Smartphone Usage by Expert Blind Users

MOHIT JAIN, Microsoft Research, India
NIRMALENDU DIWAKAR, Microsoft Research, India
MANOHAR SWAMINATHAN, Microsoft Research, India

People with vision impairments access smartphones with the help of screen reader apps such as TalkBack for Android and VoiceOver for iPhone. Prior research has mostly focused on understanding touchscreen phone adoption and typing performance of novice blind users by logging their real-world smartphone usage. Understanding smartphone usage pattern and practices of expert users can help in developing tools and tutorials for transitioning novice and intermediate users to expert users. In this work, we logged smartphone usage data of eight expert Android smartphone users with visual impairments for four weeks, and then interviewed them. This paper presents a detailed analysis that uncovered novel usage patterns, such as extensive usage of directional gestures, reliance on voice and external keyboard for text input, and repurposed explore by touch for single-tap. We conclude with design recommendations to inform the future of mobile accessibility, including hardware guidelines and rethinking accessible software design.

CCS Concepts: • **Human-centered computing → Accessibility**; *Empirical studies in HCI*.

Additional Key Words and Phrases: accessibility, blind users, vision impairment, touchscreen, smartphone, TalkBack, usage pattern, app usage, text entry, security, gestures, TTS, expert users.

ACM Reference Format:

Mohit Jain, Nirmalendu Diwakar, and Manohar Swaminathan. 2021. Smartphone Usage by Expert Blind Users. In *CHI Conference on Human Factors in Computing Systems (CHI'21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 23 pages. https://doi.org/10.1145/3411764.3445074

1 INTRODUCTION

Smartphones are inherently visually demanding, limiting their usage to people with vision impairments (PVIs). To make smartphones accessible, modern smartphones have pre-installed screen reader software, such as TalkBack in Android devices developed by Google and VoiceOver in iPhones developed by Apple. Screen reader software reads the contents of a smartphone's screen out loud, and supports a variety of touch-based gesture interactions, thus enabling the user to use the smartphone without looking at the screen, in an eyes-free manner. Although screen readers are mainstream technology, PVIs have reported problems, including privacy issues, steep learning curve, need of frequent assistance, and disturbance in work/social life due to constant relay of audio prompts [39, 45, 50]. To improve accessibility of smartphones, it is important to understand real-world smartphone usage pattern of PVIs.

Prior accessibility research work have tried to understand concerns, expectations, challenges, and barriers of smartphone usage by PVIs, in particular focusing on smartphone adoption [45], text entry performance [3, 39], gesture learning effect [28], and privacy and security concerns [2]. Different HCI methods have been used, including survey, interview, lab-based experiment, in-the-wild usage data, and their combination. Logged data has been found to be

---

more reliable than self-reporting as users tend to overestimate [16]. Most works with PVIs [3, 11, 28, 34, 36] has been conducted in lab-settings and/or use survey or interview data. More recently, a few studies [39, 45] logged in-the-wild smartphone usage data of PVIs. However, these studies were limited to novice users.

Understanding smartphone usage pattern of expert users with visual impairments is highly valuable, as expert users tend to create efficient workarounds not only to gain maximum benefit of a system, but also to overcome any limitations. By expert users, we mean computer tech-savvy individuals using TalkBack for five or more years. Understanding the hierarchy and frequency of app and gesture usage by expert users can provide insights to further enhance the capabilities of expert users, to develop tools and tutorials to help in transitioning novice or intermediate users to expert users, to help identify content relevant to teach first-time smartphone users, and overall, to help designers in developing superior user experiences for blind users.

In our work, we focus on understanding smartphone usage pattern of expert users with vision impairments. We logged smartphone usage—TalkBack, app usage, battery charging, lock/unlock and call—data of eight expert Android smartphone users with visual impairments for four weeks. We followed it with semi-structured interviews to further contextualize their smartphone usage pattern. This paper presents a detailed analysis of their gestures and app usage pattern, keyboard and voice-based text input behaviour, journey of learning TalkBack and its challenges, privacy and security concerns, and battery management. In contrast to prior works, we found expert users used directional gestures extensively instead of soft-keys, relied on voice and external keyboard for faster text input, utilized explore by touch in combination with directional gestures for quick access, switched between multiple TTS (Text-to-Speech) engines frequently, and used communication, video streaming and gaming apps similar to sighted users. They also discussed several tips, tricks and shortcuts to learn and master smartphone accessibility usage. All these usage patterns were motivated by their quest to achieve interaction speed equality with sighted users. We further contribute towards identifying challenges and proposing design recommendations for the future of mobile accessibility, such as hardware guidelines, rethinking accessible software design, AI+Crowdsourcing for accessibility, and continuous training for gesture-learning.

2 RELATED WORK

Our work is mainly informed by two areas of relevant research: accessibility for mobile devices and understanding phone usage.

2.1 Mobile Accessibility

The first smartphone with a multi-touch screen was the Apple's iPhone, introduced in 2007. It was completely inaccessible to blind users. Prior to that, there have been several works to make touch screens accessible [30, 57]. The seminal work which arguably influenced the design of Apple's iPhone screen reader, VoiceOver, was the Slide Rule work by Kane *et al.* [26]. They designed "*a completely non-visual interface that re-purposes a touch screen as a 'talking' touch-sensitive surface*". Since the release of VoiceOver in 2009, accessibility of smartphones for PVIs in resource-rich environments has been improving. However, due to affordability, iOS market share in developing regions is low. In India, it is ~2% [52]. On the other hand, India is home to one-third of the world's blind population [13], and Android is the default OS for a vast majority of them. This smartphone adoption among PVIs happened only after the release of TalkBack screen reader in Android in late 2009.

Though VoiceOver and TalkBack have matured, mobile accessibility still remains a growing field of research. Recent works in mobile accessibility have mostly focused on gestures [28, 40], text entry [3, 32, 39, 51], privacy and security [2, 4, 24], and app-specific accessibility for images [9, 63] and buttons [47].

First, it is challenging for PVIs to learn how to perform touch-screen gestures [28, 45]. To perform even a simple swipe gesture correctly, the swipe's location, speed, and angular trajectory needs to be within expected constraints. This results in a steep learning curve [31]. Approaches like corrective verbal feedback and gesture sonification has been proposed to teach touchscreen gestures [40].

Second, eyes-free text entry methods have received a lot of attention even outside the accessibility community. Even sighted people when occupied with other tasks (such as attending a meeting, or walking) that require a user's visual attention experience *situational visual impairment* [49]. In such cases, a gesture-based [23, 56] or a thumb-to-finger based [61] input method can help entering text eyes-free. Such solutions have not been evaluated with blind participants [23, 56, 61], and may not be suitable for them [41]. Hence, for PVIs, methods such as chorded braille soft keyboard [51] and gesture corresponding to Braille code [32], have been proposed. These solutions have been found to outperform standard touch QWERTY keyboard in terms of speed and accuracy. Moving beyond keyboard and gestures, researchers reported PVIs using speech modality for entering text, and found it to be faster than keyboard [3].

Third, prior work has highlighted privacy and security needs and struggles of PVIs [2, 33] in various contexts, including authentication [4, 29] and CAPTCHAs to differentiate humans from automated agents [8]. PVIs mentioned aural and visual eavesdropping as a major challenge. Also, CAPTCHAs being predominantly visual are unsolvable by them; CAPTCHAs are among the most inaccessible content on the Internet [8]. Various solutions have been proposed, such as PassChords which use a chording-based technique for unlocking and authenticating phones by detecting input finger(s) [4], and audio-based CAPTCHAs to make them accessible [8, 24].

Fourth, even if screen readers have made OS-level interaction accessible, there are still shortcomings in app-level accessibility. Analysis of Android apps found various accessibility barriers, such as missing button labels in apps [47], and missing alternative text in images on (social media) apps [38]. Solutions have been proposed for making images accessible [9, 21, 63]. Caption Crawler [21] uses reverse image search to find existing captions on the web. For missing button labels, crowdsourcing-based solutions may work [47]. Hint Me! uses crowdsourcing for generic app-level accessibility assistance, not limited to buttons/images [44].

Combining results from the above-mentioned prior work and more, several papers [15, 20, 58] provide a systematic literature review of smartphone accessibility problems faced by PVIs, propose recommendations to improve mobile accessibility, and identify new research directions. However, none of these works has focused on expert smartphone users with vision impairments.

In our work, we discuss the above usability challenges and more from the perspective of expert TalkBack users, and provide insights to assist designers in creating better user experiences for smartphone users with vision impairments.

2.2 Phone Usage

A smartphone today is not limited to communication, but has transitioned to be a general-purpose computing device. Understanding smartphone usage provide insights into user behaviour, which is crucial for improving mobile device interactions. Studies have been conducted to understand the various aspects of real-world phone usage pattern, including text messaging behaviour, app usage style, session duration, network and energy consumption, and even battery management [6, 7, 16–18, 35]. In such studies, a logging app is usually installed on the participants' phone that

PId	Sex	Age	City	Education	Occupation	Blind Years	Phone Model	Talk-Back Years	Data Days	Phone Usage (hrs/day)
P1	M	42	Bengaluru	High School	Computer Instructor	35	Redmi 8	11	21	4.6
P2	F	25	Bengaluru	Master's	Education Coordinator	25	Redmi Note7 Pro	7	31	3.2
P3	M	26	Wayanad	Master's	Student	18	Samsung Galaxy J7	6	19	2.3
P4	M	21	Bilaspur	Bachelor's	Student	21	Samsung Galaxy J7	5	30	5.4
P5	M	26	Pune	Master's	Accessibility Tester	10	Samsung Note 10	8	30	5.8
P6	M	53	Ahmedabad	Master's	Disability Expert	25	Moto G6	13	22	3.5
P7	M	38	Mumbai	Bachelor's	Legal Head	38	OnePlus 5	9	29	5.2
P8	M	22	Chennai	Bachelor's	Student	8	OnePlus 3	5	27	1.7

Table 1. Participants demography, along with data collection number of days and active hours of phone interaction per day.

runs in the background logging all or a subset of user interactions. Logged data has been found to be more reliable than self-reporting as users tend to overestimate their phone (app) usage [16].

The ProactiveTasks project [6] characterizes mobile device use sessions, by logging data of 10 participants for 18-36 days. Mathur et al. [35] collected 215 users data for 8 months in India to understand app usage, interaction with notifications, and charging pattern. Analysing mobile phone usage data has been found to be challenging [12], as it is difficult to attribute differences in usage pattern to the time frame when the data was collected, to the recruitment procedure, or to the experimental method. Hence, most recent works accompany logging data with user interviews to further contextualize the usage pattern [12, 35].

Though understanding phone usage research has mostly focused on sighted users, a few studies [39, 45] logged smartphone usage data of PVIs to further improve mobile accessibility. Rodrigueus *et al.* [45] conducted a longitudinal study with five novice smartphone users over eight weeks to understand challenges of transitioning from feature phone to smartphone, using the TinyBlackBox logger [37]. Similarly, text entry data from five blind users were logged for eight weeks to identify typing behaviour and learning effect [39]. Both these studies were with novice users. Large-scale longitudinal data with PVIs has been limited to data collected while using a specific navigation app, iMove [25] and NavCog3 [48], and does not reflect on generic smartphone usage. Other works to understand smartphone usage [42], gesture usage [11, 28, 34], using voice as input usage [3], preference in selecting text-to-speech engines [36], *etc.*, have been conducted in lab-settings and/or use survey or interview data, lacking in-the-wild usage data.

To the best of our knowledge, understanding smartphone usage pattern of blind users using logged data has been limited to novice users. In our work, we focus on expert users, log their data for a month to understand their smartphone interaction, including TalkBack and app usage, security and privacy workarounds, switching between multiple TTS engines, and battery management. We follow the logging by an interview to delve deeper into these, identify challenges and propose design recommendations.

3 STUDY DESIGN

We conducted a study during May-Aug 2020 (COVID-19 period) in India to collect logging data and conduct interviews.

3.1 Participants

Eight expert smartphone users (1 female) with average age of 31.6±1.5 years and TalkBack usage experience of 8±2.9 years participated in our study (Table 1). Three of them were blind from birth, while the others lost their eyesight later

Data Type	**Description**
Gesture	Gesture type, gesture length, gesture duration, number of items accessed (in explore by touch)
Keyboard	Number of characters (and words) typed and deleted
Voice input	Speech duration, number of characters (and words) transcribed
TTS	Number of characters (and words) processed to be spoken
Screen events	Phone locked/unlocked, method for unlocking (None/Fingerprint/Face/Pattern/PIN/Password)
Battery	Battery level announcement, battery charging on/off
Call	Type of call (incoming/outgoing/missed), call duration
App usage	Package name (unique for each app), app event type (app moving into foreground/background)

Table 2. List of data collected from the user's smartphone. *Call* and *App usage* were obtained using Android *CallLog* and *UsageStatsManager* API, respectively, while rest were obtained from *logcat*. Note: Each event was logged with Timestamp.

due to accident or disease; on an average, our participants have been blind for 22.5±10.7 years. Only one participant (P7) has a family member who is also blind. The participants were from across seven different cities of India. Four participants have graduate degrees, three have undergraduate degrees, and one has a high school diploma. Three of them were students, while the other five participants have been working for 14.2±13.3 years. All the participants have been using a touchscreen smartphone for more than 5 years. They were typical smartphone users, using the phone for 3.9±1.5 hours daily for making calls, watching videos, and use WhatsApp extensively. Three participants (P3, P5, P6) have also developed tutorials to teach TalkBack, and have shared it on YouTube and WhatsApp groups.

3.2 Logging App

To collect smartphone usage data, we developed an Android logging app. The app starts a service which runs constantly in the background and captures the users' interactions (i.e., the *logcat* data). Logcat is an Android tool that dumps a log of system messages. Reading logcat data is necessary to capture TalkBack usage. Along with logcat data, the logging app collects the app usage and call data (Table 2). As the app runs in the background, it shows a message permanently in the Notifications shade stating the same. All the data gets stored locally on the phone and anonymized before getting uploaded to the cloud storage. Note: For anonymization, our logging app removes contact information, keys pressed, content of the TalkBack speech, and any details related to content consumed on the phone. Every five minutes, the app communicates with the Microsoft Azure cloud services to securely transfer the outstanding set of anonymized log files.

Due to security reasons, Android does not allow any external app to read logcat data continuously. The only way to give that permission is by connecting the phone with a computer having Android SDK (Software Development Kit) installed, and running a shell permission command on adb (Android Debug Bridge).

3.3 Procedure (and COVID-19 Workarounds)

The study was conducted during the COVID-19 lockdown period in India, with intervals of restricted mobility. This added the constraint that the study needs to be conducted completely online. Though our study only required signing an IRB-approved consent form, installing the logging app, and an interview after four-weeks of phone log data collection, one of the major challenges in our offline study was obtaining the *logcat* reading permission to install the logging app remotely. Thus, we decided to recruit tech-savvy PVIs, who can *'connect the phone to their computer with Android Studio installed on it'*. Moreover, to ascertain expert Android users, the participant must be using TalkBack for five or more years. To add, there is a strong correlation between tech-savvy computer users and expert smartphone users [59], so this requirement might have helped us in recruiting the target demography. Two local NGO's helped us to recruit the participants, by circulating the recruitment email in their respective networks and WhatsApp groups. One of the email listing consisted of people who recently participated in a hackathon for visually impaired developers, industry professionals, and computer science students.

Thirteen individuals responded to our recruitment email. We sent them the consent form, and after they accepted it, we asked them to install Android Studio and TeamViewer software on their computer and TeamViewer QuickSupport app on their phone. After receiving a response from them that they are ready with all the required software/app installed on their devices, we conducted an hour-long online session, wherein we accessed the participants' computer and phone using the TeamViewer software, provided logcat permission and installed the logging app. We also changed a few settings on their phone, such as enabled USB debugging, warranted logging in TalkBack, allowed maximum number of background processes, and permitted unrestricted data usage by the logging app.

Out of the thirteen responses, two of them did not respond to our installation email. Three of them had Oppo and RealMe phones, which restricted reading logcat data continuously even after providing the required permissions. We had multiple remote debugging sessions with them. We were successfully able to collect 30-60 mins of continuous log data. However, it required the logging app to be manually restarted multiple times every day. As we were not able to resolve these issues remotely, we decided to continue the study with only the remaining eight participants.

After the installation, we asked the participants to let the logging app run for the next four-weeks. Whenever we stopped receiving data from a participant's phone—either when the phone was restarted or our app was (mistakenly) killed from the Notifications shade—that participant received an email asking to restart by clicking a button on the logging app. After four weeks, we conducted a semi-structured phone interview with the participant to understand their gesture and app usage behaviour, security and privacy issues, battery management, and journey with TalkBack from a novice to an expert user. We also collected demographic information. The interviews were conducted in English, audio-recorded and later transcribed. The interview lasted for ~60 mins. Participants were paid Rs 1500 (~20 USD) for participation.

3.4 Data Analysis

We conducted a mixed-methods analysis. The log files were quantitatively analyzed using within-subject paired-samples t-test and one-way repeated measures ANVOA. As we had logged data from only eight participants, we emphasize on the qualitative interview data. We subjected our interview data to open coding and rigorously categorized our codes to understand user behaviour. Two authors participated in the coding process and iterated upon the codes until consensus was reached, and then condensed the codes into high-level themes.

3.5 TalkBack Background

In Android, a single finger gesture on the screen is interpreted by its screen reader, TalkBack, while two or more fingers gesture (*e.g.*, dragging two fingers for scrolling, pinching two fingers for zooming) is interpreted by the respective app, rather than by TalkBack [55]. Our logging app only captured TalkBack gestures, hence we do not have log data for multi-finger gestures. TalkBack supports three types of gestures – (1) *Basic gestures*: swipe up/down/left/right and double-tap, (2) *Back-and-forth gestures*: swipe up then down, swipe left then right, *etc.*, and (3) *Angle gestures*: swipe up then left, swipe down then left, *etc.* Apart from these 16 gestures, TalkBack supports *Explore by touch* gesture [53].

Each gesture event is associated with an action. For instance swipe right to move to the next item, swipe up/down to cycle through navigation settings (*i.e.*, *Default, Characters, Words, Lines, Paragraphs, Headings, Links, and Controls*), swipe right then down to pull the notifications shade, double-tap anywhere on the screen to select the item under focus, and explore by touch to hear which item is under the finger. Refer Figure 1 for a brief description of each gesture. Most of them are self-explanatory, except global and local context menu, and screen search. The global context menu has commands (such as open TalkBack settings, open Text-to-Speech settings, hide screen by dimming it, repeat/spell the last utterance, read from next item, etc.) that work anywhere on the smartphone, while the local context menu commands (such as navigation settings, text edit options, controls to add/remove/edit custom labels, etc.) vary depending on the item under focus [54]. Screen search allows searching for specific text on the current screen.

To type a letter using TalkBack, a user needs to slide their finger over the keyboard until they hear the letter that they would like to type, and lift their finger up to type the focused key.

4 RESULTS

In total, we logged phone usage data for 209 days, and our participants used their phones for 976 hours. One of the broad findings that we will keep revisiting is that expert users aim to achieve speed equality with sighted users: "*Speed is crucial... I will use anything that can help me get done things faster*" - P3.

4.1 Gestures Usage Pattern

The associated actions with a TalkBack gesture have evolved over the years. One of the major changes, mentioned by four participants, was the gesture to move to the next or previous item on the screen. In older Android versions, swipe up and down gestures were associated with moving to previous and next item, respectively. Now, swipe up and down gestures are associated with cycling through navigation settings, while swipe left and right gestures are associated with moving to the previous and next item, respectively. (Note: Our participants referred to the 'navigation settings' as 'granularity level'). This resulted in severe challenges in adapting newer TalkBack versions.

> "*up down swipe which I initially learned for previous next item, later moved to left right swipe... and for up down a new thing of granularity got added... it was very confusing*" - P3.

TalkBack offers customization wherein a different action can be assigned to an existing gesture. Three participants customized their default TalkBack gestures, to resemble it with their prior learning. P1 and P2 customized swipe up/down to move to the previous/next item and swipe left/right to cycle through navigation settings, while P6 configured both swipe up/down and left/right to move to the previous/next item, as "*I rarely change my TalkBack granularity level... whenever I need to, I use the local context menu to do that*" - P6. Apart from these, no other participant customized any gesture. For log analysis, we assigned appropriate action to the logged gesture.

Overall, explore by touch was the most frequently used gesture with a total count of 104,526 events (including 56,433 text entry events, discussed in Section 4.4). Prior work suggests that novice TalkBack users rely solely on explore by touch rather than directional gestures, and after a few months of TalkBack usage, users almost exclusively use gestures to navigate around the interface avoiding explore by touch [45]. In contrast, we found our expert users using explore by touch in combination with directional gestures. Our participants stated that they remember the spatial location of most items, including apps, battery, and keyboard layout. They use explore by touch to directly hit the target. When they miss the target, they perform a quick gesture to move to the previous/next item to reach the target. Explore by touch was used in the true 'exploration' sense only while exploring a new app.

> *"I know the location of the app, so I directly click that location. If I miss, I do max one left or right swipe to reach the correct app."* - P3.

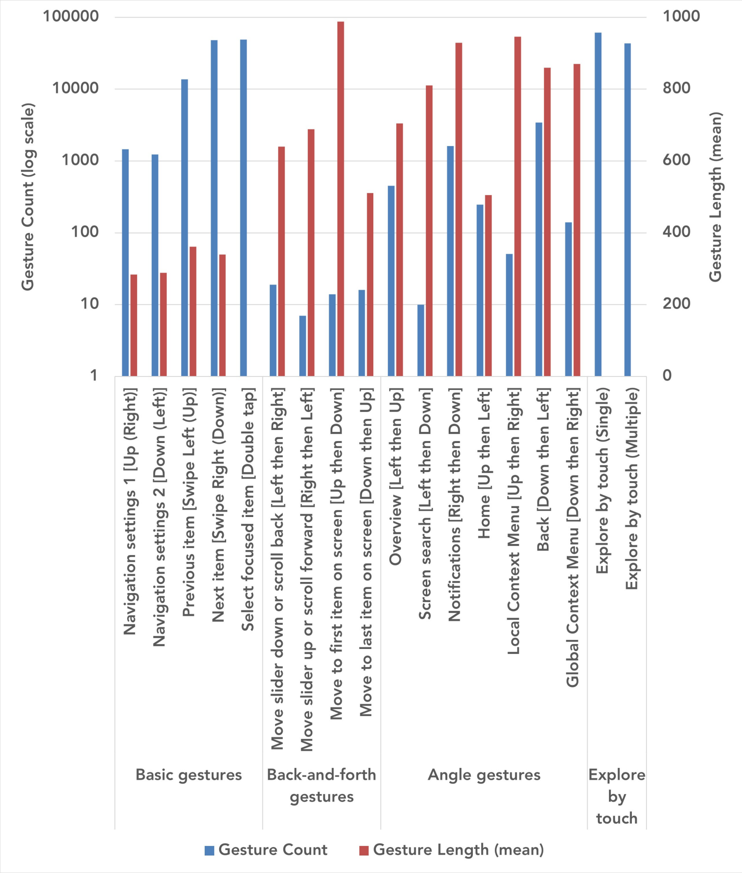

Fig. 1. Overall gesture count (in log scale) and average gesture length (in pixels)

Thus, explore by touch was mainly used to directly tap an item (similar to a single-tap), and once the item was in focus, double-tap was used to select it. Note: TalkBack does not support single-tap. We divided explore by touch in two categories: (a) *Explore by touch (Single)*: when a single item was touched (61,123 times) in an explore by touch gesture, and (b) *Explore by touch (Multiple)*: when two or more items were touched (43,403). We conducted a paired t-test analysis and found that Explore by touch (Single) was used significantly more than Explore by touch (Multiple) with t(7)=2.6, p<0.05, thus implying explore by touch was used primarily for target acquisition, not exploration.

The double-tap gesture to select an item was the second most frequently used with a total of 49,052 logged events. It was closely followed by the gesture to move to the next item (by swiping right by 5 participants, swiping down by 2 participants, and swiping right/down by 1 participant) with a total of 48,229 recorded events. The top two angle gestures were for back (3424) and pulling the notifications shade (1616). All the back-and-forth gestures were among the five least frequently used gestures. ANOVA test showed a significant main effect of gesture type on gesture usage, with F(3,21)=14.8, p<0.01. On investigating pairwise differences, we found usage of basic gestures and explore by touch to be significantly more than angle gestures (p<0.001) and back-and-forth gestures (p<0.001), and usage of angle gestures to be more than back-and-forth gestures (p<0.05), thus highlighting the importance of difference gestures.

Figure 1 also plots the average gesture length (in pixel) for all the gestures. The ANOVA test showed a significant main effect of gesture type on the gesture length, with F(2,14)=7.9, p<0.01. This prompted us to investigate pairwise differences. We found gesture length of basic gestures (swipe up/down/left/right) (318.5±75.2 pixels) to be significantly shorter than the angle (803.5±150.6) and back-and-forth gestures (706.7±353.5). We found similar results for gesture duration. These results are expected as angle and back-and-forth gestures are a combination of two basic gestures.

Most participants relied on gestures even when alternative options were present, as using gestures was much faster. For instance, they used Home, Back, and Overview gestures instead of the three equivalent soft-keys present at the bottom of the screen. Clicking the Back soft-key, at best, requires explore by touch (single) to select the key followed by a double-tap anywhere on the screen to activate it. On the other hand, a swipe down then left gesture triggers Back, which is much faster. The gestures performed by our participants were not limited to TalkBack gestures. Two participants use a richer gesture set to enhance their interaction speed.

> "*I use Nova app for more gestures... pinch up gesture for opening the File Manager, two-finger clockwise rotation for closing all apps, 'o' gesture to open camera.*" - P8.

During the interview, we asked about the most important gestures, and participants responded with previous/next item (8 participants), double-tap (8), explore by touch (8), notifications (5), changing granularity level (4), two-finger swipe up/down to scroll (4), two-finger swipe left/right to switch screens (3), back (3), home (3), overview (3), global and local context menu (2), screen search (1), and long pressing the Home soft-key to trigger Google Assistant (1). There is a relation between the perceived importance of a gesture and its frequency of usage.

Diving into the three less used, but described by our participants as *"advanced features"* – global context menu, local context menu, and screen search. In total, global context menu was accessed 140 times, local context menu 51 times and screen search 27 times. Participants use global context menu mainly to access Text-to-Speech (TTS) settings (6 participants), to dim the screen (4), and to change TalkBack settings (3). Local context menu was mainly used inside the edit text box to select/copy/paste text (3 participants), to label unlabelled buttons (2), and to change granularity level (2).

4.2 Learning TalkBack

Though learning these swipe gestures seem trivial for sighted users, Google TalkBack help pages states, "*For all gestures, use a single motion, a steady speed, and even finger pressure.*" [55], which is non-trivial for PVIs [31, 40]. P3 confirms this:

> *"Learning even the simplest gestures is hard. How much pressure to apply? How to move finger? From where to where? Ensuring that other fingers are not touching the screen... The same goes with double-tap, people may hit the screen like a hammer or with lot of delay between the two taps."*

Our participants have been using TalkBack for 5+ years. Their learning phase was broadly in agreement with earlier studies of novice users [14, 45], mentioning problems with performing basic gestures, and challenges in performing simple tasks like picking a call. During our interviews, we dug deeper into their growth in TalkBack expertise over the years. We found all of them learned TalkBack by experimentation. All, except P4, were aware of the TalkBack tutorials provided by Google. However, only P5 and P8 used the tutorials. As per World Wide Web Consortium (W3C) mobile accessibility guidelines [60], tutorials are required to teach users what gestures (and alternatives) can be used to control a given interface, and these tutorials should be easily discoverable and accessible. Prior work reports several shortcomings of the TalkBack tutorial, which were corroborated by our participants [45]. Instead, they took help from their sighted friends, family members and/or colleagues to get started. The sighted person helped them to hold the phone in the correct orientation, explained the layout of the phone, and provided a commentary of what was happening

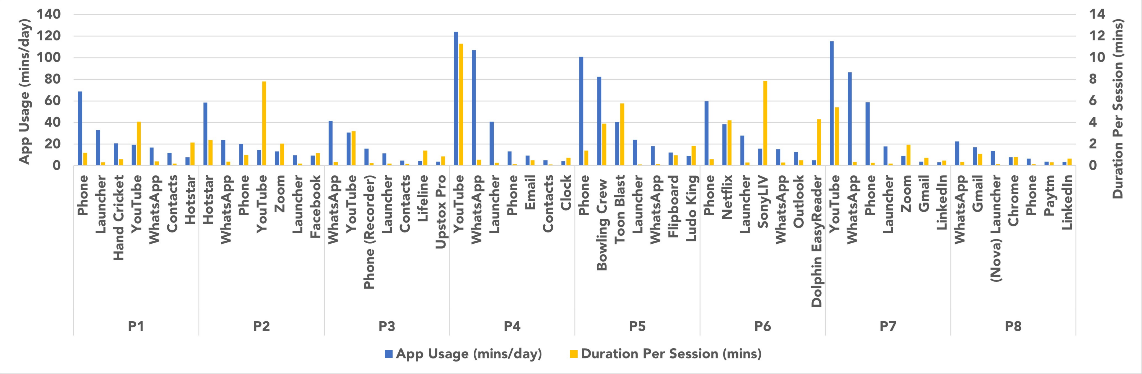

Fig. 2. Top seven apps of each participant based on average usage (mins/day). Also plots average duration per session for those apps.

on the screen on performing different gestures. Three participants explored the TalkBack tutorials much later, just out of curiosity. We found a divide among their opinions.

> *"Starting afresh just using the tutorial, well... is not possible, as the person may not even be familiar with terms like 'swipe', especially in India."* - P2.
>
> *"I teach Android to others. I ask them to go through those tutorials. It has some benefits... it is easy to introduce TalkBack using that... specifically to people familiar with Eloquence or eSpeak* (in computer)*."* - P3.

All participants were in sync that the first few gestures they learned were explore by touch, gestures to move to the next/previous item, and double-tap for selection. It took them 3-6 months to get comfortable with TalkBack. Prior work reports even after two-month of TalkBack usage, participants were struggling with it and mastering screen readers on smartphones is *"an arduous and long task"* [45]. Our participants learned non-basic gestures much later, via podcasts (P2), online articles (P5, P6), and friends and family (P1, P3). Self-learning has been very challenging. To avoid mistakes, P1 started learning TalkBack without a SIM card on the phone, and P2 started on a tablet, as it offered larger buttons.

> *"I remember, initially when I opened an app, I didn't know how to close it. So, I would just restart my phone...* (laugh)*... after a few months, my sister taught me the back button."* - P1.

Learning never stops, as with every new app interface, participants mentioned self-experimenting to learn how to use it. Moreover, phone switch has a learning curve. As per P3, gesture learning is not generalizable across phones, as different phones have different screen size, different pressure sensitivity, and different lag (due to hardware differences).

Based on our participants' responses, to learn TalkBack, one needs to be very patient (P5), and spend time practicing and mastering the basic gestures (P3). Moreover, all participants pushed for relying more on gestures, compared to explore by touch or Home/Back/Overview soft-keys (P6), because gestures increase the interaction speed.

4.3 Apps Usage Pattern

Across the eight participants, they used 352 different apps over a period of 242 days. (Note: Though we logged data for 209 days, the app usage statistics API returns last few days of app usage data. Similarly, call records API also returns historical data.) On an average, they used 19.9±7.9 apps daily. The highest used app across participants was WhatsApp (42.1 mins/day), closely followed by Phone (41.7), YouTube (39.5) and Launcher (21.2). Launcher is the app running on the home screen, providing ways to arrange, organize, and interact with the apps. Interestingly, this is similar to app usage pattern of sighted users in India with highest usage of communication apps [35]. Also, prior work, wherein people with blindness or low vision self-reported their app usage pattern [19], found the most frequently used apps to be for social networking, entertainment, and e-mail, which is well-aligned with our findings. However, in self-reporting, it seems people tend to underestimate their Phone and Launcher app usage. In Figure 2, for each participant, we plot the seven most used apps based on the average number of minutes used per day. Phone, WhatsApp and Launcher were in the top seven most used apps for all the participants. Users spent most of the time interacting with a few apps (Figure 2). The graph declines exponentially with a long tail. Users spent 63.4±14.2% of their time interacting with the top three apps, and overall each participant accessed 78±33.7 apps during the logging period.

Several video streaming services were in the top seven list. Three participants were into smartphone gaming, and only one of them limited it to text-based games (like Random Adventure Roguelike). Due to COVID-19, a few participants reported communicating more over phone/Zoom calls while working from home, and also using more emails than usual. A few apps in that top list were very specific to the individual’s interests. For instance, P3 enjoyed investing in stocks (Upstox Pro) and mental health (Lifeline), P5 was into news (Flipboard), P7 was searching for a new job (LinkedIn), and P6 enjoyed novels (Dolphin EasyReader). None of the participants had any browser in the 20 most used apps. They mentioned that they prefer browsing on a computer, mainly because JAWS is faster.

> *“I like using apps instead. I find browsing on phone very unpredictable. You have to swipe many times. In app its very predictable and more importantly, you know where buttons are.”* - P7.

Figure 2 also plots the average length of each session for those apps. The app usage time accounts only for when the app is in the foreground. *E.g.*, during a phone call, if the user starts playing a game, then only the game will be logged as it is in the foreground, as phone call goes to the background. Hence, Figure 2 shows that apps which stop working when they are in background like video streaming services (YouTube, Hotstar, SonyLIV, *etc.*), and apps which require constant attention like gaming (Toon Blast, Bowling Crew, *etc.*) have average session duration of greater than 2 mins. This is in contrast to sighted users with games having the longest average session length of 195 sec, followed by shopping apps of 101 sec [35]; YouTube or similar apps for streaming services did not have long session length.

During the interview, we asked the participants about the most important apps, and we found a disconnect with the most used apps. A few important apps mentioned by them: Google Pay for mobile transactions (P1, P4), Rapido for booking bike rides (P1), Uber/Ola for booking cabs (P3, P6), Instagram (P2), Be My Eyes, Supersense and Envision AI to understand the visual world (P3, P6), @voicealoudreader for reading books (P5), and Nearby Explorer Online and Lazarillo to know about nearby places (P2, P6). We dig deeper into the Instagram usage by P2.

> *“Everyone is intrigued by my Instagram usage. People wonder what would a blind person do on Insta as its about pictures. Mostly I use it for book and music recommendations, poetry, inspirational quotes... People rarely add descriptions to their images, but I read the captions... Book and music reviews appear in captions, so it doesn’t matter. Also, a lot of people who write poetry are on Insta and they post a lot.”* - P2.

4.3.1 Calls: The incoming and outgoing call duration were separately logged. On an average, our participants talked using the phone app for 2.3±1.4 hours daily. While talking on the phone, most participants locked their phones, thus leaving no app running in the foreground. This led to a few discrepancies. *E.g.,* for P7, the average time actively interacting with his phone was 5.2 hours/day based on the foreground app activity (Table 1), the average time taking phone calls was 5.3 hours/day, while the Phone app logged usage was only 0.9 hours/day (Figure 2).

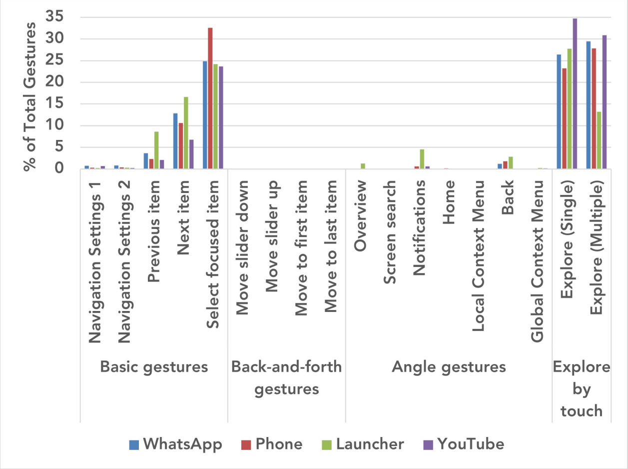

Fig. 3. Gesture count (in % of total gestures) for the four top used apps.

4.3.2 App-specific Gestures Usage Pattern. We analyzed the gestures used in the four most used apps - WhatsApp, Phone, YouTube and Launcher (Figure 3). Data reveals that expert users performed minimal exploration on the home screens, as the Launcher app has significantly less Explore by touch (Multiple) gestures compared to the other apps (F(3,21)=4.5, p<0.05). This further strengthens the argument that expert users remember the spatial location of apps on their home screens. Hence, they perform Explore by touch (Single) followed by either a double-tap (if they hit the target), or a quick Next/Previous item gesture (if they missed the target) and then a double-tap.

Similarly, it seems the gestures for showing the Notifications shade and Overview of recently used apps were called mainly from the home screen. Comparatively, YouTube interaction consists mainly of exploration with minimal double-tap selection, may be because the content gets dynamically updated and the user does not have to move sequentially. In contrast, the Phone app has maximal double-tap selection gesture, to dial a contact.

4.4 Text and Voice Input

Log data shows 47,615 characters were typed by our participants using the keyboard, plus 8,818 characters were deleted using the backspace key, significantly more than the text entered by novice user [39]. Text entry using keyboard was triggered by Explore by touch (Single/Multiple) gestures, depending on the number of keys accessed before hitting the target letter. Out of the total 352 accessed apps, text was entered using the keyboard in 69 apps. On a daily average, participants typed 184±287 characters and 14.9±23.8 backspace. Participants entered 82.7% of the total characters on WhatsApp, and deleted 7.0% of those characters. Other apps with considerable text entry were Contacts (2.5%), Phone (1.9%), YouTube (1.0%), and Gmail (0.9%). Compared to WhatsApp, the backspace key was used more often in Contacts (11.2%), YouTube (9.1%), and Gmail (7.1%). This can be mainly because searching for a specific contact/video or typing a formal email requires more precision compared to typing an informal WhatsApp message.

The Google keyboard, Gboard, provides an option to enter text using voice. Due to a bug in our logging app, we have voice-based text entry data only from five participants (P4-P8). The voice input was used 578 times in a total of 15 apps; a majority (94.5%) of which was used in WhatsApp. The average length of a voice input was 12.3±7.5 sec, inputting

9.3±7.2 words. Similar to prior findings [3], our participants preferred voice input as they found it to be faster than typing. Interestingly, 80.5% of the voice input was by P7.

> *"For formal messages, I prefer to type using keyboard... If its a long, informal message, I use STT* (speech to text) *for typing. As informal messages doesn't have to be super accurate, STT is much faster... "* - P7.

Though participants were more familiar with the keyboard, whenever they wanted to type something fast or a lot of text (as in Google docs, WhatsApp, or email), they used voice input. For shorter content, like searching on Google/Amazon, they preferred keyboard. P3 complained that during intermittent Internet, he was not able to use voice input, which is a challenge in developing regions. Also, it was hard to correct mistakes in the generated text of the voice input, highlighting the speed-accuracy trade-off.

> *"Voice-based input is much faster... The only problem is that it is difficult to make corrections to whatever is being entered wrong. You have to go back... correct it... come back... overall, that takes time... Whenever a lot of error happens, I switch to keyboard."* - P2.

4.4.1 Google Assistant. All, except one, participants used voice-based Google assistant for several tasks, including googling (4 participants), making phone calls (4), and sending WhatsApp messages (2), opening apps (2), and putting alarms and reminders (2). Participants used Google assistant as it was faster than the other approaches to perform that task, however its has its own limitations. (Note: We could not log voice commands for Google assistant.)

> *"I do use it* (Google Assistant) *to call someone, as its easier than finding people in contact. However, in crowded places or when I am on the road as I need to travel a lot for my job, it is hard to use voice assistant, so speed dial is better... I also use it to send WhatsApp messages without even touching the phone."* - P6.

4.4.2 External Keyboard. Though typing long messages on the phone was painful (as described by P3 below) and participants had various workarounds, only P5 and P6 entered text using an external Bluetooth/USB keyboard. P7 and P8 stopped using external keyboard, as P7 found it a hassle to carry keyboard, and P8 believed voice-based text entry to be faster than his external keyboard typing skills. (Note: We did not log external keyboard events.)

> *"When I need to type a long message on the phone, I usually type it on my computer and email it to myself. Then on my phone, I copy the text from the email and paste it on WhatsApp and send."* - P3.
>
> *"TalkBack offers all kind of keyboard shortcuts. So I can use the whole phone, similar to my computer, from the keyboard in a super fast manner... Its really fast. I have even created tutorials on how a combination of swiping on the phone screen and keyboard shortcuts together is the best way to use TalkBack."* - P6.

4.5 Text-to-Speech (TTS) Engines

TTS is a key ingredient of an accessibility suite, as overall our participants heard 361,741.5±348,799.8 characters daily. The maximum characters were spoken in WhatsApp (22.5% of total), followed by Phone (8.6%), YouTube (7.9%), and Email (7.6%). All, except P7, used multiple TTS engines. We found them using a combination of Google TTS (all), eSpeak (4 participants), Auto TTS (4), Vocalizer (4), Eloquence (3), Shine Plus (2), Jsue Commentary (1), and Samsung TTS (1). There were several reasons to use multiple TTS.

First, our participants were multilingual. They mentioned receiving WhatsApp messages in non-English languages and reading news in local languages, including Hindi, Kannada, Bengali and Malyalam. eSpeak and Vocalizer were used to read local languages, while Auto TTS helped to switch between TTS engines based on language. Not only language, two participants pointed that Google TTS has an American accent, hence they use Eloquence for English as it has an Indian accent. Most of them switched to Google TTS for reading emojis.

> *"Auto TTS recognizes the language and automatically switches the language easily. Google TTS should ideally do the same... while reading a sentence with two languages, like Hindi and English, Google just reads one language and skips content in the other language."* - P5

Second, speed was a factor in choosing a TTS. In spite of Google TTS supporting multiple Indian local languages, four participants found Google TTS to be slow with significant lag, in comparison to eSpeak which has *"only a few milliseconds of lag"* (P3). Finally, participants preferred TTS with rich features. P3 mentioned using the Jsue Commentary TTS, as it has "*advanced AI features*" like solving CAPTCHAs and reading unlabeled buttons.

Apart from switching between TTS engines, participants tend to update speech rate. They mentioned gradually increasing the speech rate with experience (similar to [43]), and also when they are in a hurry. Two participants mentioned that they have reached the maximum allowed speech rate. Participants showed sympathy towards TTS as it is hard to convert text into speech during certain conditions, and suggested future improvements. For instance,

> *"In WhatsApp, people write Hindi using English alphabets... TTS will obviously fail. Human need to fix themselves, its not the TTS fault... For emojis, TTS are pathetic... its highly irritating. Imagine listening to 'rolling on the floor laughing' 6 times in a row. People need to think before sending a message."* - P2

Each TTS engine support a variety of voices. However, our expert users placed little emphasis on the voice quality, and their choice was driven by performance, similar to prior findings [36]. P6 shared his extensive knowledge about that:

> *"TTS selection should suit you... People don't even know that different voices exists. Female voices are more clear as it has more treble... very shrill, so even works in crowded places. But when used for long, its tiring, might become irritating. Male voices are heavier, have high base, so better for longer use like when reading a book. However, they are not clear at high speech rate."* - P6.

4.6 Miscellaneous

4.6.1 Security and Privacy. Our study broadly concurs with prior studies in terms of concerns and workarounds devised on securely entering passwords with TalkBack [2, 33], shoulder surfing [62], challenges of using headphones, and the privacy compromises of not using headphones [2, 27, 33, 46]. Participants mentioned using 'speak password' setting, which ensures password characters are spoken only while using headphones. Five participants mentioned using headphones through out the day, while three were not comfortable wearing it, due to health and/or safety concerns, thus compromising their security.

> *"Ideally, I should use headphones everywhere... But using it all the time is not healthy for my ear. Also I never use headphones on the road as it is not safe. However, on the road, I have to book Ola* (cab). *I need to enter my Google Pay PIN for the booking... anybody standing next to me can overhear."* - P4.

Participants have devised various ways to solve these challenges, such as configuring fingerprint as password wherever allowed (P5), using the Google password manager to minimize manual password entry (P6), limiting password entry on phone only while at home (P8), using bone conduction ear plugs to ensure that ears are not covered in public places (P6), and using their phones with the minimum brightness level for security and privacy reasons (P1-P8).

Prior work suggests blind users are reluctant to use auto-lock [62], a feature that automatically locks the phone after being idle for a pre-defined time. In contrast, six participants reported using auto-lock, with fingerprint for unlocking the phone. Two of them use auto-lock with Google Smart Lock to minimise the number of times phone getting automatically locked. Only P3 and P4 disabled auto-lock, as their new phones did not have a fingerprint scanner. Log data showed that participants unlocked their phones 18.9±14.8 times daily. None of the unlocks were reported by face unlock. Only P5 phone supports face unlock, but he was more comfortable with fingerprint. Participants were able to successfully unlock their phones using fingerprint 99.4% of the times in two attempts, with a failure rate of 9.1% during the first attempt. When fingerprint fails multiple times, the fallback option is to enter the 4-digit PIN; our participants took 3.6±1.1 attempts to enter the correct PIN.

4.6.2 Battery. Prior work established that blind as well as sighted users get anxious when the battery of their smartphones drop below 20% [35, 62]. Our study broadly concurs with these concerns. E.g.:

> *"I always try to keep my phone fully charged, as I am highly dependent on it. Without phone, I practically cannot do anything... TalkBack consumes a lot of battery as it is always running in the background... When my battery reaches around 50% I put it on charging."* - P4.

Certain decisions were solely taken to optimize battery. P6 switched to Shine Plus TTS as it *"consumes very little battery"*, and P5 bought a premium phone to minimize the impact of TalkBack on battery and performance. Log data showed that our participants charged their phones 3.5±2.7 times daily, the length of a phone charging session was 41.8±132.5 mins, and they put their phones on charging when the battery level was 60.2±24.8%. This indicates that a user's decision to initiate a charging session does not depend on their current battery level, similar to sighted users charging pattern [35]. COVID-19 also impacted their charging behaviour. P2 keeps her phone fully charged during daytime to ensure that she can take *"extended Zoom calls"*. Participants used to check battery percentage level often (10.2±7.7 times daily). P5 asked for a dedicated gesture to check the battery level.

4.6.3 Making Phone Accessible. Most participants mentioned accessibility issues with several apps, mainly unlabeled buttons, which has been discussed in prior work [47]. Providing clear labels to interactive elements is part of the W3C mobile accessibility guideline [60]. Participants highlighted different ways to deal with it:

> *"I label all buttons which are unlabeled. For example, in the Gaana app, I labeled Play, Pause, Fast Forward, etc. all by myself. I will play a song, press a button and then label it... I worry that if I change my phone, I will lose all those labels... Paytm has a lot of accessibility issues. However I cannot do trial and error on Paytm, as I may end up losing money, so I took sighted help for labeling."* - P2.

P5 also discussed labels, and mentioned a way to back-up custom labels in TalkBack settings. Moreover, he also know how to import and export custom labels, and has shared custom labels with others to help them make apps accessible.

Another main issue pointed by two participants is the ability to find accessible phones, as most phone reviews do not discuss the accessibility aspect. For P2, buying a phone was a strenuous task:

> *"I went to several stores, and tried a lot of phone before buying... even the very basics of accessibility were missing. For example, the Dialer app has unlabeled buttons, rearranging apps was not accessible, etc."* - P2.

According to P5, phones need to have high processing capabilities to minimize TalkBack and TTS-related lags. He had negative experience with low-end phones, *"it triggers explore by touch, when you try to do back gesture"*.

4.7 Tips, Tricks and Shortcuts

We enquired with the participants about any tricks and/or features that they would teach novice or intermediate smartphone users. Along with the generic wisdom of listening carefully to TalkBack instructions and be patient while learning TalkBack, we found the following insights from our expert users.

First, three participants mentioned organizing the home screens properly for quick app access. P1 mentioned that he has 10 home screens, and he remembers spatial location of all apps on each of the home screens. Similarly, P6 stated:

> *"The first home screen is completely dedicated to direct dial numbers. The second home screen is well organized with my daily use apps. The first row has all news apps; second has music; third entertainment... fourth row has vision-related Supersense, Be My Eyes and the last row has important ones like WhatsApp, Outlook... I remember the location of each app... that makes interacting with them very fast."* - P6.

Second, participants mentioned a variety of shortcuts – long pressing both volume keys together to toggle TalkBack on/off (4 participants), gestures for Home, Back and Overview (3), long pressing app icons to access quick menu options (2), and ways of fast dialing (2). P5 mentioned speed dial, while P6 talked in detail about the Direct Dialing widget.

> *"I extensively use Direct Dialing. It creates an icon on my home screen, which I can just tap to start a call. It is the fastest way! I don't have to go to the Dialer app for speed dial... neither I have to select from the favorites, nor I have to click that phone number shortcut on home screen which also requires to open and then dial."* - P6.

Third, three participants talked about the different TalkBack settings, including turning off 'usage hints' in explore by touch (to disable hints like 'double-tap to activate' on every item in focus), and access to multiple TTS engines and voices. Finally, we discussed several tricks through out the paper, such as manually labeling buttons to make apps accessible, using a CAPTCHA reader, and connecting external keyboard to enhance text entry.

4.8 Limitations and Future of Smartphone Accessibility

Participants cited several instances where the smartphone accessibility failed, and they took sighted help.

To begin with, TalkBack tutorials are very limited, and PVIs need sighted help to get started with TalkBack.

Second, participants mentioned frequently receiving emails and WhatsApp messages with text embedded in an image. A few participants stated using Speak! or KNFB Reader app to read text in an image, while others took sighted help. Participants asked for OCR integration and CAPTCHA solver in future versions of mobile accessibility tools.

Third, participants suggested intelligent solutions for automatically labeling buttons.

> *"TalkBack should be able to read the button label template from the apk file.* - P3.
> *If the icon looks like a spanner, Android knows thats icon for Settings, and should auto-label it."* - P5.

To extend it further, P8 suggested developing a tool to automatically convert any app into an accessible app.

Fourth, participants complained about the camera app, as it does not provide enough feedback to help them capture images. In the contrary, P5 mentioned that the Samsung camera app provides feedback like *"one person in focus, two person in focus at the upper or middle part of the screen"* and clicking can be triggered just by saying *"smile"*. Still, even P5 felt that more AI capabilities can be added to guide further in capturing good quality images, for instance, give instructions on how to move the phone, provide elaborate details of what is on the screen, *etc.* Moreover, two participants (P4, P8) raised concern about the usability of the Gallery app to view and share photos. P4 mentioned that instead of referring to the photo as '*img15*', it should state metadata details, including date, time, location, and description. Apart from the Camera app, participants found issues with the Google Play store app.

> *"Apps are not categorized properly... There should be a place where it lists all useful apps for blind... Also, it should have a better recommender system. E.g., with Maps, Lazarillo should be recommended."* - P6.

Fifth, participants suggested various additions to enhance their interaction speed, such as enabling triple-tap gesture, showing differing options based on the time duration of the long press action, and developing completely conversational-based interface (similar to the JustSpeak project [64]).

> *"Everything should be completely voice-based... that will be super fast. That can help in achieving time equality with a sighted user. E.g., 'Ok Google, book an Ola from Silk Board to Bellandur'. I can't do that now. I have to do that in so many steps. But with such a UI, there won't be a difference between me and you."* - P4.

Participants wanted these new gestures to be associated with battery level, dimming screen, and announcing time.

Finally, participants complained about TalkBack and TTS response time, hindering their interaction with the phone. The responsiveness has improved over the years, and participants expected future versions of TalkBack to be *"much faster"*. At times, such updates to TalkBack and apps were not appreciated by the participants. For instance, three participants complained about the change in the mapping of TalkBack gestures to actions (as discussed earlier), P7 complained about the removal of dimming screen shortcut from Android 8.0, and P3 complained that the newer version of Ola app not being accessible. Most of these updates require re-learning, which was troublesome.

5 DISCUSSION

In this paper, we studied the smartphone usage pattern of tech-savvy users with visual impairments, who have been using TalkBack for five or more years. We logged their smartphone usage data for four weeks, and followed that with an interview to delve deeper into their usage pattern. We found several aspects of our participants' smartphone usage to be in accordance with prior findings [2, 27, 33, 39, 46, 62]. We also discovered usage patterns that were novel and/or in contrast to previous reports, including usage of explore by touch with directional gestures for quick app access (emulating single tap), extensive directional gestures use instead of soft-keys, reliance on voice and external keyboard for text input, using apps (for communication, gaming and video-streaming) similar to sighted users, usage of a combination of TTS engines and voice, and several tricks to learn and master smartphone accessibility usage. The key motivating factor for our expert users is to achieve speed equality with sighted users, which has not been previously reported as prior studies focus on novice/intermediate smartphone users with vision impairments.

To achieve mobile accessibility, World Wide Web Consortium (W3C) has defined guidelines under four broad principles: Perceivability, Operability, Understandability, and Robustness [60]. In spite of these guidelines being bare minimum, to our surprise, we found that current smartphones are still lacking on several aspects of the Understandability guideline, such as missing button labels, updates to gesture mapping, shortcomings in TalkBack tutorials, etc. In future, we expect smartphones to abide by these guidelines. Moreover, the guidelines also need to be updated to include mobile hardware accessibility related features, such as tactile fingerprint scanner, default touch screen sensitivity, etc.

Learning spatial location is not only limited to help with quickly accessing apps, but can also help users to access any button whose position is fixed. For example, in WhatsApp, Gmail and SMS app, the '*Compose*' button is at the same bottom-right position on the screen. Uniformity in design is crucial for accessibility [10]. Moreover, there is uniformity across platforms, for instance, Microsoft Office on a computer and on a phone have a similar ribbon structure at the top, to reduce the learning curve for both sighted and blind users.

Innovative accessible apps have the ability to make people with vision impairments 'independent', minimizing their reliance on sighted support. We found a mismatch between our participants perception of the most important apps versus the most frequently used apps. Our participants used advanced-AI and crowd-backed apps, like Be My Eyes and Envision AI, to help them in their day-to-day work in solving CAPTCHAs to finding their glasses. Apps like Google Pay and Upstox helped them achieve financial independence, while Uber and Ola helped with mobility independence. A few participants were hesitant in exploring banking and financial apps due to accessibility barriers and privacy, security and/or financial concerns, which need to be mitigated by improved design.

Eyes-free text entry for touchscreen smartphones have been a key area of HCI research [23, 32]. It is fundamental in effectively using smartphones. Given the multiplicity of options and tools available, including virtual and external keyboards, voice input with increasing accuracy rates, and word completion tools, there are various ways for people with vision impairments to gain competence in text entry. Its important to make them aware of the available technologies, and up-skill them to adapt the latest technology.

We found our participants aspire to reach parity with sighted users in terms of their smartphone usage speed. Earlier work with novice TalkBack users mainly focused on achieving accessibility [45]. To bridge the efficiency gap between expert blind users and expert sighted users, this gap, if any, needs to be quantified. It requires a task-based comparative study between the two sets of users, similar to the gesture comparison study between blind and sighted users [28]. Such a study in future will help to inform the design of accessibility apps to achieve the desired parity.

5.1 Limitations of the study

We acknowledge several limitations of this work. First, due to COVID-19, we recruited participants with competence in computing technology, which is not a requirement to gain expertise in smartphone usage. A related homogeneity of our participants was that for them smartphone was the second digital platform of everyday use, after a desktop or a laptop (similar to participants of [45]). In comparison, for the vast majority of smartphone users in India, smartphones are the first digital device providing access to the Internet [1]. Second, the small sample size limited our analyses. A larger number of participants is required to identify broader trends. However, most prior longitudinal works logging PVIs smartphone data has been limited to ∼5 users [39, 45]. Third, our findings are limited to TalkBack in Android. For instance, two of our participants with prior iPhone experience praised VoiceOver response time and camera accessibility, but complained about iPhone's affordability and minimal support for local languages. As there is no prior work on understanding expert iPhone users with vision impairment, a direct comparison is not possible. However, we believe that our study methodology can be adopted to study expert iPhone VoiceOver users. Moreover, comparing and learning about mobile accessibility from the usage of the two most widely used smartphone OS's has future research potential. Fourth, 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 [22]. Fifth, due to COVID-19, the movement of our participants was minimal, hence we can not comment on their phone usage pattern on-the-move (which has been previously discussed in [43]). Finally, as all the participants were from India, the findings may not generalize globally. However, India has one-third of the world's blind population [13], making these findings valuable for a large target audience.

5.2 Design Considerations

Hardware Accessibility: Our participants successfully managed their security using the tactile fingerprint scanner, sidestepping the difficulties encountered in using PIN, pattern or password. However, an egregious development, that parallels the shift from physical keyboards to touch-screens in order to increase the screen space for sighted users, is that of recent phones removing fingerprint scanners to reduce cost (as mentioned by two participants), or shifting from tactile fingerprint scanners to in-screen fingerprint scanners. This will impact PVIs, requiring them to either find workarounds or compromise their phone's security. Diverse solutions are needed to enable PVIs to ensure protection of their passwords and personal information while they move to an increasingly digital world. Apart from security, differing hardware act as a barrier in switching smartphones for PVIs as pointed by our participants. For instance, variation in default touch screen sensitivity across smartphones require them to re-learn pressures required to perform gestures. Standardisation and specification of smartphone hardware, including minimal processing and memory requirements for smooth functioning of TalkBack, should be part of mobile accessibility guidelines.

Software Accessibility: Apps are designed predominantly for sighted users. In most cases, after the app design phase, accessibility is added by labeling buttons and adding image descriptions. Our participants found critical apps (like Paytm for financial transaction and Gaana for listening to music) missing even these basic accessibility measures. Complying with the accessibility guidelines is a good start, but is not enough to bring PVIs at par with sighted users with respect to app usage efficiency, which multiple expert users aimed for. To achieve that, apps need to be developed with screen reader access in mind. Developers need to propose alternate user interfaces optimized for screen reader use that can be toggled manually or automatically on detecting if a screen reader is enabled. For instance, automatically remove/reduce YouTube videos quality, game graphics, and website images, which consume most of battery and bandwidth. Such optimizations can reduce the lag for PVIs, thus helping them achieve speed neutrality with sighted users. These accessibility features should be part of advertisements and app/phone reviews, to make the PVIs aware of such progress and help them adopting it. Moreover, extending hardware accessibility, more apps should support fingerprint as password, to not only increase the functionality of the fingerprint scanner from just unlocking the device, but also to help narrow the gap between PVIs and sighted users.

AI+Crowdsourcing for Accessibility: To achieve software accessibility, in apps with missing button labels, we found participants meticulously labeling each button to make it accessible. However, currently these efforts are not scalable, as each individual performs the same labeling task on their phone. Also, with every update of the app, re-labeling may be needed. There should be a central repository where such labels can be uploaded to benefit a wider audience. The central repository should be able to perform an intelligent merge if different labels are associated with the same button. Future screen readers should also have capabilities to automatically label such unlabeled buttons using a combination of image processing and de-compilation of source code. It can take crowd help in up- or down-voting the auto-generated labels to finalize the labels. This data can help in further training and strengthening the AI models. Such a solution has potential to increase the software accessibility of a range of apps.

TTS engines: Smartphones act as the gateway to Internet access around the world [1], and for a vast majority of users, similar to our participants, English is not their first language. Even for the same language, the accent varies across geography. In spite of Google TTS offering local language and English in Indian accent, most of our participants used other TTS like Eloquence, eSpeak and Vocalizer, as they were not aware of Google TTS's offerings. Ideally, default TTS language and accent should be chosen based on the user's demography. Moreover, with the high prevalence of code-mixing [5], TTS needs to effectively support two or more mixed languages in the same text. Our participants used Auto TTS as a workaround reading code-mixed text. Future versions of TTS needs to be smart, especially for colloquial content and emoticons. For instance, to read '*goooooooood*', or '*:) :) :) :)*', instead of repeating the same content multiple times, TTS can convey the content followed by the number of times it has been repeated. Apart from language support, speech rate is important for TTS [43], especially for expert users, aspiring to parity with sighted users. A few participants reached the maximum allowed speech rate. Ideally the TTS engine should support higher speech rate to help expert users in pushing boundaries. Thus, working with code-mixed text, supporting advanced features, enabling higher speech rate, and effective ways of showcasing these capabilities, is needed for TTS.

Gestures and Training: We found expert users relying heavily on gestures to achieve maximum interaction speed. Furthermore, they demanded more gestures, including triple tap and 'o'/'v' gestures, associated with newer actions such as knowing battery level and dimming screen. On the other hand, novice and intermediate TalkBack users rely on explore by touch and soft-keys [45]. Prior work proposed sonification to teach touchscreen gestures [40]. Progressing from novice to expert is non-trivial and has a steep learning curve. As users constantly learn and improve their TalkBack usage skills, there could be a TalkBack training app running in the background tracking their progress. On reaching a certain level of expertise, the next set of gestures/shortcuts could be taught to them, maybe in a playful gamified manner. It can be similar to teaching new gaming skills after reaching a certain game level.

6 CONCLUSION

To understand the smartphone usage pattern of expert Android smartphone users with vision impairments, we recruited eight tech-savvy expert TalkBack users. We logged their smartphone usage data for four weeks, and conducted an interview to further contextualize the findings. Overall, we logged 209 days (~976 hours) of smartphone usage data. On analyzing the data, we found support for several prior findings, while we also discovered novel usage patterns or findings in contrast to prior beliefs, such as repurposing explore by touch for single-tap, inter-mixing different TTS engines and voices, and dependency on voice and external keyboard for text input. We also inquired about their learning journey, challenges faced, and accessibility tricks learned along the way. Based on these, we recommend design considerations informing the future of mobile accessibility.

ACKNOWLEDGMENTS

We would like to thank our participants for their time and patience.

REFERENCES

[1] Ravi Agrawal. 2018. *India Connected: How the Smartphone is Transforming the World's Largest Democracy.* Oxford University Press.

[2] 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

[3] 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

[4] Shiri Azenkot, Kyle Rector, Richard Ladner, and Jacob Wobbrock. 2012. PassChords: Secure Multi-Touch Authentication for Blind People. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility* (Boulder, Colorado, USA) *(ASSETS '12)*. Association for Computing Machinery, New York, NY, USA, 159–166. https://doi.org/10.1145/2384916.2384945

[5] Kalika Bali, Jatin Sharma, Monojit Choudhury, and Yogarshi Vyas. 2014. "I am borrowing ya mixing?" An Analysis of English-Hindi Code Mixing in Facebook. 116–126. https://doi.org/10.3115/v1/W14-3914

[6] Nikola Banovic, Christina Brant, Jennifer Mankoff, and Anind Dey. 2014. ProactiveTasks: The Short of Mobile Device Use Sessions. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Toronto, ON, Canada) *(MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 243–252. https://doi.org/10.1145/2628363.2628380

[7] Agathe Battestini, Vidya Setlur, and Timothy Sohn. 2010. A Large Scale Study of Text-Messaging Use. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (Lisbon, Portugal) *(MobileHCI '10)*. Association for Computing Machinery, New York, NY, USA, 229–238. https://doi.org/10.1145/1851600.1851638

[8] Jeffrey P. Bigham and Anna C. Cavender. 2009. Evaluating Existing Audio CAPTCHAs and an Interface Optimized for Non-Visual Use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1829–1838. https://doi.org/10.1145/1518701.1518983

[9] Jeffrey P. Bigham, Ryan S. Kaminsky, Richard E. Ladner, Oscar M. Danielsson, and Gordon L. Hempton. 2006. WebInSight: Making Web Images Accessible. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility* (Portland, Oregon, USA) *(Assets ’06)*. Association for Computing Machinery, New York, NY, USA, 181–188. https://doi.org/10.1145/1168987.1169018

[10] Syed Masum Billah, Vikas Ashok, Donald E. Porter, and I.V. Ramakrishnan. 2017. Ubiquitous Accessibility for People with Visual Impairments: Are We There Yet?. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI ’17)*. Association for Computing Machinery, New York, NY, USA, 5862–5868. https://doi.org/10.1145/3025453.3025731

[11] Maria C. Buzzi, Marina Buzzi, Barbara Leporini, and Amaury Trujillo. 2017. Analyzing visually impaired people’s touch gestures on smartphones. 76, 4 (2017), 1573–7721. https://doi.org/10.1007/s11042-016-3594-9

[12] Karen Church, Denzil Ferreira, Nikola Banovic, and Kent Lyons. 2015. Understanding the Challenges of Mobile Phone Usage Data. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI ’15)*. Association for Computing Machinery, New York, NY, USA, 504–514. https://doi.org/10.1145/2785830.2785891

[13] 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

[14] Rafael Jeferson Pezzuto Damaceno, Juliana Cristina Braga, and Jesús Pascual Mena Chalco. 2016. Mobile Device Accessibility for the Visually Impaired: Problems Mapping and Empirical Study of Touch Screen Gestures. In *Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems* (São Paulo, Brazil) *(IHC ’16)*. Association for Computing Machinery, New York, NY, USA, Article 2, 10 pages. https://doi.org/10.1145/3033701.3033703

[15] Rafael Jeferson Pezzuto Damaceno, Juliana Cristina Braga, and Jesús Pascual Mena-Chalco. 2018. Mobile device accessibility for the visually impaired: problems mapping and recommendations. *Universal Access in the Information Society* 17, 2 (01 Jun 2018), 421–435. https://doi.org/10.1007/s10209-017-0540-1

[16] Tao Deng, Shaheen Kanthawala, Jingbo Meng, Wei Peng, Anastasia Kononova, Qi Hao, Qinhao Zhang, and Prabu David. 2018. Measuring smartphone usage and task switching with log tracking and self-reports. *Mobile Media & Communication* (04 2018), 205015791876149. https://doi.org/10.1177/2050157918761491

[17] Trinh Minh Tri Do, Jan Blom, and Daniel Gatica-Perez. 2011. Smartphone Usage in the Wild: A Large-Scale Analysis of Applications and Context. In *Proceedings of the 13th International Conference on Multimodal Interfaces* (Alicante, Spain) *(ICMI ’11)*. Association for Computing Machinery, New York, NY, USA, 353–360. https://doi.org/10.1145/2070481.2070550

[18] Hossein Falaki, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. 2010. Diversity in Smartphone Usage. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services* (San Francisco, California, USA) *(MobiSys ’10)*. Association for Computing Machinery, New York, NY, USA, 179–194. https://doi.org/10.1145/1814433.1814453

[19] Nora Griffin-Shirley, Devender R. Banda, Paul M. Ajuwon, Jongpil Cheon, Jaehoon Lee, Hye Ran Park, and Sanpalei N. Lyngdoh. 2017. A Survey on the Use of Mobile Applications for People who Are Visually Impaired. *Journal of Visual Impairment & Blindness* 111, 4 (2017), 307–323. https://doi.org/10.1177/0145482X1711100402 arXiv:https://doi.org/10.1177/0145482X1711100402

[20] William Grussenmeyer and Eelke Folmer. 2017. Accessible Touchscreen Technology for People with Visual Impairments: A Survey. *ACM Trans. Access. Comput.* 9, 2, Article 6 (Jan. 2017), 31 pages. https://doi.org/10.1145/3022701

[21] Darren Guinness, Edward Cutrell, and Meredith Ringel Morris. 2018. Caption Crawler: Enabling Reusable Alternative Text Descriptions Using Reverse Image Search. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI ’18)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3173574.3174092

[22] 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

[23] 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

[24] 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

[25] Hernisa Kacorri, Sergio Mascetti, Andrea Gerino, Dragan Ahmetovic, Valeria Alampi, Hironobu Takagi, and Chieko Asakawa. 2018. Insights on Assistive Orientation and Mobility of People with Visual Impairment Based on Large-Scale Longitudinal Data. *ACM Trans. Access. Comput.* 11, 1, Article 5 (2018), 28 pages. https://doi.org/10.1145/3178853

[26] Shaun K. Kane, Jeffrey P. Bigham, and Jacob O. Wobbrock. 2008. Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-Touch Interaction Techniques. 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, 73–80. https://doi.org/10.1145/1414471.1414487

[27] 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

[28] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable Gestures for Blind People: Understanding Preference and Performance. 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, 413–422. https://doi.org/10.1145/1978942.1979001

[29] Ravi Kuber and Shiva Sharma. 2010. Toward Tactile Authentication for Blind Users. 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, 289–290. https://doi.org/10.1145/1878803.1878875

[30] Steven Landua and Lesley Wells. 414-418. The Merging of Tactile Sensory Input and Audio Data by Means of The Talking Tactile Tablet. *Proceedings of the EuroHaptics* (414-418), 1–3. http://www.touchgraphics.com/publications/eurohaptics-paper.pdf

[31] Barbara Leporini, Maria Claudia Buzzi, and Marina Buzzi. 2012. Interacting with mobile devices via VoiceOver: Usability and accessibility issues. 339–348. https://doi.org/10.1145/2414536.2414591

[32] Mingzhe Li, Mingming Fan, and Khai N. Truong. 2017. BrailleSketch: A Gesture-Based Text Input Method for People with Visual Impairments. 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, 12–21. https://doi.org/10.1145/3132525.3132528

[33] Sylvan Lobo, Ulemba Hirom, V. Shyama, Mridul Basumatori, and Pankaj Doke. 2017. Coping with Accessibility Challenges for Security - A User Study with Blind Smartphone Users. 3–22. https://doi.org/10.1007/978-3-319-68059-0_1

[34] Vikas Luthra and Sanjay Ghosh. 2015. Understanding, Evaluating and Analyzing Touch Screen Gestures for Visually Impaired Users in Mobile Environment. In *Universal Access in Human-Computer Interaction. Access to Interaction*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 25–36.

[35] Akhil Mathur, Lakshmi Manasa Kalanadhabhatta, Rahul Majethia, and Fahim Kawsar. 2017. Moving Beyond Market Research: Demystifying Smartphone User Behavior in India. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 82 (Sept. 2017), 27 pages. https://doi.org/10.1145/3130947

[36] Ted McCarthy, Joyojeet Pal, and Edward Cutrell. 2013. The "voice" has it: screen reader adoption and switching behavior among vision impaired persons in India. *Assistive technology : the official journal of RESNA* 25, 4 (2013), 222–229. https://doi.org/10.1080/10400435.2013.768719 24620705[pmid].

[37] Kyle Montague, André Rodrigues, Hugo Nicolau, and Tiago Guerreiro. 2015. TinyBlackBox: Supporting Mobile In-The-Wild Studies. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 379–380. https://doi.org/10.1145/2700648.2811379

[38] Meredith Ringel Morris, Annuska Zolyomi, Catherine Yao, Sina Bahram, Jeffrey P. Bigham, and Shaun K. Kane. 2016. "With Most of It Being Pictures Now, I Rarely Use It": Understanding Twitter's Evolving Accessibility to Blind Users. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5506–5516. https://doi.org/10.1145/2858036.2858116

[39] Hugo Nicolau, Kyle Montague, Tiago Guerreiro, André Rodrigues, and Vicki L. Hanson. 2015. Typing Performance of Blind Users: An Analysis of Touch Behaviors, Learning Effect, and In-Situ Usage. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 273–280. https://doi.org/10.1145/2700648.2809861

[40] Uran Oh, Shaun K. Kane, and Leah Findlater. 2013. Follow That Sound: Using Sonification and Corrective Verbal Feedback to Teach Touchscreen Gestures. 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 13, 8 pages. https://doi.org/10.1145/2513383.2513455

[41] João Oliveira, Tiago Guerreiro, Hugo Nicolau, Joaquim Jorge, and Daniel Gonçalves. 2011. Blind People and Mobile Touch-Based Text-Entry: Acknowledging the Need for Different Flavors. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee, Scotland, UK) *(ASSETS '11)*. Association for Computing Machinery, New York, NY, USA, 179–186. https://doi.org/10.1145/2049536.2049569

[42] Joyojeet Pal, Anandhi Viswanathan, Priyank Chandra, Anisha Nazareth, Vaishnav Kameswaran, Hariharan Subramonyam, Aditya Johri, Mark S. Ackerman, and Sile O'Modhrain. 2017. Agency in Assistive Technology Adoption: Visual Impairment and Smartphone Use in Bangalore. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5929–5940. https://doi.org/10.1145/3025453.3025895

[43] Gisela Reyes-Cruz, Joel E. Fischer, and Stuart Reeves. 2020. Reframing Disability as Competency: Unpacking Everyday Technology Practices of People with Visual Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376767

[44] André Rodrigues, Kyle Montague, Hugo Nicolau, João Guerreiro, and Tiago Guerreiro. 2017. In-Context Q&A to Support Blind People Using Smartphones. 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, 32–36. https://doi.org/10.1145/3132525.3132555

[45] André Rodrigues, Kyle Montague, Hugo Nicolau, and Tiago Guerreiro. 2015. Getting Smartphones to Talkback: Understanding the Smartphone Adoption Process of Blind Users. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 23–32. https://doi.org/10.1145/2700648.2809842

[46] M.C. Rodriguez-Sanchez, M. Moreno-Alvarez, Estefanía Martín, S. Borromeo, and Juan Antonio Hernández Tamames. 2014. Accessible smartphones for blind users: A case study for a wayfinding system. *Expert Systems with Applications* 1.854 (11 2014), 7210–7222. https://doi.org/10.1016/j.eswa.2014.05.031

[47] Anne Spencer Ross, Xiaoyi Zhang, James Fogarty, and Jacob O. Wobbrock. 2018. Examining Image-Based Button Labeling for Accessibility in Android Apps through Large-Scale Analysis. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS ’18)*. Association for Computing Machinery, New York, NY, USA, 119–130. https://doi.org/10.1145/3234695.3236364

[48] Daisuke Sato, Uran Oh, Kakuya Naito, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. 2017. NavCog3: An Evaluation of a Smartphone-Based Blind Indoor Navigation Assistant with Semantic Features in a Large-Scale Environment. 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, 270–279. https://doi.org/10.1145/3132525.3132535

[49] Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. 2003. When Computers Fade … Pervasive Computing and Situationally-Induced Impairments and Disabilities. (01 2003).

[50] Kristen Shinohara and Jacob O. Wobbrock. 2016. Self-Conscious or Self-Confident? A Diary Study Conceptualizing the Social Accessibility of Assistive Technology. *ACM Trans. Access. Comput.* 8, 2, Article 5 (Jan. 2016), 31 pages. https://doi.org/10.1145/2827857

[51] Caleb Southern, James Clawson, Brian Frey, Gregory Abowd, and Mario Romero. 2012. An Evaluation of BrailleTouch: Mobile Touchscreen Text Entry for the Visually Impaired. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services* (San Francisco, California, USA) *(MobileHCI ’12)*. Association for Computing Machinery, New York, NY, USA, 317–326. https://doi.org/10.1145/2371574.2371623

[52] Statista. 2017. *Market share of mobile operating systems in India from 2012 to 2019*. Retrieved Dec 1, 2019 from https://www.statista.com/statistics/262157/market-share-held-by-mobile-operating-systems-in-india/

[53] Google Android Team. 2020. *Navigate your device with TalkBack*. Retrieved Feb 7, 2020 from https://support.google.com/accessibility/android/answer/6006598

[54] Google Android Team. 2020. *Use global and local context menus*. Retrieved Feb 7, 2020 from https://support.google.com/accessibility/android/answer/6007066

[55] Google Android Team. 2020. *Use TalkBack gestures*. Retrieved Feb 7, 2020 from https://support.google.com/accessibility/android/answer/6151827

[56] H. Tinwala and I. S. MacKenzie. 2009. Eyes-free text entry on a touchscreen phone. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*. 83–88.

[57] Gregg C. Vanderheiden. 1996. Use of Audio-Haptic Interface Techniques to Allow Nonvisual Access to Touchscreen Appliances. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 40, 24 (1996), 1266–1266. https://doi.org/10.1177/154193129604002430 arXiv:https://doi.org/10.1177/154193129604002430

[58] Radu-Daniel Vatavu and Jean Vanderdonckt. 2020. What Gestures Do Users with Visual Impairments Prefer to Interact with Smart Devices? And How Much We Know About It. In *Companion Publication of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS’ 20 Companion)*. Association for Computing Machinery, New York, NY, USA, 85–90. https://doi.org/10.1145/3393914.3395896

[59] Emily A. Vogels. 2019. *Millennials stand out for their technology use, but older generations also embrace digital life*. Retrieved Feb 1, 2020 from https://www.pewresearch.org/fact-tank/2019/09/09/us-generations-technology-use/

[60] W3C. 2015. *Mobile Accessibility: How WCAG 2.0 and Other W3C/WAI Guidelines Apply to Mobile*. Retrieved Feb 1, 2020 from https://www.w3.org/TR/mobile-accessibility-mapping/

[61] Eric Whitmire, Mohit Jain, Divye Jain, Greg Nelson, Ravi Karkar, Shwetak Patel, and Mayank Goel. 2017. DigiTouch: Reconfigurable Thumb-to-Finger Input and Text Entry on Head-Mounted Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 113 (Sept. 2017), 21 pages. https://doi.org/10.1145/3130978

[62] Jian Xu, Syed Masum Billah, Roy Shilkrot, and Aruna Balasubramanian. 2019. DarkReader: Bridging the Gap Between Perception and Reality of Power Consumption in Smartphones for Blind Users. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS ’19)*. Association for Computing Machinery, New York, NY, USA, 96–104. https://doi.org/10.1145/3308561.3353806

[63] Yu Zhong, Walter S. Lasecki, Erin Brady, and Jeffrey P. Bigham. 2015. RegionSpeak: Quick Comprehensive Spatial Descriptions of Complex Images for Blind Users. 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, 2353–2362. https://doi.org/10.1145/2702123.2702437

[64] Yu Zhong, T. V. Raman, Casey Burkhardt, Fadi Biadsy, and Jeffrey P. Bigham. 2014. JustSpeak: Enabling Universal Voice Control on Android. In *Proceedings of the 11th Web for All Conference* (Seoul, Korea) *(W4A ’14)*. Association for Computing Machinery, New York, NY, USA, Article 36, 4 pages. https://doi.org/10.1145/2596695.2596720