# Exploring the Experiences of Individuals Who are Blind or Low-Vision Using Object-Recognition Technologies in India

Gesu India
Swansea University
Swansea, Wales, United Kingdom
g.india.2132154@swansea.ac.uk

Simon Robinson
Swansea University
Swansea, United Kingdom
s.n.w.robinson@swansea.ac.uk

Jennifer Pearson
Swansea University
Swansea, Wales, United Kingdom
j.pearson@swansea.ac.uk

Cecily Morrison
Microsoft Research
Cambridge, United Kingdom
cecilym@microsoft.com

Matt Jones
Swansea University
Swansea, United Kingdom
always@acm.org

## Abstract

Assistive technologies, such as smartphone-based object-recognition (OR) apps, provide visual assistance to people who are blind or low-vision to enable increased independent participation in society. While previous research has explored the functional accessibility of object-recognition technologies, little attention has been given to their social accessibility, particularly in interdependent socio-cultural contexts of the Global South. Through a mixed-methods approach, employing a seven-day diary study followed by one-on-one interviews with seven OR app users in India, we explore their experiences in depth. Our findings highlight the nuances of what interdependence looks like in a multicultural, Indian society, as people navigate public and private spheres with a camera-based assistive technology designed for independent, western contexts. We argue for the necessity to design assistive technologies following the interdependence framework that accommodates the social and cultural context of the Global South. Additionally, we propose design guidelines for assistive technologies in community-oriented societies, emphasizing community-centered approaches, cultural alignment, and locally adaptable designs.

## CCS Concepts

• **Human-centered computing → Accessibility**.

## Keywords

Object-recognition, vision impairments, human-computer interaction, assistive technologies, low-resource environments

**ACM Reference Format:**
Gesu India, Simon Robinson, Jennifer Pearson, Cecily Morrison, and Matt Jones. 2025. Exploring the Experiences of Individuals Who are Blind or Low-Vision Using Object-Recognition Technologies in India. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3706598.3713107


*CHI '25, Yokohama, Japan*

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713107

## 1 Introduction

In the last decade, numerous camera-based technologies have been developed to provide visual assistance to people who are blind or low-vision to enable increased independent participation in society. These include: the facilitation of remote sighted assistance (VizWiz app [12], BeMyEyes app [19], Aira app [6]) as well applications underpinned by computer vision (Talking Goggles [20], TapTapSee [55]). Part of this suite of tools are 'object recognition' (OR) apps that recognize 'things', such as doors, chairs, or a hat, e.g. Lookout app [36] and Seeing AI [38]. Most OR apps have been developed in North America, with the datasets and design paradigms reflecting a socio-cultural context that values individual independence, emphasizing the use of technology to achieve personal autonomy and self-sufficiency.

However, with the global increase in smartphone usage, these apps are becoming available free-of-charge throughout the world; they are moving into places where social/cultural norms differ from the original technology-development context. The Global South presents a significantly different socio-cultural context and view towards independence than the countries of the Global North, valuing more highly interdependent interaction over personal autonomy [34, 43, 63]. As the Global South is home to 90% of the world's population of people with blindness [41], this prompts reflection on the suitability of technologies designed for independence in the context of cultures that emphasize interdependence. Indian, for example, values collective well-being, harmony and interconnectedness above an individual's needs [58]. Interdependence, rather than independence, shape interpersonal dynamics and thus how technologies may be used in daily life.

To address this gap, this paper presents a study of the socially situated use of smartphone-based object recognition apps, or OR apps, by individuals who are blind or low-vision in India. We conducted a seven-day diary study during which we tracked the use of OR apps by seven adults, followed by a semi-structured in-depth interview exploring their current and past experiences with OR apps. This included motivations, roadblocks, and the varied socio-cultural expectations associated with the use of camera-based assistive technologies in both public and private spaces. Our findings address our research question: **"How are OR apps, originally designed to support independence, being adapted and employed within a cultural context of interdependence, such as India?"**

We highlight how people navigate public and private spheres with a camera-based assistive technology in a social context that values interdependence in the multicultural Indian society. We identify the similarities and differences in their experiences to those reported in the similar studies conducted in the Western countries, arguing for the necessity to design assistive technologies that accommodate the socio-cultural context of the Global South. We reframe the Interdependence framework for the perspective of a community-oriented society, like India, and suggest design guidelines for developing AI systems for societies that value interdependence.

## 2 Prior Work

In this section, we overview prior research on: a) object-recognition technologies for people who are blind or low-vision, and b) innovative community-led adaptations of technologies developed for high-resource contexts in the interdependent, low-resource contexts of the Global South.

### 2.1 Object-recognition technologies for people who are blind or low-vision

Individuals who are blind or low-vision use sensory cues such as smell, sound, and touch, along with mental maps, to navigate their surroundings. However, for tasks where visual identification is crucial, such as distinguishing food items with similar tactile sensation, or when information gathered from non-visual senses and mental maps isn't enough, like finding an umbrella in the house that was moved by someone else or coordinating outfits for a job interview, sighted assistance is sought. Given the limitations (time, availability, context of task, etc.) of accessing sighted help in-person, people who are blind or low-vision frequently turn to social media platforms like Facebook for remote sighted assistance from their loved ones [14, 18]. However, the associated social costs—such as coming across as "too needy," "vulnerable," or a "spammer" [65]—can lead them to actively limit such public interactions.

To address this, platforms like VizWiz [12], Be My Eyes [19], and others [6, 24] offer anonymous, remote sighted assistance. VizWiz [12] enables users to send images with questions to sighted friends or crowd workers on Amazon Mechanical Turk [9], while Be My Eyes connects users to untrained volunteers via free audio/video calls for tasks like shopping and navigation [10]. However, remote assistance can be expensive, involve connection delays [10], and raise safety and privacy threats when interacting with strangers [4].

Advancements in computer vision and machine learning have enabled object-detection and recognition technologies to partially or fully automate visual assistance in Remote Sighted Assistance (RSA) applications [13, 30, 47]. For instance, Vizlens [21] uses crowd workers to label elements of an inaccessible interface, like an information kiosk, for computer vision tasks, while Camfind [17] and Taptapsee [55] combine human operators and image databases for assistance. Smartphone-based apps like Lookout [36], Seeing AI [38], and Envision AI [5] rely solely on trained AI models to detect, recognize, and identify objects and extract information like object size, shape, color, and more, providing output through auditory or haptic channels. Such apps also offer features like face recognition and Optical Character Recognition (OCR), making them a cost-effective, portable, and feasible alternative to wearable devices like FReAD [2]. Despite potential errors, OR apps remain popular due to their affordability and accessibility, thus this paper focuses on how these AI-driven apps, designed for high-resource settings, are adopted within the low-resource context of India.

Prior work has explored how people who are blind or low-vision use camera-based assistive technologies for hobbies like photography and practical visual tasks like OCR or barcode scanning [1, 31]. A major challenge is the lack of accessible feedback on camera positioning, lighting, and output quality, which is crucial for technologies like OR apps that rely on framing objects accurately. Brady et al. [14] found that over 82% of images submitted by blind or low-vision users to visual question-and-answer platforms had errors such as blurring, poor lighting, or improper framing and composition. While features like VoiceOver on iPhones provide feedback on camera positioning and framing, such accessibility features are not yet widely available on Android devices. This is particularly significant as Android dominates the Global South, including India, where it accounted for 70.16% of smartphone users as of December 2023 [25, 53]. To overcome these issues, individuals who are blind or low-vision resort to do-it-yourself workarounds [31]. For instance, OCR app users [46] use techniques like distancing, positioning and aligning the camera with respect to their body and the object, though the effectiveness of such methods vary with the task, such as recognition of small-sized versus large-sized objects.

The lack of accessible feedback also raises privacy concerns, as users may inadvertently expose Private Identifiable Information (PII), especially when using apps like Aira [6] and Be My Eyes [19] for 'anonymous' remote assistance. For instance, over 10% of the images in the Viz-Wiz dataset [22] contained private visual information. Despite users taking precautions, like avoiding sensitive information in the camera's view and controlling camera movements, privacy risks remain inevitable, particularly when sensitive details are required for accessing sighted assistance for tasks like finding private medicines [3]. Social context also plays a role in users' preferences, with assistance sought from family, friends, or strangers, both in-person and online [8]. Additionally, PII exposure prevalent in object-recognition datasets also leads to manual editing or data loss [22, 56]. To address this, initiatives like Zhang et al.'s [66] ImageAlly use a human-AI hybrid approach to detect and redact private content, and datasets are being developed to train AI models to automatically identify PII in images captured by blind or low-vision users [22, 48].

Beyond privacy, social perception impacts the adoption of camera-based assistive technologies. Users worry about breaching bystanders' privacy [7, 8, 35] or being perceived differently in social situations [49, 50]. Lee et al. [35] found that people who are blind or low-vision avoid technologies that attract unwanted attention, while Shinohara et al. [50] noted that assistive technology use in public spaces can prompt intrusive questions from sighted individuals about the user's disability and the technology's purpose. While the functional accessibility of object-recognition apps has been studied, research on their social acceptability, especially in non-western cultures like India, remains underexplored and is crucial for understanding the community-focused societal dynamics affecting their use and adoption in the Global South.

### 2.2 Community-led innovations in assistive tech in the Global South

India, housing one-third of the world's blind or low-vision population [41], forms a significant market for assistive technologies. As several assistive technologies are designed for use in the high-resource settings, prior work reports innovative approaches to make such systems work in low-resource settings, where necessity drives innovation. For instance, a community of people who are blind or low-vision in India facing the acute shortages of braille and high-quality audio books for many subjects, leaving students scrambling for ways to continue their education, established an informal network of peer-produce and share audio content via Bluetooth, memory cards and CDs [63]. India et al. [28] adapted Code Jumper [33, 40], originally developed for computational thinking curriculum in the UK schools, for teaching in schools for the blind (i.e., schools exclusively for children who are blind or low-vision) in India, where resource constraints limit computer education to basic tasks like using MS Office, checking email, and browsing the internet [29]. While the Global South faces economic disparities and limited access to education and technology, these examples highlight the potential of community-driven solutions to bridge these gaps. As Pal et al. [42] note, assistive technologies can open doors to opportunities and help individuals realize their aspirations.

Community values and support play a pivotal role in driving innovations in assistive technology. While such technologies are typically designed to promote independence [11], communities rooted in a culture of interdependence often devise innovative alterations to adapt them for use to their unique socio-cultural context. Hurst and Tobias [26] highlight how personalization, passion, and cost motivate individuals with disabilities to create or adapt assistive technologies, resulting in solutions that work better and have higher adoption rates than existing ones. Inspired by disability justice activists like Mia Mingus, who argue that the societal concept of independence is fundamentally flawed, Bennett et al. [11] proposed the 'Interdependence framework' for designing assistive technologies. The framework advocates for a paradigm shift in the field of assistive technology design, prioritizing collaborative access and recognizing the vital contributions of individuals with disabilities. This approach complements the traditional focus on independence in assistive technology design.

In this paper, we embrace the lens of 'Interdependence framework' to explore the use and adaptation of object-recognition (OR) apps within the diverse socio-cultural context of India.

## 3 Study Design

Our study was designed to address two key research objectives:

(1) A comprehensive seven-day diary study aimed at delving into the mental models of OR app users who are blind or low-vision, collecting examples of OR tasks and prompt real-world use, that could be later used as prompts for reflection and discussion during the interviews.
(2) Followed by semi-structured one-on-one interviews to further explore how these apps assimilate into their daily routines and the intricacies of using them in the presence of other people.

### 3.1 Study Methodology

The entire study took place remotely. The recruitment message was circulated on online social groups, including accessibility-focused WhatsApp groups. The message invited participation from adults who were legally blind or low-vision, spoke English and/or Hindi (a language spoken by 43.63% of the Indian population [60]), had a personal smartphone with internet access, and at least six months of experience using OR apps. Participants' consent was registered via email, and their demographic information like age, gender, schooling, experience using smartphones and screen readers, was recorded via an online survey.

For the diary study, participants were asked to complete at least one object-recognition task per day using the OR app they were currently using. To simulate real-life use, participants were asked to come up with their own tasks, providing insight into the variety of tasks and the physical and social contexts of the tasks. As apps like Lookout and Seeing AI offer additional features, like text recognition, we provided examples of object recognition tasks to our participants to distinguish them and ensure clarity. To get a simultaneous overview of the kinds of tasks being performed, participants were asked to submit a 'diary entry' for each task attempted. This diary entry included a screenshot of the app screen, ideally captured immediately after the task completion to ensure the object was in frame, and an audio message detailing task goal and the app's effectiveness in assisting them. These entries provided insights into the task goal and user experience (via the audio messages), as well as contextual information such as object selection, placement, distance from the user/camera, object size, and environmental conditions like lighting and clutter (via the screenshots). It's important to note that due to potential camera movements while taking a screenshot, the information gathered from screenshots may be imprecise but still valuable for gathering additional insights during the follow-up interviews. All participants confirmed being familiar with the process of taking a screenshot on their phones. Considering the privacy concerns emerging from sharing captured images, participants were informed that sharing screenshots was optional and be only shared if they felt comfortable and confident the image contained no private identifiable information (PII). As the diary study spanned seven days and many participants were working professionals or students, a daily WhatsApp reminder was sent at noon to prompt completion of the OR task(s), if not completed already.

After completing the diary study, we conducted one-on-one semi-structured interviews with all participants to explore their past experiences with OR apps. We discussed the challenges, motivations, and decision-making processes related to using these apps. All interviews were conducted remotely via Zoom in English. Each interview was entirely audio-based, lasted for an average of 45-min, and was audio-recorded with participant consent. Zoom's transcription feature was used, and transcripts were manually reviewed and corrected by one of the authors. The study was approved by our institutional ethics board. Upon completion of study, each participant received a gift voucher worth INR 1000 (~12 USD).

**Table 1: Demographic information of study participants, including level of visual impairment, smartphone and screen reader usage, and object-recognition apps used in past or currently (i.e., at the time of study). The table columns include: Participant ID, Sex, Age, Visual Impairment (VI) level, Smartphone usage (years), OR app (using currently), OR apps (used in past), Number of attempted tasks, and Number of screenshots shared. Abbreviations used: TB (Totally Blind), PS (Partially Sighted).**

| P.Id. | Sex | Age | Level of VI | Smartphone usage(years) | OR app (using currently) | OR apps (used in past) | Number of attempted tasks | Number of screenshots shared |
|---|---|---|---|---|---|---|---|---|
| P1 | F | 27 | TB | 8 | SuperSense | Lookout, SuperSense, Eye-D Pro | 7 | 4 |
| P2 | M | 24 | TB | 10 | Seeing AI | Seeing AI | 10 | 10 |
| P3 | F | 24 | TB | 7 | SuperSense | Supersense | 15 | 18 |
| P4 | F | 45 | PS | 10 | Seeing AI | Eye-D Pro, TapTapSee, Seeing AI | 10 | 12 |
| P5 | M | 28 | TB | 10 | Lookout | Lookout | 7 | 11 |
| P6 | M | 39 | PS | 10 | SuperSense | Eye-D Pro, SuperSense | 13 | 13 |
| P7 | M | 32 | TB | 15 | Lookout | Seeing AI, Lookout | 5 | 2 |

### 3.2 Data Analysis

For analyzing task-related data from the diary study, several factors were considered, including the 'task goal', 'natural tasks (i.e., a real need for OR apps) or engineered (performed for the sake of the study)', 'object of interest', 'contents of screenshot', 'quality of screenshot', 'audio message', 'task failed or successful', and 'reason for task success/failure'. Using Thematic Analysis [16], interview transcripts and audio messages were open coded by two authors. Codes were refined iteratively, until a consensus was reached among the authors. Over the analysis, coding plans were discussed, preliminary codebooks were developed and reviewed, codes were refined/edited, and categories and themes were finalized to understand the user behavior. Preliminary codes (like 'camera placement', 'self-search techniques', 'use of RSA apps', 'using OR apps in public places', 'concerns for bystander's privacy', 'feeling conscious using OR apps with sighted people', 'feeling confident using OR apps with blind/low-vision people', etc.) emerged to codes like 'recommendations for OR app design', 'alternatives to the use of OR apps', 'disability stigma', 'concerns for self/device safety', etc., which were further grouped into three broad themes: 'disability stigma and interpersonal relationships', 'safety and privacy concerns with OR apps in public places', and 'usability and accessibility challenges with OR apps'.

### 3.3 Participant Demography

Seven adults (five totally blind) who were legally blind or had low-vision participated in our study. The average age of our participants was $31.3 \pm 7.9$ years. Three participants identified as cis-women and the remaining as cis-men. All participants spoke English and multiple Indian languages. Three participants resided in tier-I cities, three in tier-II cities, and one in a tier-III city of India, as per the Government of India's three-tier city classification system [62] where tier-I cities have higher populations, more developed socio-economic schemes, and better technological infrastructure compared to tier-II and tier-III cities. All participants were all full-time working professionals, with three holding graduate and four holding post-graduate degrees. Five participants attended integrated schools, and two had attended schools for the blind. All had access to stable internet and were smartphone users with screen readers for an average of $9.6 \pm 1.6$ years.

## 4 Findings

In this section, we report the findings from the diary study and qualitative interviews.

### 4.1 Findings from the Diary Study

Overall, participants attempted an average of $9.6 \pm 3.6$ object-recognition tasks over seven days. P7 missed two study days due to health reasons. Among the OR apps used for the study, SuperSense [54] was used by three, Seeing AI [38] by two, and Lookout app [36] by two. Table 1 details participants' current and past OR app usage.

Based on participant's task goals, the object-recognition tasks during the diary study were categorized into three groups: object-detection (29 tasks), object-finding (25 tasks), and scene exploration (13 tasks). Of the 67 tasks performed, the majority (41) were deemed unsuccessful by our participants, indicating insufficient support from the OR apps for task completion. This was primarily due to a) object misrecognition (22 tasks), b) lack of feedback (ten tasks), c) failed to detect the object of interest while detecting nearby objects (seven tasks), or d) vague descriptions (two tasks). Unsuccessful tasks included 20 object-finding, 14 object-detection, and 7 scene exploration tasks.

Participants chose a variety of objects for the tasks, categorized by the authors as follows: Food items (6), Furniture (4), Laptop (4), Water bottle/flask (4), Earphones (3), Keys (3), Handbag (3), Plants (3), AC remote (2), Cardboard boxes (2), Mask (2), Medicine (2), Phone charger (2), Switchboard (2), Beard trimmer, Currency notes, Documents, Earrings, Elliptical exercise device, Folders, Game coins, Handkerchief, Mobile phone, Oil bottle, Pair of scissors, Pen, Salt jar, Spoon, Towel.


**Figure 1: Some app screenshots shared by participants during the diary study, with text description by the authors. From left to right: 1(a) A heavy food bag on a weighing scale in an Indian grocery store; 1(b) Top-view of a cluttered table drawer with a folded white cane, ink stamp, and stationery; 1(c) Women's leggings of various colors stacked inside an Indian clothing store; 1(d) Close-up top-view of a steel flask, showing half of the lid.**

During the study, participants attempted both naturally occurring tasks (e.g., using the OR app to find a dropped surgical mask or exploring a guest room at friend's house) but majority were engineered by our participants for the study (e.g., scanning items in a clothing store, looking for a specific object in a cluttered space, finding a chair in a room with low lighting, and more). Participants shared an average of 10 ± 5.4 app screenshots, with a total of 70 screenshots submitted. Of these, five (out of 70) contained private or sensitive information, such as bank papers with participant addresses, fully or partially captured in the image. Figure 1 displays some of the app screenshots shared by participants, with text descriptions provided by the authors.

On average, our participants had been using an OR app for a minimum of ~2.5 years. They reported using OR apps occasionally in daily life, primarily for text-recognition tasks (e.g., reading text on a laptop screen when the screen reader stops working or reading medicine labels). The following section presents insights from one-on-one interviews, where participants shared their overall experience using OR apps in daily life.

## 4.2 Findings from the Qualitative Interviews

Participants reported that the use of OR apps in daily life was context-dependent, typically in unfamiliar environments or when sighted assistance was either unavailable (e.g., busy friends or family) or least preferred (e.g., surrounded by strangers, not wanting to bother others). In familiar environments, participants preferred non-visual sensory information and spatial mental maps for object-recognition over using OR apps. Access to guidance from sighted friends, family members, or peers was also quicker and easier in familiar settings. Based on the interviews, we present the findings on the use of OR apps by people who are blind or low-vision, categorized into three themes.

*4.2.1 Disability stigma and inter-personal relationships.* Social stigma around disability and prevalent stereotypes in the Indian society significantly influenced participants' behaviors and experiences with OR apps, especially when accompanied by other people. Five out of seven participants felt self-conscious, uncomfortable, and often "pretentious" for using OR apps around their sighted friends and family members. As participants used OR apps more often than usual during the study, they feared being perceived as "trying hard" to be "too independent". Many participants' sighted friends and family members remained ignorant of the need for assistive technologies for a person with disability and viewed their use of OR apps as a threat to their closeness with the participants. Some participants felt pressured to rely on sighted assistance instead of using OR apps. To navigate this, some participants dedicated themselves to the unpaid, often unrecognized effort of explaining the nature and value of assistive technology, while others, weighed down by frustration, chose to avoid the conversation entirely.

Prior research in HCI and Accessibility has highlighted the negative impact of social stigma surrounding disability on the use of assistive technologies by people who are blind or low-vision [35, 49, 50, 57], and found that it can negatively affect their behavior, self- perception, confidence, and opportunities for social engagement. The stigma often discourages individuals from fully embracing their disability identity in public spaces. Our participants, too, navigated the tension between their access needs and interpersonal relationships, feeling compelled to seek assistance from family members to showcase it as a "sign of trust". For instance, P5, who has been cheated in past with financial transactions, relies on OR apps to verify groceries and reading shopping receipts (using the apps' Optical Character Recognition feature), feels socially compelled to not use it at a local grocery store run by a sighted family friend. It's intended to not appear as overly independent and preserve the "good relationships maintained over the years". As P5 puts it, using the app in that context felt like "breaking someone's trust". This aligns with prior work [35], like Shinohara et al.'s findings [50], that assistive technologies, while functionally accessible, are socially inaccessible in such situations. Likewise, P4, who also sought to use the apps for grocery shopping, felt unsure of how to explain the technology to sighted individuals unfamiliar with it, leading to avoidance. *"What do I say when they (shop-owner) ask me about it? What is it? How is that situation managed? How does one handle it?"* In such scenarios, participants often chose to seek assistance from sighted loved ones instead of using the OR apps, despite the app's functional value.

The discomfort around using OR apps was felt mostly in the company of sighted peers and not mirrored in company of friends/peers who are blind or low-vision. Participants felt at ease, even curious and excited, to use a novel assistive technology with others who shared similar experiences. For them, assistive tech was seen as a necessity and often served as a tool for social bonding. For instance, P2 liked using OR apps with his friends who are blind or low-vision at a restaurant to explore the items on the table. Similarly, P7 felt supported using OR apps to explore nearby scenes, when in company with friends who are blind or low-vision.

*4.2.2 Safety and privacy concerns with the use of OR apps in public places.* Common public spaces in India, such as parks, bus stops, railway stations, temples, and shops, can be overcrowded and noisy, making them a challenging and overwhelming space for people who are blind or low-vision [27, 51]. As, P7 said, *"As a blind person, I'm always overwhelmed in public spaces because of the crowd and noise, and do not have the mental bandwidth to use the apps."* The noise often led to sensory overload, causing participants to avoid using OR apps in crowded areas. Furthermore, since OR apps provide audio feedback, participants had to use earphones, which many avoided for safety reasons to maintain spatial awareness in public spaces.

All participants were white cane users and often found it challenging to use both a white cane and a smartphone, simultaneously. When using an OR app, participants held the cane in one hand and the smartphone in the other, extending their arm away from their body to point the camera at the object or scene of interest. In crowded public spaces, the lack of space made this difficult. Additionally, the necessity of having both hands occupied heightened their sense of vulnerability in unfamiliar environments. As P2 says, *"It's not easy to stand there, take you phone out, scan for a while, and then see the results."*

Participants also felt anxious holding the phone away from their body due to safety concerns in public spaces, particularly in crowded areas. Four out of seven participants reported fear of theft while using their phones for object recognition. The risk of losing valuable items like smartphones to theft in public spaces, combined with the inaccessibility of pavements and roads in India [27, 51], added to these challenges. A participant from our pilot study recounted a theft at a bus station, where a sighted individual, pretending to help with directions, stole their phone. Bus stations in India are usually crowded and lack accessible markings on the floor for blind or low-vision to safely and comfortably walk through and explore [27]. Safety threats in public spaces were particularly exacerbated for participants with marginalized identities. For instance, P4, a woman, felt more vulnerable to such threats than her male peers and preferred avoiding smartphone use in public. As P2 noted, *"First of all, it's not comfortable asking somebody else for help in such a scenario. Also, what I mean is it also depends on what has fallen. If your cash has fallen, your wallet, your ring. If it's expensive, you cannot use the app nor ask for help."* As reported in prior work [3], these safety concerns led our participants to a deadlock or to a continual negotiation between their physical and device safety and access needs.

Participants also felt concerned for breaching a bystander's privacy when using OR apps, a sentiment echoed in previous studies [7]. Using an OR app typically involves pointing the smartphone camera at an object/scene for 3-5 seconds, often longer, while waiting for the app to recognize. During this time, participants feared being perceived as "creepy" by sighted bystanders and being misunderstood as "recording in public spaces". As a result, three out of seven participants avoided using cameras (and camera-based apps) in public spaces. Interestingly, this concern also arose in relatively safer, familiar environments like indoor office spaces. As P5 mentioned, *"Let's see, a blind person enters your office and stands in front of you pointing his camera at you. You'd think a hundred wrong things about him. And the app takes one minute to scan the object, you'd think, 'Oh God, why is he scanning me?'"*

In line with Akter et al.'s [7] work on augmented reality glasses, our participants prioritized protecting the privacy of sighted bystanders. While some of this anxiety could be credited to the novelty of the device's form factor and the overt nature of the smartphone camera, our participants shared similar concerns about using a smartphone camera around sighted individuals in public. The concern was also echoed by our male participants who have used OR apps in the past in company of female sighted strangers. As P6 recounted,*"Once, this lady approached me at the railway station when she saw me using my phone. (she) asked me if I was taking a picture or video of her... I had used this app before at the same station, but this incident scared me. Now, I avoid such things."*

We also asked participants to compare their experiences with OR apps to human-powered object-recognition technologies, such as remote sighted assistance apps. Five participants occasionally used the BeMyEyes app and mentioned the unique privacy challenges posed by sharing their camera view with sighted strangers. Similar concerns have been reported in the Global North context [8, 10]. For example, three out of five participants expressed feeling self-conscious about their appearance and the surrounding environment during video calls with sighted volunteers. One female participant noted, *"The problem is sometimes you're not dressed properly. If you have to call on Be My Eyes, then that volunteer will look at you, and that's a bit embarrassing."* Furthermore, P1 feared unknowingly disclosing PII and felt OR apps were often the safer option to use. In line with Akter et al. [7], two out of five participants preferred sharing sensitive information with AI-based tools rather than human-powered ones. As noted earlier (in Section 4.1), 7% of the submitted screenshots (5 out of 70) contained PIIs, similar to [22], where 10% of images submitted by people who are blind or low-vision contained sensitive information. When asked if they were aware of PII in the images before sharing, two participants were unaware, while the others were aware and trusted the researchers with their information. Despite study instructions emphasizing to not share screenshots with potential PIIs, these participants chose to share the images.

Overall, various safety and privacy related concerns associated with the use of OR apps, a camera-based assistive technology, prohibited our participants from fully taking advantage of them in public spaces.

*4.2.3 Usability and accessibility challenges with OR apps.* Participants provided feedback on the various usability and accessibility challenges faced while using their respective OR apps, sharing recommendations on improving the overall user experience with such apps for people who are blind or low-vision.

Object-recognition tasks and screenshots: Out of the 67 tasks performed in the diary study, 41 were considered unsuccessful by our participants. To understand the scenarios leading to these failures, we analyzed the app screenshots submitted for each unsuccessful task. Out of 70 screenshots (60 were unique), 33 belonged to unsuccessful tasks, out of which only 14 were clear, well-lit images fully containing the object-of-interest. The remaining 19 screenshots showed issues with photo quality, with ten having the object partially or completely out of the frame, three with the object too close to the camera, three with poor lighting, two with hand occlusion, and one blurred image. None of the participants reported being aware of these photo quality issues when submitting the screenshots, with a few emphasizing the need for accessible feedback on object placement, lighting conditions, and more. P1, for instance, assumed the camera was focused on the object (e.g., a chair) when it was too close to capture it fully. P4, who reported all (ten) tasks as unsuccessful, was surprised to learn that in five cases, the object-of-interest was entirely out of the frame. This suggests that the lack of accessible feedback from the OR apps regarding camera positioning, object placement, and lighting conditions may have contributed to the unsuccessful completion of tasks.

Adapting techniques for effective OR app use: Participants reported using different techniques and strategies, as stated in [46], when using camera-based assistive tech. For example, when looking for small objects like eyeglasses on a table among larger items like a water bottle, P3, P4, and P5 found that bringing the camera closer to the surface helped. P5 shared,*"Surprisingly what I noted was, when the device was farther from the table it only recognized the larger objects such as my laptop, mobile phone, and not my eyeglasses."* P5 also found rotating the eye drop bottles in front of the camera to identify the right one as an "interesting learning". P7 also tried moving around and showing the camera different sides of his elliptical exercise bike, though the app ended up only identifying its individual units (such as display unit, seat).

Feedback on inaccuracy and delay in OR app results: The inaccurate and delayed results were the most common issue faced by the participants, making it difficult to use the apps, particularly in public places where avoiding bystanders' attention is usually a priority. Participants also struggled to understand how the apps work and felt "clueless" when the apps performed differently with changing environments.

Challenges with cultural specificity in OR app recognition: Participants noticed the apps performed inadequately, particularly when the object-of-interest was unique to the Indian set-up/culture and not common to western cultures. For instance, P5, a long-time Lookout user, found the app "less useful" in India but appreciated its ability to recognize Indian foods like 'poha' (a popular Indian dish made from flattened rice [61]). Similarly, P3 noted that improved recognition of Indian kitchen items could better support the users in India. *"Usually, the visually impaired prefer to be independent and do things on their own, and we do label things, touch and feel things, and I know nobody takes their phone into the kitchen, but it'd be great if the app can recognize an item from my kitchen."*

Enhancing scene description and real-time feedback: When scanning scenes, OR apps often missed the relative positioning of objects, which participants found crucial for building mental maps and locating items. For example, a white coffee mug near books on a table was described simply as a "white coffee mug" and "table", omitting the books and their spatial relation. Prior research highlights the value of adding descriptive results to object recognition [7], especially when describing scenes.

Five participants, who occasionally used BeMyEyes [10], preferred its human-assisted guidance for providing detailed descriptions, often also helping adjust objects in the camera frame, a common problem faced among OR app users [1, 31]. *"So, somebody would look at your object and they would ask you tilted pitch towards right, left, or move the device slightly to the top right, etc. That gives us an idea about where the camera is focusing."* (P3). Participants also recommended adding real-time audio feedback for dynamic scenes, reducing the need to capture multiple pictures. P2 enjoyed using OR apps in places like train stations or coffee shops to learn about mundane activities but found continual photo-taking inconvenient.

## 5 Discussion

This paper explores the use of object-recognition (OR) apps by seven blind or low-vision adults in India. Participants engaged in a seven-day diary study, attempting at least one OR task daily with an app of their choice, averaging $9.6 \pm 3.6$ tasks and sharing $10 \pm 5.4$ screenshots. Follow-up interviews revealed challenges related to safety, privacy, and interpersonal relationships in public and private settings. While our findings offer valuable insights, the study has several limitations: our participants were from relatively privileged backgrounds, primarily English-speaking and highly educated individuals in tier-I or tier-II cities, limiting the scope of our findings. Secondly, the remote nature of the study limited our ability to capture factors impacting app performance, such as camera placement and lighting, while the short study duration restricted our insight into long-term usage patterns. Given these limitations, our findings highlight the research question: **"How are OR apps, originally designed to support independence, being adapted and employed within the social context of interdependence, such as India?"** Building on the study findings, in this section, we unpack the practices of "interdependence" in a community-oriented society, like India, and propose guidelines for designing assistive technologies tailored to this context.

### 5.1 Unpacking interdependence in a community-oriented society

India, home to a third of the world's blind or low-vision population [41], is a country rooted in community-focused values where collective interests take precedence over individual ones [34]. Drawing on

Bennett et al.'s Interdependence framework [11], our study findings highlight the limitations of designing assistive technologies, like camera-based OR apps, that assume users with disabilities operate in isolation from their social, cultural, and environmental contexts. The frequent task failures encountered in our study prompt us to question what 'independence' means within assistive technologies, particularly those designed in North American contexts that prioritize individualism, when applied to societies like India, deeply reliant on community and familial support?

When using OR apps in the company of their sighted loved ones, participants were worried to come across as 'too independent' or 'pretentious'. What is it about such assistive technologies that makes the sighted people perceive them as foreign and 'alienating'? Is this solely coming from a place of lack of awareness about disability and assistive technologies, or is it also rooted in the design of such technologies that further isolate the users from their surroundings? Our participants struggled to bridge the gap that technology creates, a challenge familiar to technology users with disabilities [35, 50]. As P4 said, grappling with the balance between asserting his 'autonomy' and negotiating social relationships, *"What is it? How is that situation managed? How does one handle it?"* The roots of these challenges often lie in the participants' concern: "Am I assimilating well in my community or not?"

The interdependence framework is a radical shift from 'forced' independence, designing technology that balances personal autonomy with community reliance, allowing individuals to choose when to rely on others. Interdependence isn't the opposite of independence, rather a complementary approach that recognizes the diverse ways people can achieve personal autonomy, while still engaging with and benefiting from supportive relationships and systems. However, in case of India, where community reliance and collective functioning are deeply ingrained, the concept of interdependence aligns more naturally with societal values. In community-oriented societies like India, reliance on others is often an inherent part of daily life, regardless of one's abilities. Therefore, designing for independence in such societies can have more challenging consequences than in individualistic, Western contexts, where self-sufficiency is more deeply ingrained.

The concept of independence for individuals with disabilities has emerged as a response to a legacy of reliance on state and cultural structures. Throughout history, people with disabilities have frequently experienced societal isolation, residing either within their family homes, dependent on familial support, or within institutional settings under state supervision. However, several disability justice scholars like Mia Mingus have repeatedly emphasized in their works that people, regardless of their abilities, function in dependence on each other, calling it access intimacy [39]. The Interdependence framework advocates for designing technologies that foster these interconnected dependencies, recognizing how people, objects, and environments interact in symbiotic relationships. In India, this mutual reliance is not only common but culturally celebrated. Our participants preferred assistance from their sighted loved ones instead of using an OR app, showcasing their 'ability to rely'. This mutual reliance was equally emphasized when using the apps in the company of loved ones who are blind or low vision, where the focus was on their role as reliable members of society. In both instances, the essence of these interactions is mutual reliance—foundational to any community-oriented society.

Designing assistive technologies with the goal of achieving independence often assumes that people with disabilities should distance themselves from reliance on others and from viewing themselves primarily as recipients of care and access. However, our findings challenge this notion, revealing that individuals of all abilities act as both 'active' and 'passive' recipients of care, facilitated by technologies designed for people with disabilities [15]. For example, in P9's task scenario, where she used an OR app to confirm whether her sighted son had cleaned his room, both she and her son were recipients of care and access. This reflects a collaborative dynamic, which also extends to the interaction with assistive technologies. While OR apps provide object recognition and detection, it is the user who plays an active role in framing objects, adjusting lighting, and managing camera movements, as seen in other studies [46]. This highlights the overlooked collaborative nature of such technologies, which often fail to provide sufficient support to nurture these dynamics. As AI systems increasingly take on autonomous roles, our understanding of collaboration is evolving, with AI-enabled systems seen not just as tools but as active teammates. Designers should consider how to enhance the efficiency of these collaborations, particularly in terms of their impact on users from diverse backgrounds.

Prior work suggests that the exchange of access is often a collaborative effort, whether between humans and technology, humans themselves, or a combination of both, and the assistive technologies should be designed to make this collaboration enriching and fulfilling [11, 64]. Could technology being designed to approach such negotiations with the aim of enriching human-human collaboration rather than limiting them? From the perspective of a society inherently entangled with community-focused values, mutual reliance is a language of love and care, and the design goal for promoting interdependence should be to support community dynamics and not replace it. As Hellen Keller has said, *"Alone we can do so little; together we can do so much."* How, as assistive technology designers, can we ensure this symphonic, fruitful and enriching dependence?

## 5.2 Guidelines for designing assistive technologies for interdependent, community-oriented societies

Our diary study findings reveal the key opportunities to address the usability and accessibility issues in OR apps. In this section, we propose a set of guidelines for designing assistive technologies for interdependent, community-oriented societies, like India. In one promising approach, the technologies could be designed to operate "implicitly", allowing users to navigate the socio-cultural expectations seamlessly within their local contexts [32]. However, in this section, we propose guidelines for technologies that are designed to function explicitly, to make them more inclusive and contextually appropriate.

*5.2.1 Community-centered design approach.* For social contexts that are community-oriented, it's crucial to design technologies that are not only user-friendly but also community-friendly. A community-centered design approach places the community at the heart of the process, aiming to create inclusive and sustainable solutions for the group as a whole. This goes beyond the user-centered design, which primarily focuses on the needs and preferences of individual users. Instead, a community-centered approach looks at the goals and needs of the community, using methods like participatory design [52] to gain a deeper understanding of the social, cultural, and emotional dynamics involved. Engaging various stakeholders also helps establish trust with the community, thus fostering a sense of ownership and collective involvement [43, 45]. Further, prioritizing interdependence and mutual reliance also helps harness the power of people coming together for a common purpose. For example, a playful, collaborative approach to teaching computational thinking to children who are blind or low-vision in schools for the blind in the South of India reimagined the role of teachers as "players" rather than mere facilitators, for the learning experience to be more immersive and deeply rooted in shared engagement [28]. A community-centered design bridges the gap between technology and the community, reducing risks such as social alienation, as also faced by our study participants using OR apps in India.

*5.2.2 Cultural alignment.* Care and attention should be given to the designed technologies that align with the local cultural norms and practices. This involves every aspect of design, from incorporating regional languages to ensuring that AI training and testing datasets reflect the local socio-cultural context. Prior research underscores that cultural alignment significantly impacts technology adoption [23]. In our study, participants expressed frustration and disappointment when OR apps failed to recognize certain Indian items, revealing the emotional toll of such disconnects. These apps are primarily trained on datasets from sighted individuals in Western countries [22, 37, 56], missing both data from blind or low-vision people and objects that are prevalent in non-western cultural contexts like India. This leads to false recognition of objects like a sari (an Indian un-stitched stretch of woven fabric popularly worn by women [59]), as a blanket. While prior work has made significant effort at collecting datasets with people who are blind or low-vision, like the VizWiz-Priv dataset [22] and ORBIT dataset [37], such initiatives should also extend to non-Western contexts. Overall, designing technologies with cultural specificity ensures they are more likely to be embraced, understood, and valued by the communities they serve.

*5.2.3 Adaptability.* The key to designing technologies for a diverse country like India, is to make them locally adaptable. In Western contexts, high-resource technologies are often designed with advanced components and a focus on scalability. This approach stems from the need to automate tasks in environments with limited human resources. However, in community-oriented contexts, the goal is to function cohesively as a unit and such high-cost, high-tech solutions may feel alien or inaccessible to households with varying levels of resources. For instance, it's a common practice in the Global South countries such as India, Pakistan, and Bangladesh, for multiple members of a family to share a single mobile device [44]. Technologies designed for such contexts should empower users to adapt them to their specific scenarios, ensuring flexibility and inclusivity. For instance, in the case of P6, for scene exploration at the railway station, an option in the app to click a picture and upload it to the cloud for feedback would have been more discreet and context-appropriate over recording their surroundings for several seconds. Such adaptability allows users to maintain control over their experience, deciding when to be discreet, whether to reveal their disability identity, or how much to rely on the technology. Designing for adaptability not only respects an individual's preferences but also aligns with the social dynamics and resource realities of their diverse communities.

*5.2.4 Reduce social stigma.* Assistive technologies hold the power to alleviate the social stigma faced by users with disabilities. Shinohara et al. [50] suggest accounting for social accessibility as much for functional accessibility when designing assistive technologies, as "There persists the notion that "you are who you are perceived to be" and that perception can be influenced by what you use." Assistive technologies should be designed to either blend seamlessly or stand out in a positive way. This promotes positive self-image among users while ensuring their full participation in society. Prior work [23] suggests creating thoughtful and artistic designs, such as creating prosthetics or white canes in vibrant, approachable colors, which may also transform them into conversation starters. Such inclusive design has the potential to shift societal perceptions, fostering environments where assistive devices are seen as empowering tools rather than markers of difference.

## 6 CONCLUSION AND FUTURE WORK

This paper presents findings from a seven-day diary study and follow-up interviews with seven blind or low-vision users of object-recognition apps (OR apps) in India, exploring: "How are OR apps, originally designed to support independence, being adapted and employed within a cultural context of interdependence, such as India?" Our findings emphasize the cultural significance of interdependence in community-oriented societies like India, where mutual reliance is not only celebrated but essential for social inclusion. We compare these experiences with those in Western contexts, where independence is often prioritized, and argue for a shift toward a community-centric approach in designing assistive technologies. Given the study's limitations, future research should explore how diverse socio-economic factors (e.g., gender, caste, digital literacy, rural areas) influence the use of OR apps. Additionally, there is merit in conducting in-person observations to gain a better understanding of how the OR apps are used in real-life settings. Observing factors such as camera usage and object placement can offer valuable insights into the cognitive processes and mental models employed by users during app usage.

## References


[1] Dustin Adams, Lourdes Morales, and Sri Kurniawan. 2013. A qualitative study to support a blind photography mobile application. In *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments* (Rhodes, Greece) *(PETRA '13)*. Association for Computing Machinery, New York, NY, USA, Article 25, 8 pages. doi:10.1145/2504335.2504360

[2] Abhay Agarwal, Sujeath Pareddy, and Manohar Swaminathan. 2017. FReAD: A Multimodal Interface for Audio Assisted Identification of Everyday Objects. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 1471–1477. doi:10.1145/3027063.3053107

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

[4] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. 2016. Addressing physical safety, security, and privacy for people with visual

impairments. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. Association for Computing Machinery, New York, NY, USA, 341–354.

[5] Envision AI. 2017. Envision - Perceive Possibility. https://www.letsenvision.com/

[6] Aira. 2014. We're Aira, a Visual Interpreting Service. https://aira.io/

[7] Taslima Akter, Tousif Ahmed, Apu Kapadia, and Swami Manohar Swaminathan. 2020. Privacy Considerations of the Visually Impaired with Camera Based Assistive Technologies: Misrepresentation, Impropriety, and Fairness. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) *(ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 32, 14 pages. doi:10.1145/3373625.3417003

[8] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. 2020. "I am uncomfortable sharing what I can't see": Privacy Concerns of the Visually Impaired with Camera Based Assistive Applications. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, Berkeley, CA, 1929–1948. https://www.usenix.org/conference/usenixsecurity20/presentation/akter

[9] Amazon. 2005. Amazon Mechanical Turk. https://www.mturk.com/worker.

[10] Mauro Avila, Katrin Wolf, Anke Brock, and Niels Henze. 2016. Remote Assistance for Blind Users in Daily Life: A Survey about Be My Eyes. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (Corfu, Island, Greece) *(PETRA '16)*. Association for Computing Machinery, New York, NY, USA, Article 85, 2 pages. doi:10.1145/2910674.2935839

[11] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. 2018. Interdependence as a Frame for Assistive Technology Research and Design. 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, 161–173. doi:10.1145/3234695.3236348

[12] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, and Tom Yeh. 2010. VizWiz: nearly real-time answers to visual questions. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. Association for Computing Machinery, New York, NY, USA, 333–342. doi:10.1145/1866029.1866080

[13] Jeffrey P. Bigham, Chandrika Jayant, Andrew Miller, Brandyn White, and Tom Yeh. 2010. VizWiz::LocateIt - enabling blind people to locate objects in their environment. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*. IEEE, San Francisco, CA, USA, 65–72. doi:10.1109/CVPRW.2010.5543821

[14] Erin Brady, Meredith Ringel Morris, Yu Zhong, Samuel White, and Jeffrey P. Bigham. 2013. Visual challenges in the everyday lives of blind people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2117–2126. doi:10.1145/2470654.2481291

[15] Stacy M. Branham and Shaun K. Kane. 2015. Collaborative Accessibility: How Blind and Sighted Companions Co-Create Accessible Home Spaces. 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, 2373–2382. doi:10.1145/2702123.2702511

[16] Virginia Braun and Victoria Clarke. 2024. Thematic analysis. In *Encyclopedia of quality of life and well-being research*. Springer, New York, NY, US, 7187–7193.

[17] CamFind. 2013. CamFind Blind Accessibility App - iAccessibility Solutions for iOS Communications. https://www.iaccessibility.com/apps/blind/index.cgi/product?ID=20

[18] Scott Carter and Jennifer Mankoff. 2005. When participants do the capturing: the role of media in diary studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. Association for Computing Machinery, New York, NY, USA, 899–908. doi:10.1145/1054972.1055098

[19] Be My Eyes. 2015. Be My Eyes - See the world together. https://www.bemyeyes.com/

[20] Talking Goggles. 2025. Talking Goggles. https://www.sparklingapps.com/goggles/

[21] Anhong Guo, Xiang 'Anthony' Chen, Haoran Qi, Samuel White, Suman Ghosh, Chieko Asakawa, and Jeffrey P. Bigham. 2016. VizLens: A Robust and Interactive Screen Reader for Interfaces in the Real World. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 651–664. doi:10.1145/2984511.2984518

[22] Danna Gurari, Qing Li, Chi Lin, Yinan Zhao, Anhong Guo, Abigale Stangl, and Jeffrey P Bigham. 2019. Vizwiz-priv: A dataset for recognizing the presence and purpose of private visual information in images taken by blind people. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, California, US, 939–948.

[23] Megan Hofmann, Devva Kasnitz, Jennifer Mankoff, and Cynthia L Bennett. 2020. Living Disability Theory: Reflections on Access, Research, and Design. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) *(ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 4, 13 pages. doi:10.1145/3373625.3416996

[24] Author Bill Holton. 2016. BeSpecular: A New Remote Assistant Service | Accessworld | American Foundation for the Blind. https://www.afb.org/aw/17/7/15313

[25] Josh Howarth. 2021. How Many People Own Smartphones? (2024-2029). https://explodingtopics.com/blog/smartphone-stats

[26] Amy Hurst and Jasmine Tobias. 2011. Empowering individuals with do-it-yourself assistive technology. 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, 11–18. doi:10.1145/2049536.2049541

[27] Gesu India, Mohit Jain, and Manohar Swaminathan. 2021. Understanding motivations and barriers to exercise among people with blindness in India. In *Human-Computer Interaction–INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30–September 3, 2021, Proceedings, Part I 18*. Springer, Springer, Bari, Italy, 444–454.

[28] Gesu India, Geetha Ramakrishna, Joyojeet Pal, and Manohar Swaminathan. 2020. Conceptual Learning through Accessible Play: Project Torino and Computational Thinking for Blind Children in India. In *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development* (Guayaquil, Ecuador) *(ICTD '20)*. Association for Computing Machinery, New York, NY, USA, Article 6, 11 pages. doi:10.1145/3392561.3394634

[29] Gesu India, Vidhya Y, Aishwarya O, Nirmalendu Diwakar, Mohit Jain, Aditya Vashistha, and Manohar Swaminathan. 2021. Teachers' Perceptions around Digital Games for Children in Low-resource Schools for the Blind. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 43, 17 pages. doi:10.1145/3411764.3445194

[30] Rabia Jafri, Syed Abid Ali, Hamid R Arabnia, and Shameem Fatima. 2014. Computer vision-based object recognition for the visually impaired in an indoors environment: a survey. *The Visual Computer* 30 (2014), 1197–1222.

[31] Chandrika Jayant, Hanjie Ji, Samuel White, and Jeffrey P. Bigham. 2011. Supporting blind photography. 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, 203–210. doi:10.1145/2049536.2049573

[32] Wendy Ju and Larry Leifer. 2008. The design of implicit interactions: Making interactive systems less obnoxious. *Design Issues* 24, 3 (2008), 72–84.

[33] Code Jumper. 2019. Code Jumper. https://codejumper.com/

[34] Cigdem Kagitcibasi. 1997. Individualism and collectivism. *Handbook of cross-cultural psychology* 3 (1997), 1–49.

[35] Kyungjun Lee, Daisuke Sato, Saki Asakawa, Hernisa Kacorri, and Chieko Asakawa. 2020. Pedestrian Detection with Wearable Cameras for the Blind: A Two-way Perspective. 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, 1–12. doi:10.1145/3313831.3376398

[36] Google LLC. 2019. Lookout - Assisted vision – Apps on Google Play. https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal&hl=en_GB

[37] Daniela Massiceti, Luisa Zintgraf, John Bronskill, Lida Theodorou, Matthew Tobias Harris, Edward Cutrell, Cecily Morrison, Katja Hofmann, and Simone Stumpf. 2021. Orbit: A real-world few-shot dataset for teachable object recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE, Montreal, Canada, 10818–10828.

[38] Microsoft. 2017. Seeing AI - Talking Camera for the Blind. https://www.seeingai.com/

[39] Mia Mingus. 2011. Interdependence (exerpts from several talks) | Leaving Evidence. https://leavingevidence.wordpress.com/2010/01/22/interdependency-exerpts-from-several-talks/

[40] Cecily Morrison, Nicolas Villar, Anja Thieme, Zahra Ashktorab, Eloise Taysom, Oscar Salandin, Daniel Cletheroe, Greg Saul, Alan F Blackwell, Darren Edge, et al. 2020. Torino: A tangible programming language inclusive of children with visual disabilities. *Human–Computer Interaction* 35, 3 (2020), 191–239.

[41] World Health Organization. 2023. Blindness and vision impairment. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

[42] Joyojeet Pal and Meera Lakshmanan. 2012. Assistive technology and the employment of people with vision impairments in India. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development* (Atlanta, Georgia, USA) *(ICTD '12)*. Association for Computing Machinery, New York, NY, USA, 307–317. doi:10.1145/2160673.2160711

[43] Jennifer Pearson, Simon Robinson, Thomas Reitmaier, Matt Jones, and Anirudha Joshi. 2019. Diversifying Future-Making Through Itinerative Design. *ACM Trans. Comput.-Hum. Interact.* 26, 5, Article 33 (July 2019), 21 pages. doi:10.1145/3341727

[44] Agnes R Quisumbing, Neha Kumar, and Julia A Behrman. 2018. Do shocks affect men's and women's assets differently? Evidence from Bangladesh and Uganda. *Development Policy Review* 36, 1 (2018), 3–34.

[45] Dani Kalarikalayil Raju, Krishna Seunarine, Thomas Reitmaier, Gethin Thomas, Yogesh Kumar Meena, Chi Zhang, Adam Pockett, Jennifer Pearson, Simon Robinson, Matt Carnie, Deepak Ranjan Sahoo, and Matt Jones. 2021. PV-Pix: Slum Community Co-design of Self-Powered Deformable Smart Messaging Materials.

In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 304, 14 pages. doi:10.1145/3411764.3445661

[46] Gisela Reyes-Cruz, Joel Fischer, and Stuart Reeves. 2022. Supporting Awareness of Visual Impairments and Accessibility Reflections through Video Demos and Design Cards. In *Nordic Human-Computer Interaction Conference* (Aarhus, Denmark) *(NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 67, 15 pages. doi:10.1145/3546155.3546697

[47] Bikash Santra and Dipti Prasad Mukherjee. 2019. A comprehensive survey on computer vision based approaches for automatic identification of products in retail store. *Image and Vision Computing* 86 (2019), 45–63.

[48] Tanusree Sharma, Abigale Stangl, Lotus Zhang, Yu-Yun Tseng, Inan Xu, Leah Findlater, Danna Gurari, and Yang Wang. 2023. Disability-First Design and Creation of A Dataset Showing Private Visual Information Collected With People Who Are Blind. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 51, 15 pages. doi:10.1145/3544548.3580922

[49] Kristen Shinohara and Jacob O. Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. 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, 705–714. doi:10.1145/1978942.1979044

[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. doi:10.1145/2827857

[51] Blind Welfare Society. 2025. Transportation Rights in India: Ensuring Accessible Travel for the Visually Impaired. https://blindwelfaresociety.in/blogs/transportation-rights-in-india-ensuring-accessible-travel-for-the-visually-impaired

[52] Clay Spinuzzi. 2005. The methodology of participatory design. *Technical communication* 52, 2 (2005), 163–174.

[53] Start.io. 2024. Start.io | Low-End Android Users in India Audience. https://www.start.io/audience/low-end-android-users-in-india

[54] Supersense. 2020. Supersense - AI for Blind / Scan text, money and objects. https://www.supersense.app/

[55] TapTapSee. 2012. TapTapSee - Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API. https://taptapseeapp.com/

[56] Lida Theodorou, Daniela Massiceti, Luisa Zintgraf, Simone Stumpf, Cecily Morrison, Edward Cutrell, Matthew Tobias Harris, and Katja Hofmann. 2021. Disability-first Dataset Creation: Lessons from Constructing a Dataset for Teachable Object Recognition with Blind and Low Vision Data Collectors. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 27, 12 pages. doi:10.1145/3441852.3471225

[57] Anja Thieme, Cynthia L. Bennett, Cecily Morrison, Edward Cutrell, and Alex S. Taylor. 2018. "I can do everything but see!" – How People with Vision Impairments Negotiate their Abilities in Social Contexts. 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, 1–14. doi:10.1145/3173574.3173777

[58] Contributors to Wikimedia projects. 2003. Culture of India - Wikipedia. https://en.wikipedia.org/wiki/Culture_of_India

[59] Contributors to Wikimedia projects. 2003. Sari - Wikipedia. https://en.wikipedia.org/wiki/Sari

[60] Contributors to Wikimedia projects. 2004. List of languages by number of native speakers in India - Wikipedia. https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers_in_India

[61] Contributors to Wikimedia projects. 2005. Flattened rice - Wikipedia. https://en.wikipedia.org/wiki/Flattened_rice

[62] Contributors to Wikimedia projects. 2006. Classification of Indian cities - Wikipedia. https://en.wikipedia.org/wiki/Classification_of_Indian_cities

[63] Aditya Vashistha, Erin Brady, William Thies, and Edward Cutrell. 2014. Educational Content Creation and Sharing by Low-Income Visually Impaired People in India. In *Proceedings of the Fifth ACM Symposium on Computing for Development* (San Jose, California, USA) *(ACM DEV-5 '14)*. Association for Computing Machinery, New York, NY, USA, 63–72. doi:10.1145/2674377.2674385

[64] Beatrice Vincenzi, Alex S. Taylor, and Simone Stumpf. 2021. Interdependence in Action: People with Visual Impairments and their Guides Co-constituting Common Spaces. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 69 (April 2021), 33 pages. doi:10.1145/3449143

[65] Shaomei Wu and Lada A. Adamic. 2014. Visually impaired users on an online social network. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3133–3142. doi:10.1145/2556288.2557415

[66] Zhuohao (Jerry) Zhang, Smirity Kaushik, JooYoung Seo, Haolin Yuan, Sauvik Das, Leah Findlater, Danna Gurari, Abigale Stangl, and Yang Wang. 2023. ImageAlly: A Human-AI Hybrid Approach to Support Blind People in Detecting and Redacting Private Image Content. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. USENIX Association, Anaheim, CA, 417–436. https://www.usenix.org/conference/soups2023/presentation/zhang