Designing Accessible Oversight of Household Robots with Blind and Low-Vision People

Mina Huh
University of California, Berkeley
Berkeley, California

Huihan Liu[*]
University of Texas, Austin
Austin, Texas

Albert Yu[*]
University of Texas, Austin
Austin, Texas

Ujjaini Das[*]
University of Texas, Austin
Austin, Texas

Roberto Martin-Martin
University of Texas, Austin
Austin, Texas

Yuke Zhu
University of Texas, Austin
Austin, Texas

Maya Cakmak
University of Washington
Seattle, Washington

Amy Pavel
University of California, Berkeley
Berkeley, California

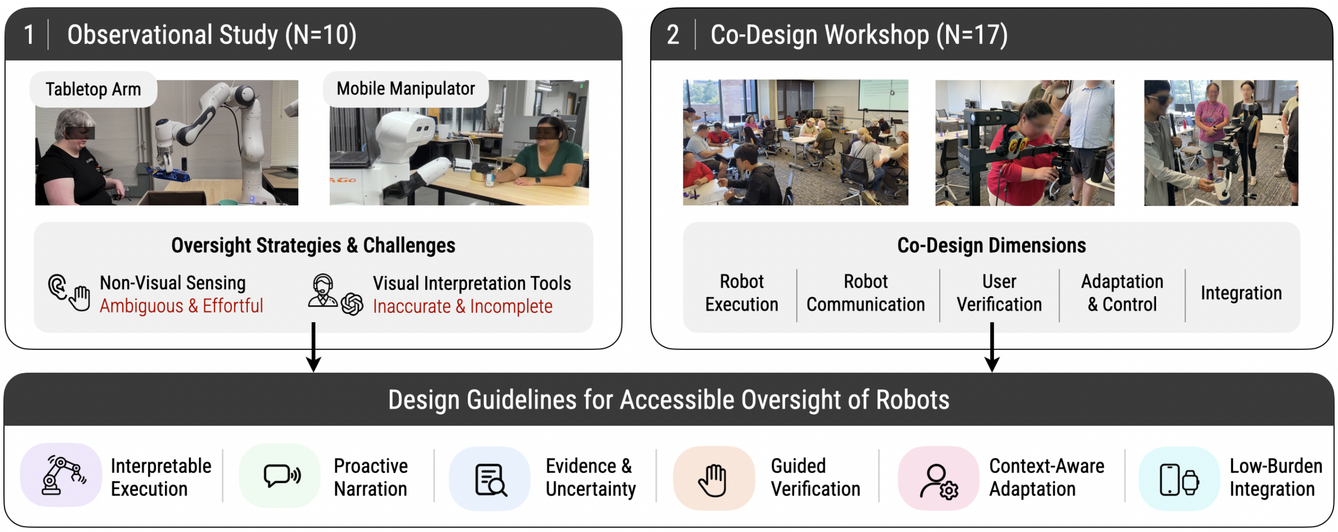

Figure 1: To understand current strategies and challenges for accessible robot oversight, we conducted 1) an observational study with 10 BLV participants. We then conducted 2) a co-design workshop with 17 BLV participants and accessibility and multimodal-interaction experts to explore how robot behavior and communication could better support oversight across everyday contexts. Across both studies, we derive design considerations for accessible robot oversight.

Abstract

Robots are moving into homes with promise to reduce barriers to housework for people with disabilities and decrease effort for everyone. As robots perform tasks autonomously, people need to oversee their progress and outcomes and know when to intervene. While sighted users can visually notice mistakes, inefficient behavior, and changes in the environment, blind and low vision (BLV) users cannot rely on visual cues to identify problems or verify outcomes. To understand current strategies and challenges in non-visual robot monitoring, we conducted an observational study with 10 BLV participants who monitored multi-step household tasks performed by two robot platforms. Participants used sound, touch, and visual interpretation tools to follow robot actions and assess outcomes, but these approaches were often ambiguous and effortful, leading to missed errors and misinterpretations. We then conducted a co-design workshop with 11 BLV participants and 6 experts in accessibility and multimodal communication. Participants proposed predictable robot behavior and proactive pauses for user verification, alongside multimodal communication with clear cue-to-meaning mappings and context-adaptive delivery. Across both studies, we characterize challenges in non-visual robot monitoring and derive design guidelines for accessible robot oversight.

[*]These authors contributed equally to this work.

Conference acronym 'XX, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

ACM Reference Format:
Mina Huh, Huihan Liu, Albert Yu, Ujjaini Das, Roberto Martin-Martin, Yuke Zhu, Maya Cakmak, and Amy Pavel. 2018. Designing Accessible Oversight of Household Robots with Blind and Low-Vision People. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX).* ACM, New York, NY, USA, 18 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Robots are beginning to take on everyday physical tasks on people's behalf (*e.g.,* tidying, washing dishes, cooking). Such robots can save time and effort for everyone, including people with disabilities [47]. Yet, even when a robot performs a task autonomously, people still need to *oversee* its work–to understand its progress, verify outcomes, and intervene when necessary. Sighted users can often do this by watching the robot or glancing at the workspace, while blind and low-vision (BLV) people may not be able to use visual observation to notice errors or identify what has changed.

To help people follow and interpret robot behavior, prior work has developed mechanisms such as legible motion [17], gaze and attention cues [7, 46], and on-robot or environmental displays [10, 66, 70]. These approaches assume a sighted observer who can already see the robot, its workspace, and the physical changes it makes. Prior research with BLV people has largely focused on assistive robots for navigation and object finding, where the robot guides the user and the user remains nearby and actively engaged [12, 29, 35, 54]. General-purpose household automation changes this relationship: the robot instead performs tasks on the user's behalf, allowing users to shift their attention elsewhere while the robot continues acting. Accessible automation should therefore preserve the benefit of offloading work while still enabling BLV users to understand progress, verify outcomes, and intervene when necessary. In this work, we investigate: *How can BLV people effectively oversee physical work delegated to autonomous robots?*

To understand current nonvisual oversight strategies and challenges, we conducted an observational study with 10 BLV participants who interacted with two robots (a tabletop arm [58] and a mobile manipulator [59]) as the robots performed household tasks. Participants used naturally available sounds to follow robot activity in real time, but these cues were often incomplete or ambiguous. Tactile inspection provided more direct information about task outcomes, but was typically only possible after the robot had stopped or completed the task. Participants also used human- and AI-powered visual interpretation tools (*e.g.,* Meta glasses), but these tools could provide inaccurate or overly generic descriptions and depended on correctly framing the scene and asking the right questions. Across 4 tasks, participants' reports contained an average of 7.5 inaccuracies per participant (SD=1.27), including missed robot errors, missed task steps, and reports of events that did not occur.

Building on these strategies and challenges, we conducted an in-person co-design workshop (N=17) to explore how robot oversight can be redesigned. The workshop brought together 11 BLV participants and 6 experts in accessibility and multimodal communication. Across 7 workshop groups, participants proposed changes to robot behavior, task information, access modalities, adaptation and user control, and integration with familiar technologies and practices. Together, the two studies reveal where nonvisual oversight breaks down and how accessible oversight might be supported across different contexts. We synthesize these findings into 6 design guidelines for accessible oversight of autonomous robot work.

We contribute:

- An observational study (N=10) characterizing how BLV people oversee household robots and the challenges they face.
- A co-design study (N=17) identifying recurring design patterns for accessible robot oversight.
- Design guidelines for accessible oversight of autonomous robot work, synthesized across both studies.

2 Related Work

2.1 Robot Communication for Oversight

Prior HRI research has examined *what* information users seek from robots during task execution [68], *when* explanations are desired [69], and *how* robot behavior and internal state can be made transparent to users [62]. Existing robots can visually convey information through legible motion [17], eye gaze [7], visual projection [13], and deictic gestures [21, 28]. Other work has made otherwise hidden robot reasoning and decision processes more transparent through policy explanations [23] and visualizations of robot control architectures [24]. Robots have also been augmented with sound cues [49] and rich language-based communication [48, 71] to explain robot failures [71], respond to user questions [68], and adapt explanations when users appear confused [48]. Notably, none of these studies report including blind participants. These approaches largely study communication as a supplement to visual observation rather than as the primary means of accessing robot activity and task outcomes. Our work explores how these oversight requirements change for BLV people when visual observation is unavailable.

2.2 Accessible Robot Interaction for Blind and Low-Vision People

Prior work has developed assistive robots for BLV people, particularly for navigation [9, 31, 32, 35, 42, 55] and object localization [12, 29, 54]. These systems guide the user to a location or object through explicit *physical guidance* (*e.g.,* the robot moves to a location as the person holds onto it) or *verbal guidance* (*e.g.,* the robot states the remaining distance to a location). Across these settings, the robot assists the user in carrying out specific tasks, with the user typically remaining nearby and actively engaged with its guidance. General-purpose household automation changes this relationship: rather than assisting the user in performing a task, the robot performs the task on the user's behalf [18, 61, 65]. As robots increasingly take on such everyday work, accessibility must extend beyond disability-specific assistive systems to general-purpose robots [53]. For BLV people, accessible automation should preserve the benefit of offloading work, while letting them oversee robot work safely and without much effort. As a step toward accessible robots for daily living, we study how BLV people currently oversee autonomous household robot work and derive design guidance for supporting this oversight.

2.3 Accessible Visual Assistance and Awareness

Existing accessibility tools have helped BLV people access visual information in everyday activities. VizWiz allows users to photograph objects or scenes and ask questions answered by remote crowd workers [11]. BeMyEyes [1] and Aira [2] similarly connect camera feeds to sighted agents, and newer AI tools such as BeMyAI [3] and Meta Glasses [4] generate scene descriptions and support follow-up questions. Beyond accessing real-world scenes, accessibility research further explored how BLV people can maintain *awareness* of other people's ongoing activity. In shared digital environments (*e.g.,* document editors), accessibility extension systems communicate who is editing what through earcons, voice fonts, and spatial audio [15, 39], and summarizes visual changes made by collaborators so BLV users can review and respond to them [51]. In physical environments, systems such as PeopleLens provide continuous information about nearby people's identity, location, and attention [45]. These systems support awareness of other people, but did not yet explore a setting in which a user delegates physical work to another actor and may shift their attention elsewhere. In this work, we examine how BLV people oversee autonomous robots acting on their behalf.

3 In-Person Observational Study

To understand how blind people make sense of robotic behaviors, task progress, and outcomes, we conducted an in-person observational study with 10 BLV participants. Researchers teleoperated the robot for safety and consistent demonstrations. We investigate the following research questions:

RQ1. *What strategies do BLV people use to oversee robot task execution and outcomes?*

RQ2. *What challenges do BLV people encounter when using these strategies for oversight?*

3.1 Method

Participants. We recruited 10 BLV participants (P1–P10, Table 3) through a local chapter of the National Federation of the Blind (NFB) and word of mouth. We asked participants to bring any visual assistive technologies they use in daily life (*e.g.,* smart glasses, visual interpretation smartphone applications) for optional use during the study. The study was approved by our institution's IRB, and participants received $70 for the 2-hour session.

Robots. We used two robot platforms with complementary form factors and capabilities: TIAGo, a mobile manipulator with a wheeled base and dual arms [59], and the Franka Emika Panda, a table-top robotic arm [58]. We selected these platforms to cover common household task demands, and given that robot morphology and noise can shape how people perceive robot behavior [34]. During the study, a researcher teleoperated both robots using a 3Dconnexion SpaceMouse. Participants were unaware of the teleoperation during the task and were informed during the post-study debrief.

Tasks & Procedure. The study took place in a lab space staged with household objects (*e.g.,* a table setting, sink, trash items, pantry items). We began with a brief description of the space and each robot's appearance, size, and capabilities. Following prior work [12], we then allowed participants a brief tactile exploration of the robots

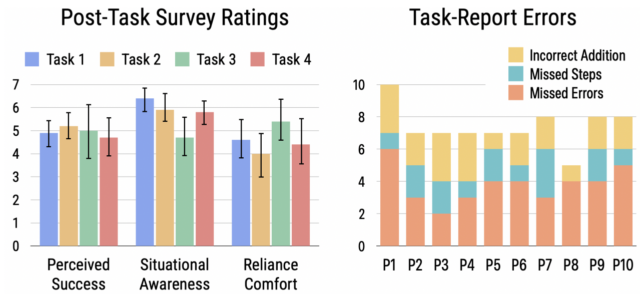

Figure 2: Average post-task survey ratings (1-7) for perceived success of the robot, situational awareness, and reliance comfort (Left). Frequency of three types of inaccuracies (missed errors, missed steps, incorrect addition) in participants' task reports (Right).

while they were stationary. Participants then supervised two robots completing four household tasks (Table 4): (1) setting the table with matching plates and cups, (2) cleaning up the table by sorting dishes and trash, (3) searching for and bringing salt, and (4) bringing pasta and pouring it into a pot. The tabletop Franka performed (1) and (2) while the mobile Tiago performed (3) and (4). The robot provided a brief verbal update at the start and end of each task to convey minimal status updates. In each task, we introduced two scripted failures spanning diverse error types following prior taxonomies of failures in HRI [26] (Table 4), to examine how participants detect, interpret, and respond to common breakdowns.

Participants could use any non-visual cues and any assistive technologies they typically use to observe the robot and verify the robots' task completion. When they wanted to touch the robot or workspace to verify task state, we paused the robot and allowed touch only when it was idle to prevent unexpected collisions. Participants were asked to think aloud while completing each task. After each task, they briefly summarized what they believed had happened and completed a short survey assessing perceived task success, awareness of the robot's behavior and outcomes, and comfort relying on the robot. After all tasks, we conducted a semi-structured interview to understand participants' strategies and challenges.

Analysis. We video-recorded all sessions. To characterize oversight outcomes, we analyzed participants' post-task ratings and compared their summaries of what happened against the recorded robot execution, identifying missed errors, and omitted steps. We then iteratively coded participants' task interactions, think-aloud comments, and interview responses to identify recurring oversight strategies (§3.3) and challenges (§3.4).

3.2 Understanding of Robot Task Execution

Post-task survey ratings show that participants perceived the robots as moderately successful, and rated their situational awareness as relatively high across all tasks (Figure 2, left). Despite high self-ratings of situational awareness, the robot's execution did not provide enough accessible information for accurate understanding of what had occurred. Figure 2 (Right) shows our categorization of different types of inaccuracies in each participant's task reports, including missed robot errors, omitted task steps, and reports of events that did not occur. Participant reports of the robot's actions contained an average of 7.5 inaccuracies (SD=1.27): 51% were missed errors, 30% were incorrect additions, and 20% were missed steps. Incorrect additions were often due to ambiguous robot or environmental sounds that could plausibly be interpreted as actions that had not occurred.

3.3 Non-Visual Oversight Strategies

Participants used 3 recurring strategies to monitor robot activity and verify task outcomes (S1-S3): auditory cues, direct sensory inspection, and AI or human visual assistance. Table 1 summarizes the oversight strategies used across tasks.

Task	Monitoring and verification approaches (# participants)				
	Auditory cues	Tactile cues	Smell/taste cues	Human assistance	AI assistance
Set table	10	5	0	2	4
Clear table	10	8	0	1	2
Fetch salt	10	2	4	3	3
Pour pasta	10	7	0	1	2

Table 1: Approaches blind participants used to monitor and verify robot task execution.

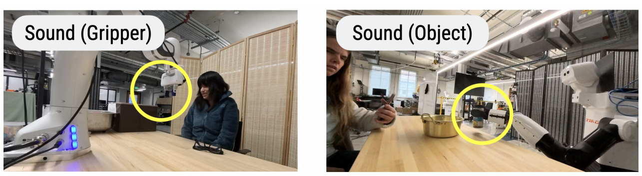

Figure 3: Participants used sound to monitor robot actions, listening for the gripper opening and closing and for objects being placed.

S1. Auditory cues for monitoring robot activity. In all 4 tasks, all 10 participants reported using sound cues to track when the robot began moving and when made contact with objects (*e.g.,* grasping and placing) (Table 1). For mobile robot, 4 participants (P1, P4, P7, P10) mentioned using wheel sound to infer where it was heading, although P1 and P4 still found its distance difficult to judge and felt uncomfortable when it moved nearby. Grasping sounds were more perceptible from the nearby tabletop robot than from the mobile robot when it was farther away.

Participants also detected failures using sound. All 10 participants noticed the robot spilling the pasta because pasta sounded different when falling into the pot versus onto the table. 8 participants noticed a cup knocked over by the robot arm. However, other errors were difficult to infer from audio, and sound could also be misleading. P9 misidentified the pepper bottle the robot brought, explaining it *"sounded like salt."* 3 participants also misinterpreted environmental or robot sounds. For example, P9 heard a building fan and assumed the robot was scanning objects, while P1 and P5 interpreted the robot's arm-retraction sound as actions that had not occurred (*e.g.,* opening the pasta box, recovering spilled pasta).

Participants varied in their ability to extract information from audio cues. P2 (a musician) explained that cups and plates make distinctive sounds and that he could tell which of the two the robot was placing. In contrast, P10 said, *"I have no clue what objects the robot is placing each time. I have to wait until I can touch it after."* P8 added: *"Cups going into the sink make a clear sound, but when some small trash goes into the trash bin, it's not really making a noise."*
S2. Touch, smell, and taste for verifying task outcomes. Participants used touch to verify task outcomes after the robot finished. Participants touched objects to inspect placements, retrieved items, and spills. However, tactile verification was often incomplete: during table clearing, 8 participants swept the table with their hands to confirm nothing was missed, yet 5 still overlooked a napkin remaining at the edge. Similarly, 2 participants attempting to collect spilled pasta still missed pieces left on the table. Tactile checking could also displace objects. For example, P6 accidentally pushed the pepper bottle to the floor while trying to locate it.

Figure 4: Participants used smell and taste to identify retrieved objects. They also used touch to verify outcomes (*e.g.,* detecting a knocked-over cup). Yet, tactile checking could introduce additional errors, such as spilling more pasta while searching for the spill.

Participants also described touch as undesirable in realistic household contexts. 8 participants checked the trash bin and sink after table clearing, but noted that such verification would be impractical in daily life. As P10 put it, *"Who would want to put their hand into the trash to check what the robot did? I expect the robot to work for me, not me working for it."* P4 noted safety concerns: *"If there was actually a boiling pot of water here, trying to touch around it would be really dangerous."* Beyond touch, participants used other senses to verify outcomes. 4 participants smelled, and in some cases tasted, the retrieved spice bottle to identify it. Yet, P9 explained, *"Touching or smelling is slower than hearing. I can only check when the task is complete because I can't interfere while the robot is still moving."*
S3. AI and human visual assistance for accessing visual information. 7 participants used AI-powered visual interpretation tools such as BeMyAI [3] powered by ChatGPT [50], Seeing AI [5], and Gemini Live [6] to capture photos or stream video and receive descriptions (Figure 5). 4 also used Meta smart glasses [4], which reduced the need to hold and aim a phone. Participants often combined AI with other strategies, such as confirming again with touch or cross-checking multiple AI tools to increase confidence.

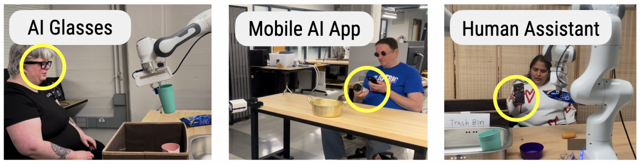

Figure 5: Participants used AI visual interpretation tools (*e.g.,* meta glasses, mobile applications) and remote human assistance to understand robot action and verify the object retrieved.

Framing the robot and target objects was difficult when using phone-based AI tools. During the salt fetching task, P6 repeatedly tried to capture the bottle label for BeMyAI, but it remained out of frame and he eventually switched to a barcode reader. Participants also noted that AI descriptions were often at the wrong level of granularity or required careful prompting. For example, P4 and P7 asked AI to describe what it saw and received generic description (" *a robotic arm on the table with kitchen items*) which required follow-up questions (*e.g., "What are those items? Any trash?"*). P7 noted *"I should ask the right question, otherwise the descriptions are not useful."* Participants further reported inaccuracies and hallucinations as models often relied on a single snapshot or narrow view and thus failed to capture motion or context. P1 noted, *"My Meta Glasses said the robot is stationary, when I could clearly hear it moving."* AI tools were also harder to use when the robot was far away, occluded, or turned away from the camera.

4 participants used human visual interpretation services including Be My Eyes [1] and Aira [2] (Figure 5). These tools allowed participants to share live video over a call and ask questions. Participants provided task context and adjusted the camera to help interpreters access relevant information, yet uncertainty remained. P3 explained: *"When she [the visual interpreter] said there's nothing on the table, I'm not sure if it's true or whether my camera was not showing it all."* Participants also highlighted that the services were often expensive ($100 for 50 minutes of Aira), unreliable with poor connectivity, and could reduce independence and privacy. P9 noted *"I want the robot to describe things, not them [visual interpreters] so that I can feel more independent."* In the interview, 9 participants expressed a preference for AI over human assistance due to more consistent response styles as well as independence and privacy.

3.4 Non-Visual Oversight Challenges

These strategies supported different aspects of non-visual oversight: auditory cues enabled ongoing monitoring (S1), direct sensory inspection supported outcome verification (S2), and visual assistance provided access to otherwise unavailable information (S3). Across these strategies, we synthesized four recurring challenges (C1-C4) that limited complete and reliable oversight.

- **C1 Naturally available sensory cues provide incomplete access to robot activity and outcomes.** Sound can reveal movement, contact, and some failures, but may be absent, ambiguous, or misleading. Touch can verify outcomes, but is slower, effortful, and sometimes undesirable or unsafe.

Figure 6: Co-design workshop materials and activities: (Left) 3D-printed human and robot figures for demonstrating spatial arrangements and interaction ideas. (Right) A participant sensing the sound and vibration from the robot gripper during hands-on exploration.

- **C2 AI and remote visual assistance can provide inaccurate or incomplete information.** Descriptions can misrepresent robot activity or miss events due to limited camera framing, occlusion, missing context, or description quality.
- **C3 Discovering missing information requires knowing when, where, and what to check.** Participants must pose questions, aim cameras, or search the workspace themselves, making unexpected errors or changes difficult to discover.
- **C4 Access to information decreases with distance from the robot and workspace.** Sound is harder to interpret, touch takes longer to initiate, and camera-based assistance is less reliable when activity is distant or out of view.

4 Co-Design Workshop

Our observational study revealed strategies and challenges BLV people encounter when overseeing robot task execution and verifying their outcomes. To engage BLV participants and accessibility experts in designing solutions to these challenges, we conducted a co-design workshop with 17 participants.

4.1 Method

Participants. We recruited 17 participants (Table 5-6), including 11 BLV participants and 6 experts in accessibility and multimodal communication (1 audio describer, 2 orientation and mobility specialists, 2 sound designers, and 1 haptics engineer). We recruited BLV participants through a local NFB chapter and word of mouth, and recruited experts through professional mailing lists. None of the participants participated in our prior observational study.

Robots & Materials. Participants interacted with two robot platforms: a Franka Emika Panda tabletop manipulator [58] and a Hello Robot Stretch mobile manipulator [33]. To broaden the design space beyond the two available robot platforms, we also provided a 3D-printed humanoid robot figure representing an emerging form factor for general-purpose household robots. During the design session, we provided 3D-printed figures representing humans and different robot embodiments that participants could arrange and manipulate to demonstrate their ideas (Figure 6).

Procedure. In the 2-hour workshop, we first summarized findings from the observational study, including the oversight strategies participants used and the challenges encountered. We then introduced the robots through a guided hands-on experience similar to our observational study. Participants tactually explored each robot and its components, learned about its capabilities and movement, and listened to the sounds it produced.

We formed 7 small groups of 2-3 participants, ensuring that BLV participants were not outnumbered by sighted participants in any group to center BLV perspectives during the design process. We used small groups to support collaborative ideation while giving each participant space to contribute, following prior accessibility co-design studies [16, 67]. Each group worked for 30 minutes on a different design scenario involving a particular robot embodiment, household task, and oversight context (see Appendix B.2 for the full scenario list). We derived the scenarios from our observational study, where participants reflected how their oversight strategies and challenges would change across different contexts. Each group designed ideas on how BLV people can effectively oversee robots in their assigned scenario. Participants could use the 3D-printed figures or role-play interactions themselves to demonstrate spatial relationships, robot behaviors, and proposed interactions. Finally, groups reconvened for a 30-minute share-out in which each group presented its design and discussed it with the broader workshop.

The study was approved by our institution's IRB, and all participants provided informed consent. We video-recorded the workshop and collected the groups' design artifacts. We qualitatively coded the recordings and artifacts and iteratively grouped related ideas across groups into recurring design patterns.

Figure 7: Co-design workshop materials and activities: (Left) Participants discussing in small groups. (Right) A participant demonstrating a scenario in which the robot picks up and hands over his mug.

4.2 Patterns Across Group Designs

Table 2 summarizes the recurring design patterns across groups. Because groups worked on different scenarios, counts indicate recurrence across designs rather than participant prevalence. Complete group designs are provided in § B.2.

Robot Execution: How robot actions could be easier to follow and interrupt. Participants wanted the robot to perform the task in a way that users could more easily predict and follow. For example, G2 suggested robots to search a tabletop in a consistent grid pattern, similar to systematic search strategies blind people are taught to use, so users could infer which areas had already been checked by the robot. G4 similarly wanted cleaning routines to begin from the same location and proceed in a consistent order, making progress easier to follow. For intervention checkpoints, 3 groups introduced different moments for users to inspect or influence execution. G3 and G6 designed robots that proactively pause and seek confirmation before consequential actions such as discarding an uncertain object.

Robot Communication: What users wanted to know during and after the task. Participants wanted information that clarified what the robot was doing, and how its actions changed the environment. 6 groups wanted updates about the robot's current activity, progress, remaining work, or estimated completion. For example, G4 proposed periodic progress updates, and G7 wanted the robot to communicate how much longer its task would take. 5 groups also wanted information about task outcomes or changes to the environment. G4 proposed a concise summary of completed work, and G7 wanted the robot to report additional changes it made to the shared space during the task (*e.g.,* moving furniture or objects).

Robot Communication: How robot updates could be conveyed. Groups proposed multiple modalities for receiving robot updates. 6 groups used language-based communication, ranging from proactive narration and task summaries to on-demand Q&A. Some also had the robot initiate questions when user input was needed (*e.g.,* confirming consequential action). The amount and timing of speech varied: G7 proposed a full description when the robot first encountered the user, then changing to ambient updates and on-demand questions as the user became familiar with its behavior.

All 7 groups also incorporated auditory or haptic cues, often for information that could be conveyed without demanding sustained attention. G3 proposed distinct sounds for movement along different axes (*i.e.,* moving vertically versus forward and backward). When users' auditory attention was occupied, groups shifted toward haptics: G5 proposed subtle vibrations while the user was in conversation, with stronger vibration patterns signaling greater urgency. Participants often paired these cues with language so that sounds or vibrations provide timely awareness and narration clarify the detailed outcome, or reason attention was needed.

Verification Support: How users could verify robot claims and outcome. 2 groups wanted the robot to provide *evidence* that users could use to assess the robots' claims. In G1, where the robot was refilling a guide dog's water bowl, participants proposed reporting sensed changes in water level or bowl weight instead of only stating that the bowl had been filled. In G2, where a robot arm was loading dishes, participants wanted it to report how many dishes it had handled and its confidence that the task was complete, helping users judge whether to further check. P12 in G2 explained *"The robot should tell me sometimes: I'm pretty certain, no need [for you] to check, or I'm not sure so you should come and check."* 2 groups also designed support to help users verify outcomes. G4 proposed bringing questionable objects to the user for confirmation or placing them in a consistent location for later tactile inspection, reducing the user burden to search the workspace.

Adaptation & Control: How oversight changed across contexts and preferences. 4 groups wanted oversight to adapt based on the user's situation, including their attention, activity, proximity, time of day, and whether other people were present. Unlike sighted users who can often maintain awareness through brief visual glances or listen to robot communication while visually focusing on another task, BLV users rely on speech and auditory cues that compete with other auditory attention (*e.g.,* screen reader, conversation). Therefore, participants wanted oversight to adapt its modality and intensity to their current context. For example, G5 proposed subtle phone or wearable notifications while the user was in conversation,

Design Dimension	Recurring Pattern	Description	Groups
Robot Execution	**Predictable execution**	Designs used consistent movement, search, or action sequences that users could anticipate and follow.	2/7
	Intervention checkpoints	Designs paused at meaningful points so users could confirm, correct, or otherwise intervene before execution continued.	3/7
Robot Communication (Information)	**Action & progress**	Designs communicated what the robot was doing, task progress, remaining work, or estimated completion.	6/7
	Outcomes & changes	Designs communicated completed work and changes the robot made to objects or the surrounding environment.	5/7
Robot Communication (Modality)	**Narration**	Designs used spoken or written descriptions, updates, summaries, or question answering to convey task information.	6/7
	Auditory & haptic cues	Designs used sounds, spatial audio, vibration, or other non-speech cues to communicate movement, location, state, or urgency.	7/7
Verification Support	**Verification evidence**	Designs exposed sensed state, counts, confidence, before–after comparisons, or other evidence users could use to assess robot reports.	2/7
	Physical verification	Designs created opportunities for users to directly inspect objects or outcomes through touch, such as bringing uncertain objects to the user or keeping them in known locations.	2/7
Adaptation & Control	**Context adaptation**	Designs changed communication timing, detail, modality, or urgency based on attention, proximity, activity, or social context.	4/7
	User customization	Designs let users configure modalities, verbosity, update frequency, or other oversight preferences.	3/7
Practical Integration	**Familiar devices**	Designs incorporated accessible devices users already used, such as phones, watches, or speakers.	4/7
	Low learning burden	Designs reused familiar cues, devices, or conventions to reduce new interaction mappings or setup.	3/7

Table 2: Recurring patterns across the 7 groups' co-designed concepts. We organize patterns by robot execution, robot communication (information and modality), verification support, adaptation and control, and practical integration. Individual designs often combined patterns across multiple dimensions. Counts indicate the number of group designs that incorporated each pattern, rather than participant prevalence.

escalating to stronger cues for urgent events. G6 similarly proposed shifting from spoken updates to private phone or watch notifications at night or around others. 3 groups also proposed user control of how the robot communicated. G4 and G7 proposed controls over modality, alert frequency, volume, and verbosity, and G6 suggested remembering communication preferences across repeated tasks. P21 in G7 highlighted *"Not everyone wants conversational robots."*

Practical Integration: How oversight fit into familiar devices and practices. 4 groups incorporated devices participants already used, including phones, smart watches, speakers, and other accessible technologies. For example, G6 proposed routing robot updates through Alexa [8] rather than requiring a separate robot-specific sensor or speaker interface. P18 emphasized *"It [monitoring the robot] should work with existing techologies and not require "another piece" of equipment."* 3 groups also tried to minimize the amount of new interaction users would need to learn to interpret. G3 drew on familiar auditory conventions such as pedestrian signals for robot movement. G4 emphasized first explaining what different sounds and cues meant and then limiting the number of distinct cues so users could remember them.

5 Design Guidelines

Drawing across the oversight challenges observed in the first study and the strategies participants used and proposed across both studies, we distill 6 design guidelines for accessible oversight of autonomous household work. These guidelines aim to preserve the benefits of participants' existing non-visual oversight strategies while addressing the challenges that limited their effectiveness. They address how robot behavior itself can support oversight (DG1), how task information is communicated (DG2), how users can assess and verify robot work (DG3-DG4), and how oversight can adapt to context and fit existing practices and technologies (DG5-DG6).

DG1. Make robot execution easier to follow through predictable behavior and nonvisual cues. Study 1 showed that participants relied heavily on naturally available sounds to follow robot activity, but these cues could be incomplete or ambiguous (**C1**). In Study 2, participants proposed robots that followed recognizable execution patterns and mapped sounds to specific actions such as movement direction or gripper state. Household robot planners often prioritize objectives such as coverage, efficiency, and obstacle avoidance [30, 56], and often adapt execution to newly sensed objects [19, 25], which consequently vary their paths across executions. Beyond such optimization, prior HRI work has explored *legible motion*, where visual trajectory differences help observers infer robot intent [17]. However for non-visual oversight, these visual trajectory differences may provide little benefit. Participants instead valued *consistency*, suggesting repeatable search, cleaning, and placement orders that users can learn and anticipate. They also wanted robot behavior to build on strategies already familiar to BLV people, such as systematic spatial search [60]. This suggests designing robot task strategies around BLV users' established practices, extending ability-based design's emphasis on building on users' existing abilities [73]. Beyond changing robot behavior, robots can preserve and strengthen informative natural sounds. As motor and contact sounds can be dampened by carpet or masked by environmental noise, robots can amplify them or add sound effects for key actions (*e.g.,* movement start or gripper state), making execution easier to track without continuous narration.

DG2. Proactively narrate task state and environmental changes users may miss. Participants could miss task events that were not conveyed through available nonvisual cues (**C1**), and user-initiated questions depended on knowing that something had happened or that information was missing (**C3**). Recent work demonstrates that robots can transform robot task trajectories into natural-language narration [43, 71], and research in explainable robots reveal what sighted users want to know about execution details and environmental state [41, 68, 69]. However, non-visual oversight may require more task and environmental state to be communicated explicitly, since users cannot rely on visual observation to fill in gaps. Rather than narrating everything, robots can model what the user can already perceive and proactively surface consequential information that is otherwise difficult to infer or discover. In practice, designers can prioritize events that are *(1) hard to perceive nonvisually, (2) consequential for task success and safety, and (3) unlikely to prompt a question because the user may not know they occurred.* Finer execution details can remain available through Q&A.

DG3. Ground robot communication in evidence and uncertainty to indicate when verification is needed. Both robots' own narrations and external AI descriptions can be incomplete, inaccurate, or delayed (**C2**), creating another source of uncertainty beyond the robot's physical execution [22, 52]. For nonvisual oversight, such errors may be difficult to detect because users cannot always visually cross-check the robot's account against its behavior or the resulting environment. Building on prior work on communicating robot uncertainty and anticipated failures [27, 41], robots should ground important claims in evidence and expose relevant uncertainty or sensing limits so users can determine *what still needs verification*. For example, rather than simply reporting that all dishes were loaded, a robot can report how many it handled and which parts of the workspace remained uncertain, helping users target further checks and recover otherwise unnoticed errors.

DG4. Guide users toward safe, targeted, and low-effort hands-on verification. When AI descriptions remained uncertain (**C2**), participants often turned to touch, but tactile checking could be effortful or unsafe (**C1**). Prior work has shown how robots can guide BLV users to objects through spatial language, haptic guidance, and structured handovers [12, 29, 36, 37]. In delegated work, these mechanisms can instead support verification of robot work. After completing a task, a robot can guide the user to an outcome rather than requiring them to search the workspace themselves. Robots can make clear *where and when* touch is safe. While still working, they can indicate the area they are occupying and nearby regions that are safe to inspect. When closer checking is needed, they could pause or retract and guide the user directly to the relevant outcome. Robots should also warn users when hands-on checking is unsafe or likely to disturb the result (*e.g.,* hot cookware, sharp knife, or spillable containers). In such cases, robots could instead provide sensor evidence or targeted visual descriptions, and low-vision users could be supported through visual highlighting of relevant or hazardous regions [14, 40].

DG5. Adapt oversight to users' attention, sensory access, and context. Study 1 showed that access to robot activity decreased when users were farther from the robot or workspace (**C4**), and co-design participants emphasized that the same communication may not remain accessible across everyday situations. For BLV users, audio often supports both computer access and environmental awareness, so robot sounds or narration can compete with screen readers, environmental sounds, or conversations rather than simply adding another channel. Robots could therefore sense available perception channel of users at the moment and escalate across modalities (*e.g.,* haptics on mobile devices) when audio is occupied. Similarly, when users are far away from the robot, it can bring uncertain objects to users for verification. While most BLV participants wanted rich narration from robots, they also wanted it to adapt to shared household contexts. To avoid interrupting household members or revealing private task details aloud, users should retain control over modality and verbosity.

DG6. Build on familiar devices and consistent conventions. This guideline emerged primarily from co-design rather than a specific Study 1 challenge. While participants wanted rich narration, sound cues, and information about the broader environment, they did not want this support to require additional cameras, speakers, or robot interfaces. Instead, they proposed integrating robot updates into devices they already use. This reflects broader accessibility work emphasizing access through mainstream technologies, rather than requiring new devices that can introduce additional cost and social stigma [63, 64]. Prior work has already explored extending robot monitoring and control beyond the robot itself through phones, wearables, and multimodal alerts [38, 44, 72]. For BLV users, robot oversight can build on an existing accessibility ecosystem (*e.g.,* smartphone visual-assistance apps, smart speakers, and smart glasses). Participants wanted robot communication to reuse familiar audio conventions and cue meanings (*e.g.,* earcons [20]), reducing the need to learn a new vocabulary. At the same time, robot narration could use a distinct voice or sound identity to distinguish it from other audio delivered through the same devices.

6 Reflections and Future Work

Our studies surface methodological considerations for future accessible HRI research. First, choices of physical props that may seem inconsequential in visually oriented studies can substantially affect nonvisual interaction. In our study, we used real household objects but substituted plastic plates for glass for safety. Yet, participants noted that glass plates would produce different sounds, suggesting that visual similarity does not imply auditory similarity. Second, while our guidelines are grounded in both observed oversight challenges and co-designed patterns, they have not yet been validated through implemented systems. We plan to build and evaluate these designs with BLV users in their homes, where household layouts, lighting, and robot sensing limitations may further shape oversight. Also, most of our low-vision participants had limited usable vision, so future work can examine how oversight strategies differ across broader spectrum of low vision. Finally, we focused on oversight rather than how BLV users command or intervene with robots. They may rely less on visually grounded references (*e.g., "the red cup"*) when clarifying objects or actions, and remains an important direction for future work.

7 Conclusion

We explored how household robot oversight can support BLV people when visual observation cannot be assumed. Our observational study (N=10) demonstrated that people used sound, touch, and visual interpretation tools to follow robot activity and verify outcomes, but these strategies were often ambiguous, effortful, and insufficient. Our co-design workshop (N=17) indicated that accessible oversight requires robots to make their behavior easier to follow, their communication easier to verify, and both adaptable to users' context. We synthesize these findings into design guidelines for accessible oversight of autonomous household work. As household robots promise to reduce effort and expand access to everyday tasks, we hope this work helps ensure that these benefits do not depend on vision, enabling BLV people to delegate work while retaining awareness, safety, and agency.

References

[1] Last visited: 2025. https://www.bemyeyes.com/
[2] Last visited: 2025. https://aira.io/
[3] Last visited: 2025. https://www.bemyeyes.com/bme-ai/
[4] Last visited: 2026. https://www.meta.com/ai-glasses/
[5] Last visited: 2026. https://www.seeingai.com/
[6] Last visited: 2026. https://gemini.google/overview/gemini-live/
[7] Henny Admoni and Brian Scassellati. 2017. Social eye gaze in human-robot interaction: a review. *Journal of Human-Robot Interaction* 6, 1 (2017), 25–63.
[8] Amazon. 2026. Amazon Alexa. https://alexa.amazon.com/about. Accessed: 2026-09-18.
[9] Shiri Azenkot, Catherine Feng, and Maya Cakmak. 2016. Enabling building service robots to guide blind people a participatory design approach. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 3–10.
[10] Kim Baraka, Stephanie Rosenthal, and Manuela Veloso. 2016. Enhancing human understanding of a mobile robot's state and actions using expressive lights. In *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 652–657.
[11] Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, et al. 2010. Vizwiz: nearly real-time answers to visual questions. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. 333–342.
[12] Mayara Bonani, Raquel Oliveira, Filipa Correia, André Rodrigues, Tiago Guerreiro, and Ana Paiva. 2018. What my eyes can't see, a robot can show me: Exploring the collaboration between blind people and robots. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. 15–27.
[13] Ravi Teja Chadalavada, Henrik Andreasson, Robert Krug, and Achim J Lilienthal. 2015. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In *2015 European Conference on Mobile Robots (ECMR)*. IEEE, 1–6.
[14] Ruijia Chen, Junru Jiang, Pragati Maheshwary, Brianna R Cochran, and Yuhang Zhao. 2025. Visimark: Characterizing and augmenting landmarks for people with low vision in augmented reality to support indoor navigation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.
[15] Maitraye Das, Thomas Barlow McHugh, Anne Marie Piper, and Darren Gergle. 2022. Co11ab: augmenting accessibility in synchronous collaborative writing for people with vision impairments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
[16] Ujjaini Das, Shreya Kappala, Meng Chen, Mina Huh, and Amy Pavel. 2026. Co-Designing Multimodal Systems for Accessible Asynchronous Dance Instruction. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. 1–18.
[17] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 301–308.
[18] Jodi Forlizzi and Carl DiSalvo. 2006. Service robots in the domestic environment: a study of the roomba vacuum in the home. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 258–265.
[19] Caelan Reed Garrett, Chris Paxton, Tomás Lozano-Pérez, Leslie Pack Kaelbling, and Dieter Fox. 2020. Online replanning in belief space for partially observable task and motion problems. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5678–5684.
[20] Stavros Garzonis, Simon Jones, Tim Jay, and Eamonn O'Neill. 2009. Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1513–1522.
[21] Atmaraaj Gopal, Arihiro Yorita, Naoyuki Kubota, and Matthias Rätsch. 2025. Nmm-hri: Natural multimodal human-robot interaction with voice and deictic posture via large language model. (2025).
[22] Zhao Han and Holly Yanco. 2023. Communicating missing causal information to explain a robot's past behavior. *ACM Transactions on Human-Robot Interaction* 12, 1 (2023), 1–45.
[23] Bradley Hayes and Julie A Shah. 2017. Improving robot controller transparency through autonomous policy explanation. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*. 303–312.
[24] Lukas Hindemith, Christiane B Wiebel-Herboth, Heiko Wersing, Britta Wrede, and Anna-Lisa Vollmer. 2025. Improving HRI through robot architecture transparency. *International Journal of Social Robotics* 17, 12 (2025), 2981–3001.
[25] Kohei Honda, Ryo Yonetani, Mai Nishimura, and Tadashi Kozuno. 2024. When to replan? an adaptive replanning strategy for autonomous navigation using deep reinforcement learning. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 6650–6656.
[26] Shanee Honig and Tal Oron-Gilad. 2018. Understanding and resolving failures in human-robot interaction: Literature review and model development. *Frontiers in psychology* 9 (2018), 861.
[27] Julian Hough and David Schlangen. 2017. It's not what you do, it's how you do it: Grounding uncertainty for a simple robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 274–282.
[28] Chien-Ming Huang and Bilge Mutlu. 2013. Modeling and Evaluating Narrative Gestures for Humanlike Robots.. In *Robotics: Science and Systems*, Vol. 2.
[29] Felix Huppert, Gerold Hoelzl, and Matthias Kranz. 2021. GuideCopter-A precise drone-based haptic guidance interface for blind or visually impaired people. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–14.
[30] Mrinal Kalakrishnan, Sachin Chitta, Evangelos Theodorou, Peter Pastor, and Stefan Schaal. 2011. STOMP: Stochastic trajectory optimization for motion planning. In *2011 IEEE international conference on robotics and automation*. IEEE, 4569–4574.
[31] Rie Kamikubo, Seita Kayukawa, Yuka Kaniwa, Allan Wang, Hernisa Kacorri, Hironobu Takagi, and Chieko Asakawa. 2025. Beyond Omakase: Designing Shared Control for Navigation Robots with Blind People. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–17.
[32] Seita Kayukawa, Daisuke Sato, Masayuki Murata, Tatsuya Ishihara, Akihiro Kosugi, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2022. How users, facility managers, and bystanders perceive and accept a navigation robot for visually impaired people in public buildings. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 546–553.
[33] Charles C Kemp, Aaron Edsinger, Henry M Clever, and Blaine Matulevich. 2022. The design of stretch: A compact, lightweight mobile manipulator for indoor human environments. In *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 3150–3157.
[34] Laura Kunold, Nikolai Bock, and Astrid Rosenthal-von der Pütten. 2023. Not all robots are evaluated equally: the impact of morphological features on robots' assessment through capability attributions. *ACM Transactions on Human-Robot Interaction* 12, 1 (2023), 1–31.
[35] Masaki Kuribayashi, Kohei Uehara, Allan Wang, Shigeo Morishima, and Chieko Asakawa. 2025. Wanderguide: Indoor map-less robotic guide for exploration by blind people. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–21.
[36] Dorothea Langer, Franziska Legler, Pia Diekmann, André Dettmann, Sebastian Glende, and Angelika C Bullinger. 2024. Got It? Comparative Ergonomic Evaluation of Robotic Object Handover for Visually Impaired and Sighted Users. *Robotics* 13, 3 (2024), 43.
[37] Dorothea Langer, Franziska Legler, Philipp Kotsch, André Dettmann, and Angelika C Bullinger. 2022. I let go now! Towards a voice-user interface for handovers

between robots and users with full and impaired sight. *Robotics* 11, 5 (2022), 112.
[38] Wing-ting Law, Ki-sing Li, Kam-wah Fan, Wang-hon Ko, Tiande Mo, and Chi-kin Poon. 2022. Two-way human-robot interaction in 5G tele-operation. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1198–1199.
[39] Cheuk Yin Phipson Lee, Zhuohao Zhang, Jaylin Herskovitz, JooYoung Seo, and Anhong Guo. 2022. Collabally: Accessible collaboration awareness in document editing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.
[40] Jaewook Lee, Andrew D Tjahjadi, Jiho Kim, Junpu Yu, Minji Park, Jiawen Zhang, Jon E Froehlich, Yapeng Tian, and Yuhang Zhao. 2024. Cookar: Affordance augmentations in wearable ar to support kitchen tool interactions for people with low vision. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–16.
[41] Gregory LeMasurier, Alvika Gautam, Zhao Han, Jacob W Crandall, and Holly A Yanco. 2024. Reactive or proactive? how robots should explain failures. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 413–422.
[42] Shuijing Liu, Aamir Hasan, Kaiwen Hong, Runxuan Wang, Peixin Chang, Zachary Mizrachi, Justin Lin, D Livingston McPherson, Wendy A Rogers, and Katherine Driggs-Campbell. 2024. Dragon: A dialogue-based robot for assistive navigation with visual language grounding. *IEEE Robotics and Automation Letters* 9, 4 (2024), 3712–3719.
[43] Zeyi Liu, Arpit Bahety, and Shuran Song. 2023. Reflect: Summarizing robot experiences for failure explanation and correction. *arXiv preprint arXiv:2306.15724* (2023).
[44] Muhammad Akmal Bin Mohammed Zaffir and Takahiro Wada. 2024. Presentation of robot-intended handover position using vibrotactile interface during robot-to-human handover task. In *Proceedings of the 2024 ACM/IEEE international conference on human-robot interaction*. 492–500.
[45] Cecily Morrison, Ed Cutrell, Martin Grayson, Geert Roumen, Rita Faia Marques, Anja Thieme, Alex Taylor, and Abigail Sellen. 2021. PeopleLens. *Interactions* 28, 3 (2021), 10–13.
[46] Bilge Mutlu, Toshiyuki Shiwa, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. 61–68.
[47] Amal Nanavati, Vinitha Ranganeni, and Maya Cakmak. 2023. Physically assistive robots: A systematic review of mobile and manipulator robots that physically assist people with disabilities. *Annual Review of Control, Robotics, and Autonomous Systems* 7 (2023).
[48] Andreas Naoum, Parag Khanna, Elmira Yadollahi, Mårten Björkman, and Christian Smith. 2025. Adapting robot's explanation for failures based on observed human behavior in human-robot collaboration. *arXiv preprint arXiv:2504.09717* (2025).
[49] Nnamdi Nwagwu, Adeline Schneider, Ibrahim Syed, Brian J Zhang, and Naomi T Fitter. 2024. The benefits of sound resound: an in-person replication of the ability of character-like robot sound to improve perceived social warmth. (2024).
[50] OpenAI. 2026. ChatGPT (GPT-4o version). Large language model. https://chatgpt.com Accessed: September 19, 2026.
[51] Yi-Hao Peng, Jason Wu, Jeffrey Bigham, and Amy Pavel. 2022. Diffscriber: Describing visual design changes to support mixed-ability collaborative presentation authoring. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–13.
[52] Pradip Pramanick, Luca Raggioli, Alessandra Rossi, and Silvia Rossi. 2024. Effects of Incoherence in Multimodal Explanations of Robot Failures. In *Companion Proceedings of the 26th International Conference on Multimodal Interaction*. 6–10.
[53] Malak Qbilat, Ana Iglesias, and Tony Belpaeme. 2021. A proposal of accessibility guidelines for human-robot interaction. *Electronics* 10, 5 (2021), 561.
[54] Adil Rahman, Md Aashikur Rahman Azim, and Seongkook Heo. 2023. Take my hand: Automated hand-based spatial guidance for the visually impaired. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
[55] Vinitha Ranganeni, Mike Sinclair, Eyal Ofek, Amos Miller, Jonathan Campbell, Andrey Kolobov, and Edward Cutrell. 2023. Exploring levels of control for a navigation assistant for blind travelers. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 4–12.
[56] Nathan Ratliff, Matt Zucker, J Andrew Bagnell, and Siddhartha Srinivasa. 2009. CHOMP: Gradient optimization techniques for efficient motion planning. In *2009 IEEE international conference on robotics and automation*. IEEE, 489–494.
[57] Laurel D Riek. 2012. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of human-robot interaction* 1, 1 (2012), 119–136.
[58] RoboDK. 2026. Franka Emika Panda Robot. https://robodk.com/robot/Franka/Emika-Panda. Accessed: 2026-01-07.
[59] PAL Robotics. 2026. TIAGo – Mobile Manipulator Robot for Research. https://pal-robotics.com/robot/tiago/. Accessed: 2026-01-07.
[60] Victor R Schinazi, Tyler Thrash, and Daniel-Robert Chebat. 2016. Spatial navigation by congenitally blind individuals. *Wiley Interdisciplinary Reviews: Cognitive Science* 7, 1 (2016), 37–58.
[61] Eike Schneiders, Anne Marie Kanstrup, Jesper Kjeldskov, and Mikael B Skov. 2021. Domestic robots and the dream of automation: Understanding human interaction and intervention. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.
[62] Svenja Y Schött, Rifat Mehreen Amin, and Andreas Butz. 2023. A literature survey of how to convey transparency in co-located human–robot interaction. *Multimodal Technologies and Interaction* 7, 3 (2023), 25.
[63] Kristen Shinohara, Cynthia L Bennett, Wanda Pratt, and Jacob O Wobbrock. 2018. Tenets for social accessibility: Towards humanizing disabled people in design. *ACM Transactions on Accessible Computing (TACCESS)* 11, 1 (2018), 1–31.
[64] 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*. 705–714.
[65] Yoshiaki Shiokawa, Winnie Chen, Aditya Shekhar Nittala, Jason Alexander, and Adwait Sharma. 2025. Beyond vacuuming: How can we exploit domestic robots' idle time?. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–17.
[66] Moondeep C Shrestha, Ayano Kobayashi, Tomoya Onishi, Erika Uno, Hayato Yanagawa, Yuta Yokoyama, Mitsuhiro Kamezaki, Alexander Schmitz, and Shigeki Sugano. 2016. Intent communication in navigation through the use of light and screen indicators. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 523–524.
[67] Stephanie Valencia, Michal Luria, Amy Pavel, Jeffrey P Bigham, and Henny Admoni. 2021. Co-designing socially assistive sidekicks for motion-based aac. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 24–33.
[68] Lennart Wachowiak, Andrew Coles, Gerard Canal, and Oya Celiktutan. 2025. What Questions Should Robots Be Able to Answer? A Dataset of User Questions for Explainable Robotics. *arXiv preprint arXiv:2510.16435* (2025).
[69] Lennart Wachowiak, Andrew Fenn, Haris Kamran, Andrew Coles, Oya Celiktutan, and Gerard Canal. 2024. When do people want an explanation from a robot?. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 752–761.
[70] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating robot motion intent with augmented reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 316–324.
[71] Zihan Wang, Brian Liang, Varad Dhat, Zander Brumbaugh, Nick Walker, Ranjay Krishna, and Maya Cakmak. 2024. I can tell what i am doing: Toward real-world natural language grounding of robot experiences. *arXiv preprint arXiv:2411.12960* (2024).
[72] Fabian C Weigend, Xiao Liu, Shubham Sonawani, Neelesh Kumar, Venugopal Vasudevan, and Heni Ben Amor. 2024. iRoCo: intuitive robot control from anywhere using a smartwatch. In *2024 IEEE international conference on robotics and automation (ICRA)*. IEEE, 17800–17806.
[73] Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)* 3, 3 (2011), 1–27.

A In-Person Observational Study

We outline the stages of our observational studies in Figure 8 and share details about our participants and the household tasks completed by the robots during the study.

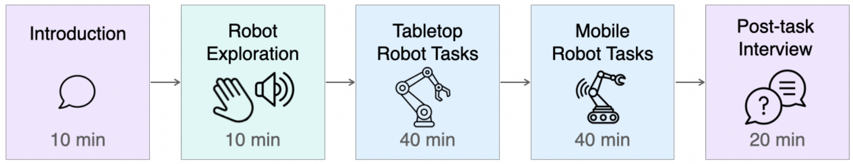

Figure 8: Stages of our observational study. The Tabletop robot and mobile robot each contained 2 tasks, with a survey and debrief for each task. The order of the robot types (*i.e.*, which robot was presented first) was randomized per participant.

A.1 Participants

We recruited 10 blind participants (P1–P10 in Table 3) through a local chapter of the National Federation of the Blind (NFB) and word of mouth. Eligibility included being able to travel to the study location in person. We ensured accessible commute support with local access door-to-door services.

PID	Gender	Age	Vision	Onset
P1	Female	38	Totally Blind	Congenital
P2	Male	53	Totally Blind	Acquired
P3	Female	38	Totally Blind	Congenital
P4	Female	40	Totally Blind	Congenital
P5	Male	47	Totally Blind	Congenital
P6	Male	46	Totally Blind	Congenital
P7	Female	44	Totally Blind	Congenital
P8	Female	37	Totally Blind	Acquired
P9	Female	30	Totally Blind	Congenital
P10	Male	41	Totally Blind	Acquired

Table 3: Demographics of blind participants in the in-person observational study

A.2 4 Household Tasks

We designed 4 everyday household tasks: two with a tabletop manipulator and two with a mobile manipulator. Each was presented with a short scenario to provide participants with a concrete goal and expectations for the outcome. During the study, a researcher teleoperated the robot following an execution script for each task. We also introduced 2 planned errors at specific moments for each task, to explore how blind participants identify and react to these errors.

Task 1: Set the Table

Scenario. Your friend is coming over, and you want to set the table with matching plates and cups. You ask the robot to help place the items neatly on the table. As the task runs, pay attention to what the robot seems to be doing and whether the final setup matches your expectations.

Steps.

(1) Two dishes (pink, teal) and two cups (pink, teal) are on the sink on the table; nothing else is on the table.
(2) The robot picks up a pink plate and places it on the table.
(3) The robot picks up a teal cup and places it next to the pink plate.
(4) The robot picks up a teal plate and places it on the table (it contacts the teal cup during placement).
(5) The robot picks up a pink cup and places it next to the teal plate.

Injected errors.

Robot	Task	Task description	Error type	Error description
Panda	**Set table**	From a sink-side staging area, the robot places two plates and two cups onto the table, arranging one cup next to each plate. The intended outcome is two color-matched plate–cup settings.	Misinterpretation	**Color mismatch + layout:** The robot forms two place settings with mismatched colors (*e.g.,* pink plate is placed next to teal cup).
			Control error	**Cup collision + risky grasp:** While placing the second plate, the robot bumps a nearby cup; the grasp appears marginal (held at the edge), raising safety concerns.
Panda	**Clear table**	Given mixed items on a cluttered table, the robot clears items by placing reusable dishware into a sink and disposing of trash into a bin. One small item remains on the table at the end.	Planning error	**Mis-sorted reusable:** The robot places a reusable cup into the trash bin instead of in the sink.
			Sensing error	**Missed paper towel:** The robot fails to notice a crumpled paper towel (outside its field of view) and leaves it on the table.
Tiago	**Fetch salt**	The robot navigates to a counter with seasonings, selects an item, returns to the user at the kitchen table, and places the item near the user's workspace.	Sensing error	**Wrong item retrieved:** The robot retrieves the wrong seasoning (lemon pepper instead of salt), likely due to label orientation.
			Control error	**Missed grasp + retries:** The robot must try multiple times to grasp the object since initial attempts fail, delaying the task.
Tiago	**Pour pasta**	The robot retrieves a pasta jar from a counter, returns to the user, tilts the jar to pour pasta into a pot, and sets the jar down afterward.	Control error	**Spillage during pour:** Some pasta spills onto the table or floor during pouring.
			Sensing error	**Residual pasta left:** The robot stops pouring before the jar is empty, leaving pasta remaining inside.

Table 4: Summary of four teleoperated household-robot tasks and injected issues.

- Misinterpretation error: The robot places two place settings with mismatched colors (the teal cup is placed next to the pink plate).
- Control error: While placing the second plate, the robot knocks over the nearby teal cup and does not pick it up.

Task 2: Clear the Table

Scenario. After your friend leaves, the table is cluttered with cups, bowls, and snacks. You ask the robot to clear the table so you can use it for work. There is a sink and a trash bin on the table. The robot should put reusable items into the sink and trash into the bin. Notice what the robot moves, what it leaves behind, and how confident you feel about the outcome.

Steps.

(1) Two plastic cups (pink, teal), one purple bowl, one empty Oreo package, one crumpled paper towel, and one leftover bread are laid out on the table.
(2) The robot picks up the pink cup and places it into the trash bin.
(3) The robot picks up the teal cup and places it into the sink.
(4) The robot picks up the bread and places it into the trash bin.
(5) The robot picks up the plastic bowl and places it into the trash bin.
(6) The robot picks up the Oreo package and places it into the trash bin.
(7) The crumpled paper towel remains on the table (out of the robot's field of view).

Injected errors.

- Planning error: The robot places the reusable pink cup into the trash bin (instead of the sink).
- Sensing error: The robot misses the crumpled paper towel and leaves it on the table.

Task 3: Fetch Salt

Scenario. You are sitting at the kitchen table while preparing pasta, and realize you need salt for boiling pasta. The robot will search for the salt, navigate back to you, and put it on the table. Notice what the robot is doing and what object it brings to the table.

Steps.

(1) A pot is on the table in front of the participant, and the robot starts near the participant's table.

(2) The robot navigates to the other side of the kitchen toward a shelf containing cereal boxes, fruit, a pasta jar, and salt and pepper. The labels on the salt and pepper are facing away from the robot.
(3) The robot attempts to grasp the pepper and makes multiple grasp attempts before succeeding.
(4) The robot grasps the pepper, navigates back to the participant's table, and places it on the table next to the pot.

Injected errors.

- Control error: The robot tries multiple times to grasp the object due to the approach angle and distance.
- Sensing/Planning error: The robot brings the wrong object (lemon pepper instead of salt).

Task 4: Pour Pasta

Scenario. You are preparing other ingredients and ask the robot to help by preparing pasta. The robot will search for a pasta jar, bring it to the table, and pour pasta into a pot. Pay attention to how the robot moves through space and how you understand what happens during pouring.

Steps.

(1) The robot navigates back to the kitchen shelf and picks up the pasta jar. The pasta jar has no lid.
(2) The robot grips the pasta jar upright and navigates back to the user.
(3) The robot adjusts its left arm joint to find a pouring angle.
(4) The robot pours pasta into the pot. Some of the pasta spills onto the table and onto the floor.
(5) The robot places the pasta jar on the table. There's remaining pasta in the jar.

Injected errors.

- Control error: The robot spills some pasta onto the table and/or floor.
- Early termination / sensing error: Some pasta remains in the jar after pouring.

Safety Measures. To support participant safety and to reliably introduce planned robot errors, we ran the sessions using a Wizard-of-Oz setup [57]. A researcher continuously supervised the robot and held the emergency stop button. We maintained a safe distance between the robot and participants throughout the study, and whenever participants wanted to touch the robot or inspect the workspace during execution, we paused the robot first to prevent collisions. We used a mix of real kitchenware (*e.g.,* cups, plates, sink) and mock props (*e.g.,* a trash bin). Before starting, we described the objects and the surrounding workspace and invited blind participants to tactually explore them (*e.g.,* materials, object locations, table boundaries) to build familiarity with the study setup. To improve manipulation reliability, we made minor task modifications (*e.g.,* using a pasta jar without a lid). To reduce spill risk and simplify cleanup during the pasta-pouring task, we used an empty pot with no boiling water.

B Co-Design Study

We outline the stages of our co-design workshop in Figure 9 and share details about our participants and each group's scenarios and solutions.

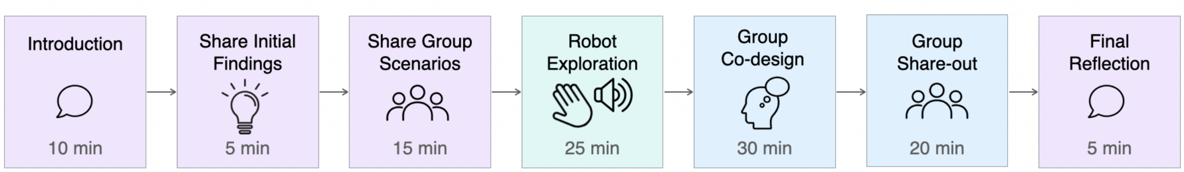

Figure 9: Co-design workshop stages. We first introduced the study and shared findings from the observational study, then presented group scenarios and facilitated hands-on robot exploration, followed by group co-design, a share-out, and final reflection.

B.1 Participants

We recruited 11 blind participants (P11–P21, Table 5) through a local chapter of the National Federation of the Blind (NFB) and word of mouth, and 6 sighted participants (S1–S6, Table 6) through an institutional mailing list. To minimize potential bias from prior exposure to the same task scenarios and planned errors, we did not recruit participants in our earlier in-person study.

Safety Measures. As with the observation study, a researcher continuously supervised the robot during the robot exploration portion of the co-design study and had immediate access to the emergency stop button. We maintained a safe distance between the robot and participants during this portion, and ensured that the robot was not running while participants touched or interacted with the robots physically.

PID	Gender	Age	Occupation	Vision	Onset	Robot Exp.	Types of Robots Used	Group
P11	Male	74	Retired	Totally blind	Acquired	2 years	Robot vacuum	1
P12	Male	34	Project manager	Totally blind	Congenital	4 years	Robot vacuum	2
P13	Male	50	Assistive technology specialist	Low vision	Congenital	12 years	Robot vacuum	3
P14	Male	63	Customer Service Representative	Totally blind	Congenital	4 years	Robot vacuum	3
P15	Female	52	Licensed Massage Therapist	Low vision	Acquired	3 years	Robot vacuum	4
P16	Female	38	Outdoor consultant	Totally blind	Congenital	None	N/A	4
P17	Female	40	Vocational Rehabilitation Counselor	Low Vision	Congenital	5 years	Robot vacuum, autonomous cars	5
P18	Female	34	Data Scientist	Totally blind	Acquired	4 years	Robot vacuum, autonomous cars	6
P19	Male	34	Product Manager	Legally blind	Congenital	2 years	Robot vacuum, autonomous cars	6
P20	Female	59	Workforce Development Manager	Totally blind	Acquired	4 years	Robot vacuum	7
P21	Female	47	Product Manager	Totally blind	Congenital	None	N/A	7

Table 5: BLV participants for the in-person co-design workshop.

PID	Gender	Age	Occupation	Vision	Onset	Robot Exp.	Types of Robots Used	Group
S1	Male	40	Musician	Sighted	N/A	None	N/A	1
S2	Female	*	OMS	Sighted	N/A	None	N/A	2
S3	Female	40	OMS	Sighted	N/A	Limited	N/A	3
S4	Male	29	Haptic engineer	Sighted	N/A	10 years	Research robots, wearables	5
S5	Male	43	Musician	Sighted	N/A	8 years	Robot vacuum	6
S6	Non-binary	50	Audio describer	Sighted	N/A	None	N/A	7

* Participant preferred not to disclose age.

Table 6: Sighted participants for the in-person co-design workshop.

B.2 Group Scenarios and Proposed Designs

We designed 7 scenarios, one for each design group. Two scenarios involved a mobile manipulator, two involved a tabletop robot arm, and three involved a humanoid robot.

Deliverables for each group: Design the end-to-end communication experience for your scenario. Show when the robot communicates, what it communicates, and how the user can respond or verify what happened. During the final share-out, demonstrate your design using role-play, the robot/3D figures, props, sketches, or another simple prototype.

Group 1: You are away from home while the robot completes a task
Robot: Tabletop robot arm
Scenario: You are away from home and ask the robot to refill your guide dog's water bowl. The robot reports that it has finished the task successfully.

- How would you want to monitor the task remotely without watching it continuously?
- How should the robot help you verify that the bowl was actually filled?
- What evidence would be useful? For example, a verbal description, status updates, a photo/video that is described to you, sensor information, or something else?
- Design the communication from request → task progress → verification of completion.

Co-designed solution: Group 1 proposed integrating the robot with a "smart" bowl equipped with sensors to detect the water level and potentially using the robot's camera to capture evidence of the bowl transitioning from empty to full. The system would communicate task completion through notification sounds on the user's phone or existing devices, providing sensor or visual evidence for the user to independently verify the outcome.

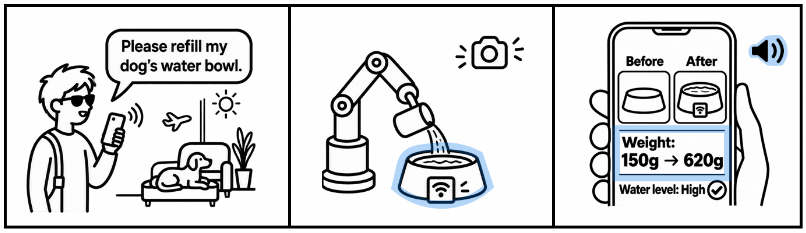

Figure 10: Group 1. Remote Monitoring Using smart objects with cameras and weight sensors to provide accessible evidence for remote task verification.

Group 2: You need to take over an incomplete robot task
Robot: Tabletop robot arm
Scenario: You ask the robot to load dirty dishes into the dishwasher. It loads most of them, but its camera fails to detect two dishes on the table. The robot stops, and you need to finish the task yourself.

- How should you know what the robot completed and what remains?
- How should the robot help you locate the two remaining dishes?
- What information would help you smoothly take over and finish the task?
- Design the handoff from robot → you.

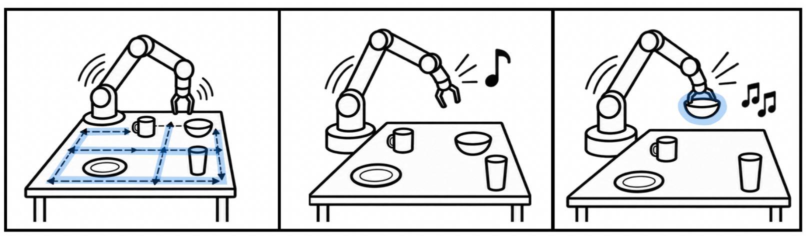

Figure 11: Group 2. Understanding Progress Using predictable search patterns and distinct movement sounds to make robot progress easier to understand.

Co-designed solution: Group 2 proposed that the robot follow a predictable grid pattern when searching, mirroring existing strategies used by BLV people to locate objects. The robot would also communicate its location through different audio cues (*e.g.*, specific sounds for when the robot is stuck, has retrieved an object, or is starting the task). The system would also report how many dishes it had loaded and communicate its confidence in task completion, allowing the user to infer what remains and locate the missed objects.

Group 3: The robot is about to take an action you may not want
Robot: Mobile manipulator
Scenario: You ask the robot to organize the kitchen counter. It sees a plastic spoon, decides that it is trash, and throws it away—but you wanted to keep it.

- At what point should the robot have communicated with you?
- Which actions can it take independently, and which should require confirmation first?
- How should the robot communicate uncertainty such as, "I think this is trash"?
- Design the interaction from planning → asking/confirming → acting → reporting what it did.

Co-designed solution: Group 3 proposed that the robot communicate object identities before acting. While simple movements could proceed autonomously, consequential actions such as picking up or discarding an object would require confirmation. The robot would also use different sounds to convey movement along different axes or planes.

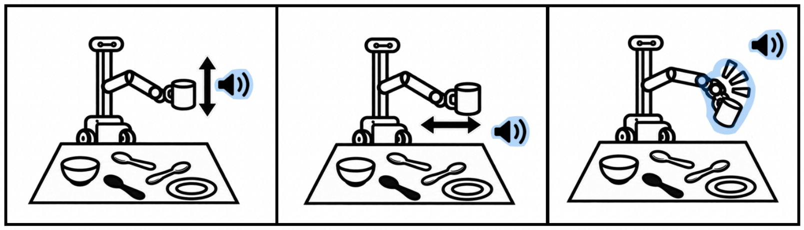

Figure 12: Group 3-1. Sound Cue Mappings Using different sounds to convey movement along different axes or planes.

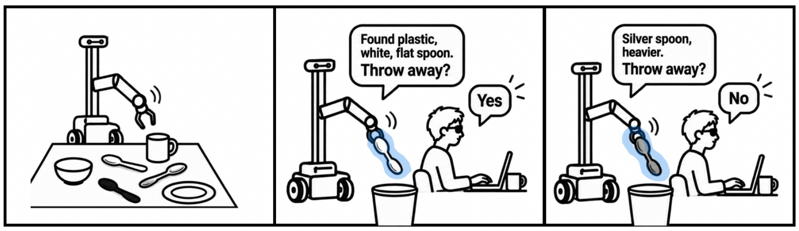

Figure 13: Group 3-2. Object Confirmation Describing visual and physical object characteristics before consequential actions.

Group 4: You are doing another task while the robot works
Robot: Mobile manipulator
Scenario: You ask the robot to clean the living room while you study in another room. The task takes about 20 minutes and involves several steps and small decisions.

- What information should interrupt you while you are studying? What information can wait until the task is complete?
- How should it communicate its progress?
- How should the robot distinguish between something that is urgent, needs a decision, or is simply a routine update?
- Design when and how the robot communicates while keeping interruptions minimal.

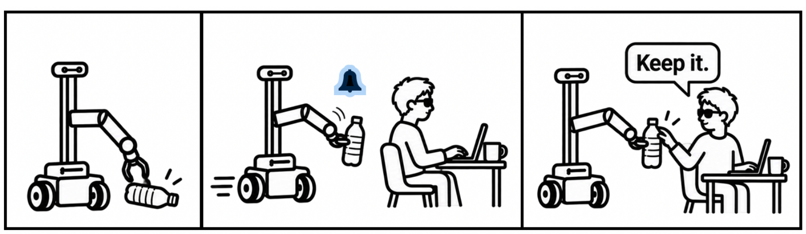

Figure 14: Group 4. Object Confirmation Bringing objects the robot is uncertain about, such as objects on the floor that may be trash, to the user for confirmation.

Co-designed solution: Group 4 proposed a customizable system where users select their preferred modalities (auditory, haptic, or verbal), alert frequency, and level of detail/verbosity. Routine progress could be conveyed through concise updates or phone notification haptics, while urgent situations would trigger specific "unpleasant" sounds. Based on user preferences, the robot may bring items to the user for confirmation, such as objects on the floor that may be trash, or keep such uncertain objects in a predetermined location for later inspection. The robot would provide an end-of-task summary, where verbosity can also be customized.

Group 5: You are socializing while the robot needs your attention
Robot: Humanoid robot
Scenario: You ask the robot to clean or organize the living room while a friend is visiting and you are having a conversation. During the task, the robot may need information from you or encounter something worth communicating.

- How should the robot get your attention without unnecessarily disrupting your conversation?
- Should its communication change because another person is present?
- How could you indicate "not now," "tell me privately," or "this is important—interrupt me"?

- Design how the robot enters, participates in, and exits the interaction.

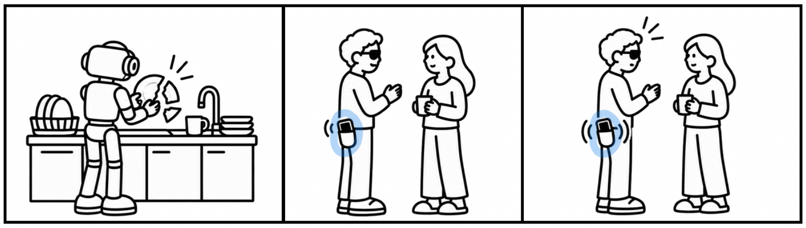

Figure 15: Group 5. Haptic Notifications Using minimally disruptive haptic notifications, where vibration frequency indicates task urgency.

Co-designed solution: Group 5 proposed a multi-level communication system that begins with minimally disruptive haptic notifications through a phone or wearable. The vibration frequency can correspond with the urgency of the task (higher frequency = more urgent). If the user does not respond to an urgent notification, the robot may escalate by approaching the user. Messages would be brief and convey the relevant location and urgency.

Group 6: You and another household member are present while the robot works
Robot: Humanoid robot
Scenario: You ask the robot to perform a household task while you and another household member are sharing the same space. The robot needs to give you progress updates, warnings, or ask you questions.

- How should the robot make it clear that a message is for you, rather than for everyone nearby?
- What information should be communicated aloud, privately, or only when necessary?
- How can the robot communicate enough for you to understand what is happening without becoming disruptive to the other person?
- Design how the robot communicates with you in this shared environment.

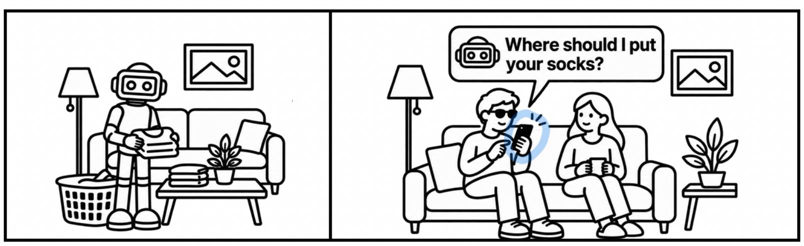

Figure 16: Group 6. Minimally Disruptive Notifications Using notifications sent to the user's phone for minimal interruptions, based on people present or time of day.

Co-designed solution: Group 6 created an example situation where one household member is sleeping while another asks the robot to retrieve juice. They proposed integrating robot communication with devices users already have, such as smartphones, smart speakers, and smartwatches, rather than introducing a new interface. The system would adapt its modality and timing based on context, such as the time of day and number of people present, while allowing users to configure communication preferences and respond through their existing devices. For example, if the robot is retrieving the user's juice at night while a household member is asleep, the robot would send questions to the user through phone notifications instead of asking them aloud.

Group 7: Someone else asks the robot to work in your shared space
Robot: Humanoid robot
Scenario: A sighted household member asks the robot to perform a task in a room that you are also using. You did not request the task, so you may not need detailed progress updates—but the robot's movements or actions could still affect you.

- What does the robot need to communicate to you even though you did not give the command?
- When should it tell you that it is approaching, moving objects, blocking a pathway, or changing the shared environment?
- What information would be unnecessary or annoying?
- Design how the robot maintains your awareness and safety without over-communicating.

Co-designed solution: Group 7 proposed that the robot announce its presence, current task, and expected completion time when it encounters the BLV user. The robot would clearly communicate changes to the shared environment, such as moving furniture or approaching

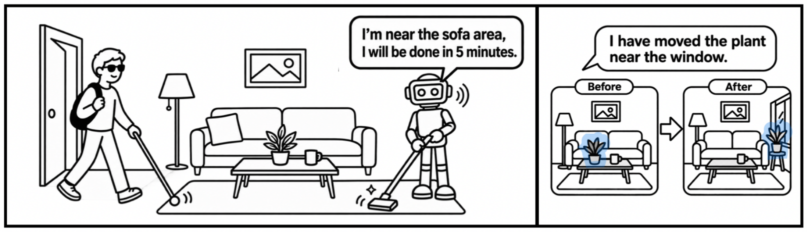

Figure 17: Group 7. Shared Space Awareness Announcing robot location, task duration, and environmental changes.

the user. The robot could also use ambient sounds and lights to indicate where it is, with more explicit warnings for urgent situations such as potential collisions. Users could customize the verbosity and frequency of task updates.