experiences of real-world navigation	30
Apps for navigation	11	Walking direction	7
Apps for POI announcements	9		