

In-context Q&A to Support Blind People Using Smartphones

André Rodrigues¹, Kyle Montague³, Hugo Nicolau², João Guerreiro⁴, Tiago Guerreiro¹

¹LaSIGE, Faculdade de Ciências, Universidade de Lisboa

²INESC-ID, Instituto Superior Técnico da Universidade de Lisboa

³Open Lab, Newcastle University

⁴Carnegie Mellon University

afrodrigues@fc.ul.pt, kyle.montague@newcastle.ac.uk, hman@inesc-id.pt,
jpvguerreiro@cmu.edu, tjvg@di.fc.ul.pt

ABSTRACT

Blind people face many barriers using smartphones. Still, previous research has been mostly restricted to non-visual gestural interaction, paying little attention to the deeper daily challenges of blind users. To bridge this gap, we conducted a series of workshops with 42 blind participants, uncovering application challenges across all levels of expertise, most of which could only be surpassed through a support network. We propose *Hint Me!*, a human-powered service that allows blind users to get in-app assistance by posing questions or browsing previously answered questions on a shared knowledge-base. We evaluated the perceived usefulness and acceptance of this approach with six blind people. Participants valued the ability to learn independently and anticipated a series of usages: labeling, layout and feature descriptions, bug workarounds, and learning to accomplish tasks. Creating or browsing questions depends on aspects like privacy, knowledge of respondents and response time, revealing the benefits of a hybrid approach.

CCS Concepts

•Human-centered computing → Empirical studies in accessibility; Accessibility systems and tools;

Keywords

Blind; Smartphone; Human Computation; Assistance

1. INTRODUCTION

Learning to use a smartphone device or application can be challenging for blind people, particularly when adopting these for the first time. Prior work has predominantly focused on improving physical touchscreen accessibility, particularly in the text-entry domain [7], either by measuring

performance [6], understanding gestures [5] or providing alternative interaction methods [9]. Rodrigues et al. [8] found that the barriers that blind people must overcome are not only related with the physical performance of touch gestures, but also in understanding the new paradigm that accompanies these devices and apps. For many, these challenges included being able to perform basic operations with the device, such as placing a call or sending a text message. Due to the lack of support for smartphones and apps, users found themselves resorting to friends and family to overcome these hurdles. However, the study results were drawn from assessing the adoption process of five blind people, putting forward the need for broader evaluations.

To better understand the extent of the challenges and current coping mechanisms, we conducted a series of workshops with 42 blind participants, from newcomers to expert smartphone users with different devices and operating systems. We found challenges that originate from the lack of guidance when using new features or layouts that are only surpassed with the support of others.

Prior research in crowdsourcing has allowed users to identify and locate objects in the real-world through visual questions using a smartphone (e.g. [1]). Following a similar approach, the Social Accessibility project [10] provided a collaborative metadata authoring mechanism to enhance webpage accessibility. Chilana et al. [4] have enabled Q&A in-context into web applications. Based on our workshop findings and inspired by these prior work, we developed *Hint Me!*, a human-powered service that allows blind users to get in-app smartphone assistance. Using *Hint Me!* as a design probe, we conducted a user study with six blind participants to elicit their perceptions on the usefulness and acceptance of human-powered networks for smartphone support.

In this paper, we extend the domain knowledge of the challenges blind users experience when interacting with smartphones. We propose *Hint Me!* as a solution to leverage the collective knowledge of others, and we share design considerations for future human-powered assistance technologies aimed at backing smartphone usage by blind people.

2. ISSUES AND COPING STRATEGIES

We recruited 42 blind people through social media, word of mouth, and through the centre for the blind where the workshops took place. Participants were required to be screenreader users. Over a two day period, we conducted

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ASSETS '17 Baltimore, Maryland USA

© 2017 ACM. ISBN ...\$15.00

DOI:

five workshop sessions for newcomers and novice users (i.e. three on Android and two on iOS), and two for experts (i.e. one Android and one in iOS). Each session lasted two to three hours and each participant attended only one. Although the workshops were conducted informally, newcomers and novices were guided throughout basic phone tasks, while experts sessions were centered around their questions and doubts. Participants were free to ask questions and collaborate during the sessions. Participant smartphone expertise ranged from newcomers (17), and novice (18) to expert (7). Newcomers were people who had never used a smartphone; novice users already owned a smartphone but were only able to do simple tasks, such as placing/receiving calls or send/read text messages; and, experts were able to use more advanced features, such as web services (e.g. Dropbox, Facebook). Our goal was to gain a broader understanding of smartphone barriers faced by blind users and their current learning and coping mechanisms. We extend prior work by including participants with different ability levels and devices. We gathered qualitative insights about initial reactions to smartphones; how participants currently use these devices; common and critical challenges; and how they cope with them. We conducted inductive thematic analysis [3] on researchers' notes of the workshops, which resulted in the following three main themes.

2.1 Challenges Beyond Touch Interaction

Participants thoroughly discussed issues they experienced with smartphones. For newcomers, their problems were related with touchscreen interactions and simple gestures. However, the cause of their struggles quickly became the lack of understanding on how the underlying interfaces were behaving. Although advanced gestures (e.g. 'L' gesture or rotor) were challenging for users of all expertise levels, it did not prevent them from accomplishing their goals. Moreover, some of their difficulties came from the lack of knowledge on how to perform the gesture rather than its execution.

All participants reported issues with smartphones, independently of expertise level and device. However, expert users focused more on application-specific issues, such as 1) unawareness of available options (*"In one app I had no way of sharing to Facebook. When I pressed More Actions nothing happened. What I found out afterwards, when I asked a friend, was that the option was there but it was not yet on the screen. I had to scroll on a new window that appeared."*), 2) using advanced features (*"I am not able to listen to music from my Dropbox in offline mode"*), or 3) accessibility problems (e.g. unlabeled buttons). Interestingly, all issues were solved with simple instructions.

2.2 Independent and Community Learners

Participants strongly rely on others to surpass challenges, often asking for help from people they consider to be technology experts. We found that users informally created communities that relied on the same specialist; two of them were present in our workshops. They were tech savvy, autodidact, and highly motivated to learn about technology. They regularly read blogs, forums, and mailing lists about assistive technologies, and even contact developers to report bugs and request features. Several participants in the workshop relied on them to cope with daily problems. They provided assistance through a variety of channels (e.g. calls, SMS, Skype) and often about the same issue but to different people.

During the workshops, experienced users would often help by guiding others step-by-step, while doing the actions on their own devices and waiting for others to finish each step. For gestural interaction, some participants went further and performed the gesture on the back of the other users' hand. Nevertheless, it was clear that people preferred an active learning approach rather than giving their device to others.

2.3 Issues in Surpassing Challenges

For some issues, the only possible solution was asking for help from a sighted friend (e.g. screenreader started speaking in a foreign language). However, participants discussed some situations where help from sighted friends and family was challenging due to their unfamiliarity with screenreaders. All but one participant mentioned how they preferred to be helped by screenreader users (*"Often the problem is not them [sighted users] not knowing how to solve the problem, the problem is not knowing how to explain to us how we can solve it"*). Although sighted people are seen as valuable sources of assistance, most of them are oblivious to the challenges of screenreader users. They usually know the steps needed to accomplish a given task, but are unaware on how to perform them using accessibility services.

Availability and over-reliance on others was recurrently mentioned. For specialists, it can become a burden in their daily lives, even though they enjoy helping others, as the wife of one of the specialists described: *"He helps everyone except me! He spends his evenings helping everyone, on the phone or on the computer, but has no time to talk with me."*

3. HINT ME! IN-APP MOBILE ASSISTANCE

The workshops extended prior work [8], revealing a frequent need for other people's assistance, despite expertise level. It is worth highlighting that people benefit from explanations given to the person next to them; this knowledge was 'contagious', spreading from one person to another, creating a collaborative learning experience. The exposure to similar doubts alongside the celerity at which people were able to learn together revealed both an opportunity and a need to enlarge the support networks beyond their current reach. Currently, users are limited to relying on others for help, or searching online for answers, which is cumbersome, takes the user out of the context the problem arouse in, relies on the user being able to portrait his issue, and often will not produce any result.

To augment current support networks, we propose *Hint Me!*, a human-powered in-app assistance tool for smartphone usage. It is an integrated Android service that enables users to connect with a support network of people willing to provide assistance. With *Hint Me!*, every question is linked with the context it was created in. Using an overlay button users may quickly browse existing questions and answers associated with the app, current screen, or with a particular element they select. Moreover, since it is an independent service it is available system-wide in any mobile application. *Hint Me!* supports, among others, workflow guidance (e.g., *how to perform an action?*), layout description, labeling, and learning workarounds for inaccessible content.

3.1 Creating a Question

Users are able to ask questions, within any application, through an ever-present quick launch overlay button (e.g. Figure 1 B). Questions are recorded and their text is ex-

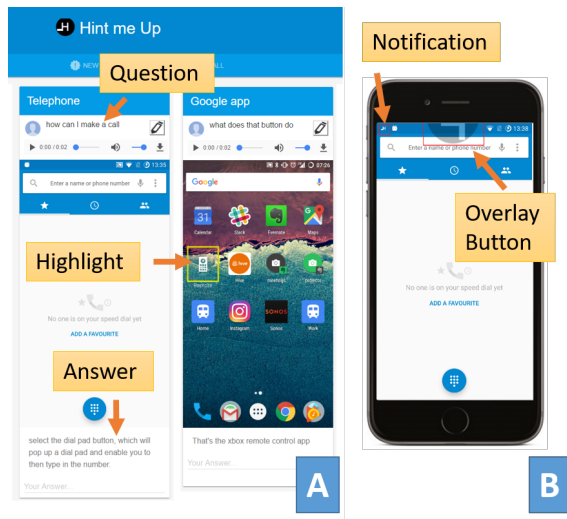


Figure 1: A) Volunteer web app. It shows two answered questions, one with a specific element of the interface highlighted. B) Hint Me! with the always available button on the top of the screen, and a notification showing the user he received an answer.

tracted relying on Google Voice Recognition technology. Additional information is collected to enable us to present the question to volunteers in diverse environments (e.g. web-app, Facebook, in-app context) augmented by, in-app contextual information aimed at providing volunteers with an enriched view over the applications and obstacles faced by the user (e.g. Figure 1 A). In order to provide context, *Hint Me!* automatically captures a screenshot of the device alongside all element’s details (e.g. alternate text, text, position, dimensions), creating the DOM tree structure of the interface and enabling its re-creation; in case the question relates to a specific element of the interface, users can select it in order to be highlighted.

3.2 Getting an Answer

Hint Me! gradually builds a shared knowledge-base with the answers to previously asked questions. Volunteers are able to edit the question for clarity or to correct errors from the speech conversion. Questions only become available when they have been validated and answered. When an answer is submitted, the question author receives a notification. Users can browse through all Q&A associated with their current context, or through their asked questions. Within their current context, Q&A will be filtered according to their current app or screen. Additionally, users can select a particular interface element to navigate content specifically associated to it. Users can select an answer to pin it to the *Hint Me!* overlay button. Long pressing the button accesses the answer, enabling users to follow long and hard answers without the need to memorize them.

4. PERCEPTIONS OF HINT ME!

We conducted a design probe study in an institution for visually impaired people where we recruited 6 participants, 3 females. Their ages were comprehended between 31 and 62

years old ($M=45.7$; $SD=12.6$). All had previous experience with smartphones: P2 a month, P1 and P3 a year, P4 two years, P5 and P6 over three years, all were legally blind and screenreader users. We used a Vodafone Smart Platinum 7 smartphone running Android 6.0 with Talkback. We relied on *Hint Me!* to allow users to explore the possibilities behind a human-powered assistant for smartphone usage in order to elicit comments and opinions for the interview thereafter.

The study was divided in two sessions, each lasting one hour: 1) posing questions; 2) browsing existing content. The content generated in the first session populated *Hint Me!* with Q&A derived from the users’ needs. One researcher acted as the volunteer through a web-app (Figure 1 A), the optimal scenario of a volunteer being an expert user. The researcher had previous knowledge of the tasks and was able to listen to the user interacting with the device. Although our focus was to understand the perceptions of the end-users receiving assistance, we also report on the volunteer experience. Each session started with a description of *Hint me!*, then participants performed four tasks, two in each of the selected apps (i.e. Facebook, WhatsApp, Du Speed Booster, and Spotify). Applications were counterbalanced between sessions and were selected from the top Play Store applications, discarding apps from the same category. Tasks were created based on the Play Store descriptions such as: check <John’s> friend profile; send <John> your location; release memory; and play music. Participants could resort to *Hint Me!* when they felt they could benefit from it. Each task started with the researcher reading aloud the task description; participants could, at any point during the task, prompt the researcher to repeat. Participants could only ask questions through *Hint Me!*. When a question was submitted, the researcher used the volunteer web-app to provide an answer. After the second session, we conducted a semi-structured interview to investigate the perceived usefulness and acceptance of *Hint Me!* and its underlying approach.

In the first session, a total of 21 questions were created (e.g. “How is the page organized?”, “How can I reach the artist since i cannot find him in the list?”, “Which button is the optimize?”), and each participant did at least two. In the 24 tasks of the second session, *Hint Me!* was opened 18 times and 16 answers were consulted, with all participants relying on them at least once. Two researchers inductively created a codebook from a set of three interviews. They coded independently and reached a Cohen’s Kappa agreement of $k=0.67$. Below, we detail our findings, anchored to the four main identified themes followed by the experience report from the researcher that acted as a volunteer.

4.1 From Aid to Self-Organized Learning

All participants reported positively about *Hint Me!* showing interest in installing the app in their devices. P1 and P4 felt that having direct answers to their very specific questions was the most useful feature, allowing them to surpass many of the barriers previously encountered. P2 stated: “For example, nowadays I don’t use the Internet on my phone. But, if I had access to *Hint Me!* I would have started using the Internet already. I am sure.” In this case, *Hint Me!* was seen as a safety net to explore new applications, knowing that he could always ask someone for assistance, if needed.

Hint Me! was seen as a learning tool that would give users autonomy to fully control their devices, as P4 stated:

"[with Hint Me!] we have greater autonomy in using the device because we are not dependent on others to tell us how something is done". Interestingly, P2 and P4 felt that Hint Me! allowed them to learn without the dependency on others. Although we explicitly told them that someone would be answering their questions online, these comments suggest that Hint Me! has the potential to reduce the social barriers associated with asking for help: "Sometimes people don't have the time to explain to us [how to do things]. If I had this service I wouldn't need to bother other people, I would just do them [the questions] here". P6 explicitly valued the active learning approach; that is, it is the users who perform the actions by learning and following a set of instructions: "I like this does not work as remote assistance, people have access to an image but can't control the device."

Participants identified several scenarios where the tool would be helpful. Four participants mentioned *Hint Me!* could be useful when exploring new apps or after an update. P1 stated *"I see myself using this app] mainly in an app that I am using for the first time, or maybe after an update, when new features are made available. Or there can even be a bug which already has workarounds available".* The system was also seen as a tool to report and deal with malfunctions or interface elements that had unexpected behaviors.

P6 saw *Hint Me!* as an in-app training tool, rather than a questioning app: *"It is useful to describe the app, it's structure and layout. It helps. A lot of blind people do not have a mental model [of the app] and can't do things easily - oh it's on the center of the screen or a little more to the right - they don't have that mental picture".* On the other hand, P2 focused on using *Hint Me!* to surpass accessibility problems, such as mislabeled or unlabeled buttons: *"I recently installed news apps and some of them are not accessible at all. With this app I could understand which button to press to get to certain sections".*

4.2 Questioning vs Browsing: A Trade-off

Participants identified value in both being able to create a new question and browsing previous stored knowledge. However, when asked about the foreseen usage of the system, they revealed different perceptions and preferences, namely in regards to the way they would retrieve knowledge.

P5 showed a preference for browsing and would only create questions if he couldn't find a response: *"I think I would check the database first. This way, I wouldn't risk making a question that was already asked. If I couldn't find it, I would then add one more question".* On the other hand, P6 considered the perceived availability of an answer to be a deciding factor: *"It is always easier to ask a question if the answer comes right away; if it is about the app's layout, I would search for an [existing] question, because that question was probably made, and it would be faster to search rather than ask a new question; if it is something that probably no one asked before, it's easier to ask."*

Other participants reinforced time of response as being relevant in their foreseen operation of the system. P4 stated *"having the list of questions is very relevant as there may not be people available to timely answer our questions".* Time was not the only reason for a browse first approach; other participants felt leaned to it as they had doubts about their ability to accurately formulate a question.

Participants expressed thoughtful concerns, namely to what relates to application versions. P1 told us: *"If the answers*

provided were for the same version, I would search for an available answer; if not, I would make a question since it is likely that the answer is no longer valid for my version".

4.3 Anonymity and Answer Quality

We asked participants about whom they would send their questions to, particularly between unknown volunteers, close friends, or their broad social network. Participants stated to prefer directing the questions to the volunteer group, choosing a bi-directional anonymous communication. The main reasons to support this were related to not overburdening their family and friends, but also related to the limited knowledge that this closer group may have (*"...in their case it would be harder to get the answer"*, P1). There is a common belief that the group of volunteers would be more qualified, both at the application and accessibility level (*"[closer] people are not aware of accessibility (i.e., Talkback), probably they will not be able to help much"*, P1). However, sharing their in-app information with unknown people was considered a possible issue, where additional contextual information is required, as stated by P5: *"I'm not sure what the screen capture shows. I think it would be important for us to understand how much of the screen is being captured."*

Asking questions to close people was considered useful when sensitive information was involved. P6 stated: *"If it had [personal] information, [...], I would be more comfortable asking someone I know. But if it was - what is this button in this app, where personal information is not shown, in this case I wouldn't mind asking a broader group of people."*

People disliked the possibility of having questions posted in their Facebook wall. In these cases, they do not feel it is private nor safe (*"Facebook is public. One thing is to ask an anonymous question to a specific group [...] A different thing is to post it on Facebook"*, P6). Similar findings were reported in Brady et. al. [2] where blind people were reluctant to use social networks for visual Q&A (i.e. a question associated with a photo) due to the perceived social cost.

When talking about performing questions to sighted or blind people, participants reinforced that volunteers should be experienced with accessibility services so that they could offer useful responses.

4.4 The Role of the Volunteer

The volunteer provided answers that accommodated different navigation methods (explore by touch or swiping) by providing both spatial and positional instructions (e.g. *"The Optimize button is on the center top of the screen; navigating from the start it is the second button without a label"*).

There is often a mismatch between the visual information and the output of the screenreader. Without using a screenreader or having additional contextual information, the volunteer would not have been able to answer all questions. One example was a mislabeled option when searching for an artist on Spotify, the first step involved opening a menu incorrectly labeled as *"Go Up"*. In some cases, even layouts with correct labels may not be enough to incite a clear answer (e.g. duplicate labels).

Without rich contextual information, sighted people will struggle to provide clear answers. On the other hand, there are questions that are only trivial to sighted volunteers (e.g. *"How is this page organized"*, *"What is this button?"*).

5. IMPLICATIONS FOR DESIGN

Participants showed interest in using *Hint Me!* to learn at their own pace, thus removing the need to rely on others' availability. Still, they expressed concerns on how to ensure the quality of the answers and their own privacy. They saw different usages for questioning and browsing, from addressing an issue to finding features or workarounds they were unaware of. The following implications derived from their desires and concerns and the insights on the volunteer role:

Enable Self-Organized Learning. Facilitating smartphone usage is not just about overcoming challenges; it is also about promoting serendipitous discovery of new features. Assistance must allow users to have control on content consumption in order to learn at their own pace.

Support the Workforce. For a successful assistance, the human supporter must be provided with enough information to become domain competent and aware of the communication needs of the end-user. As such, we should compensate the potential mismatch between the user's experience and the volunteer, augmenting his understanding of the user's context and doubts (e.g leveraging DOM trees to portrait the information available to the screenreader).

Gather Knowledge. Technologies that rely on human input should not waste contributions in single use, but instead iteratively build a shared knowledge-base. Moreover, we must look for opportunities to preemptively generate knowledge (e.g. describe layout structure) enabling better coverage and availability.

Nurture Knowledge. The variety of mobile devices, applications versions and frequent updates demands a continuous re-assessment of the gathered knowledge validity.

Respect Privacy. Smartphones are inherently private and hold personal data. Human-powered approaches must provide users with control over what they share and with whom, awareness of what is being shared, and selection of supporter-group based on information sensitivity. Alternatively, we must find novel ways to take advantage of context by removing all private and identifiable information.

6. LIMITATIONS

With the guarantee of the quality of the answers we were able to understand the potential of the approach. However, it limited our understanding of the issues the users face with answers of variable quality. We relied on a screenshot of the user interface to provide context, limiting the solution to sighted volunteers. However, we also collect DOM tree structure enabling the recreation of the contents accessible to the screenreader. With it, the pool of volunteers can be expand to users with a wide range of abilities.

7. CONCLUSIONS

Our findings show that support networks are essential to surpass everyday barriers. Although existent human-powered tools help blind users making sense of the real-world, supporting smartphone usage has been an unexplored avenue. We present *Hint Me!*, a human-powered assistance tool that mimics the qualities of in-person support and collaborative learning. Initial perceptions showed positive and promising results related to in-app support and self-organized learning. Future work will need to address privacy issues, leverage created knowledge, guarantee adequate assistance, and target a larger segment of users (e.g. older adults).

8. ACKNOWLEDGEMENTS

We thank the Fundação Raquel e Martin Sain in Lisbon. This work was partially supported by Fundação para a Ciência e a Tecnologia (FCT) through funding of the scholarship and ref. SFRH/BD/103935/2014, LaSIGE Research Unit, ref. UID/CEC/00408/2013.

9. REFERENCES

- [1] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, and T. Yeh. Vizwiz: Nearly real-time answers to visual questions. In *Proc. of ACM Symposium on User Interface Software and Technology*, UIST '10, pages 333–342, 2010.
- [2] E. L. Brady, Y. Zhong, M. R. Morris, and J. P. Bigham. Investigating the appropriateness of social network question asking as a resource for blind users. In *Proc. of Conference on Computer Supported Cooperative Work*, CSCW '13, pages 1225–1236, 2013.
- [3] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.
- [4] P. K. Chilana, A. J. Ko, and J. O. Wobbrock. Lemonaid: Selection-based crowdsourced contextual help for web applications. In *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1549–1558, 2012.
- [5] S. K. Kane, J. O. Wobbrock, and R. E. Ladner. Usable gestures for blind people: Understanding preference and performance. In *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 413–422, 2011.
- [6] H. Nicolau, K. Montague, T. Guerreiro, A. Rodrigues, and V. L. Hanson. Typing performance of blind users: An analysis of touch behaviors, learning effect, and in-situ usage. In *Proc. of ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '15, pages 273–280, 2015.
- [7] J. Oliveira, T. Guerreiro, H. Nicolau, J. Jorge, and D. Gonçalves. Blind people and mobile touch-based text-entry: Acknowledging the need for different flavors. In *Proc. of ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '11, pages 179–186, 2011.
- [8] A. Rodrigues, K. Montague, H. Nicolau, and T. Guerreiro. Getting smartphones to talkback: Understanding the smartphone adoption process of blind users. In *Proc. of ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '15, pages 23–32, 2015.
- [9] C. Southern, J. Clawson, B. Frey, G. Abowd, and M. Romero. An evaluation of brailletouch: Mobile touchscreen text entry for the visually impaired. In *Proc. of the International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 317–326, 2012.
- [10] H. Takagi, S. Kawanaka, M. Kobayashi, T. Itoh, and C. Asakawa. Social accessibility: Achieving accessibility through collaborative metadata authoring. In *Proc. of ACM SIGACCESS Conference on Computers and Accessibility*, Assets '08, pages 193–200, 2008.